

# ANÁLISIS DE EVALUABILIDAD

DICIEMBRE DE 2017

## ÍNDICE

ANÁLISIS DE PERTINENCIA DE EVALUACIÓN DIAGNÓSTICA CON LA METODOLOGÍA DE MARCO LÓGICO.....	3
PROGRAMA PRESUPUESTARIO COMO PROGRAMA PÚBLICO .....	6
CRITERIOS DE NO EVALUABILIDAD CON ESTAS HERRAMIENTAS .....	7
BIBLIOGRAFÍA.....	10

## **ANÁLISIS DE PERTINENCIA DE EVALUACIÓN DIAGNÓSTICA CON LA METODOLOGÍA DE MARCO LÓGICO**

La evaluabilidad de los programas y las intervenciones públicas es un tema que ha ganado centralidad en el debate internacional sobre evaluación y consolidación institucional de la gestión del desempeño. Como ejemplo, destaca la adopción del criterio de evaluabilidad como una de las cinco dimensiones imprescindibles de la evaluación profesional y de alta calidad, señalados en los Estándares de Evaluación para América Latina y el Caribe. Este documento, publicado en 2017, por la Red de seguimiento, evaluación y sistematización en América Latina y el Caribe (Relac), con el apoyo de Fomento de Capacidades en Evaluación (FOCEVAL) y el Instituto Alemán de Evaluación de la Cooperación para el Desarrollo (DEval), es producto de una consulta muy amplia en toda la región y refleja un consenso disciplinario emergente. El proyecto surge con el objetivo de “contribuir al desarrollo de un marco común de referencia que sirva de orientación para la obtención de evaluaciones con altos estándares de calidad, la formación y práctica profesional, facilitar la comunicación entre los actores participantes, el aprendizaje y generación de conocimientos, y el fomento de una cultura de evaluación y responsabilización social” (Rodríguez Bilella et al 2017:3) Las dimensiones que destaca el documento son las siguientes: 1) evaluación rigurosa; 2) adecuada evaluabilidad; 3) evaluación conducida con principios éticos y jurídicos; 4) adecuada comprensión cultural y 5) relevancia y utilidad.

La evaluabilidad se refiere a la valoración previa de condiciones básicas preexistentes que permitan suponer que la evaluación se puede gestionar efectivamente, que existen procedimientos prácticos que permiten realizar la valoración, que el contexto social y político hace viable la evaluación y que se puede hacer una valoración realista con la información disponible. Esto busca asegurarse que la evaluación y sus hallazgos sean relevantes y útiles para la mejora. En este contexto, un programa con una teoría de cambio no identificable o sin información básica sobre su operación no debería evaluarse porque en estas condiciones puede anticiparse que los hallazgos serán muy limitados, poco confiables y de baja utilidad para la mejora. En el caso de las evaluaciones que utilizan la metodología de marco lógico como herramienta central, es importante considerar criterios adicionales de evaluabilidad asociados al uso de esta metodología.

La gestión adecuada del desempeño gubernamental depende en gran medida de la utilización idónea de las diversas herramientas disponibles para el monitoreo y la evaluación de programas públicos. Por esta razón, es bien sabido que la utilidad de la Matriz de Indicadores para Resultados (MIR) depende del cumplimiento de una serie de condiciones previas. En primer lugar, que la MIR sea efectivamente el resultado de un ejercicio abierto y participativo de planeación estratégica como el que señala la metodología de marco lógico, en cualquiera de sus variantes (Aldunate y Córdoba, 2011), y que los participantes cuenten con los conocimientos técnicos necesarios para aplicar la metodología de manera correcta. En segundo lugar, que la información que documenta un programa y su orientación a resultados esté realmente apegada al proceso y la lógica real de la intervención pública y sea gestionada por los actores y los procesos en ella descritos. Finalmente, que el programa presupuestario, como categoría contable, corresponda al concepto de proyecto o programa público derivado del enfoque de política pública. Es decir, una intervención sistemática (generalmente producción de bienes y/o servicios) orientada a la atención de la población para resolver un problema público claramente definido, mediante una teoría de causal expresa, o un proyecto de implementación discreta (categórica) (ej. construcción de un puente, implementación de una institución u organización nueva).

Evidentemente, en países en proceso de adopción de la gestión orientada a resultados, como México, estas condiciones no suelen cumplirse cabalmente. Lo que se espera, en un modelo incremental de reformas orientadas a mejorar el desempeño gubernamental, es que los programas públicos se aproximen gradualmente a estas condiciones idóneas para poder reflejar con claridad y transparencia las lógicas, los modelos de gestión y los resultados para el ciudadano alcanzados por cada una de las intervenciones. Así, por ejemplo, a pesar de que muchas MIR no son producto de procesos participativos de planeación estratégica o suelen ser realizadas de manera unilateral por algún grupo con capacidad técnica heterogénea, los errores y problemas que de esto derivan se pueden identificar con facilidad mediante un ejercicio de validación de la matriz. La razón es simple, muchas inconsistencias de la MIR son clara consecuencia del proceso de construcción y la ausencia de un diagnóstico de calidad que cumpla con los estándares mínimos de validez y las especificaciones técnicas de los conceptos de la metodología. Por lo tanto, cuando las evaluaciones detectan estos problemas se detona una curva de aprendizaje que permite fortalecer la evaluabilidad de los programas, aunque en un primer momento los hallazgos de la evaluación en sí misma no sean muy sólidos por ese problema de origen. Igualmente, el monitoreo de los indicadores de gestión y las evaluaciones de procesos pueden identificar con relativa facilidad la distancia entre la norma y la

operación real de un programa, promoviendo la congruencia sustantiva y administrativa entre los documentos programáticos y la operación real. Finalmente, cuando se identifican problemas técnicos en el diseño y la concepción de la teoría del programa expresada en la MIR, o el nivel de objetivo al que corresponden cada uno de los indicadores, así como el cumplimiento de los criterios básicos de validez de los indicadores (criterios CREEMA), el ejercicio evaluativo permite subsanar esas fallas y apoyar a los operadores de los programas en el fortalecimiento de su orientación estratégica y la claridad y transparencia con las que se maneja la información del programa.

No obstante, la ruta de mejora a través de evaluaciones que fortalecen la evaluabilidad sin necesariamente estar en condiciones de medir el desempeño por problemas de origen en el planteamiento del programa son una ruta de muy alto costo para la mejora, porque sin condiciones básicas de evaluabilidad, la evaluación arrojará hallazgos sobre el grado de evaluabilidad y no resultados de evaluación directamente. Un caso particularmente importante de este problema se refiere a programas que no tienen una ruta de salida clara en la presentación de bienes y/o servicios para la población. ***En estas situaciones, por las razones que se expondrán más adelante, no solo no hay condiciones idóneas para evaluar utilizando las herramientas de la metodología de marco lógico, sino que incluso podría ser contraproducente, ya que se sometería a un programa presupuestario sin orientación clara a la solución de un problema público mediante la provisión de bienes y servicios directos a la población, a un proceso forzado de adaptación a unas herramientas para medir desempeño que no son las óptimas.***

En estos casos, las alternativas disponibles son las siguientes: la revisión del diseño y la clasificación de los programas presupuestarios, es decir, una revisión de la clasificación de estos programas en la estructura programática, o la adopción de otras herramientas para realizar la evaluación y el monitoreo. Estas otras herramientas incluyen la selección de indicadores de gestión y estratégicos para monitorear al programa presupuestario sin adoptar el paquete completo de la evaluación anclada en la MML; la adopción de estándares o *benchmarks* de índole institucional y/u organizacional o la generación de mecanismos mixtos para el monitoreo y la evaluación, como la adopción de indicadores de desempeño que capturen la dimensión interna y organizacional del programa presupuestario, sin presuponer que se están entregando bienes y servicios directamente a una población objetivo predeterminada.

## PROGRAMA PRESUPUESTARIO COMO PROGRAMA PÚBLICO

La metodología de marco lógico está pensada para monitorear y evaluar proyectos, programas e inversiones que proveen bienes y/o servicios con salida directa a la ciudadanía, para resolver un problema público previamente identificado y conceptualizado, en los términos de la propia metodología. Es decir, que el adecuado uso de la metodología supone que los programas presupuestarios entran claramente en la definición de programa público/intervención, que supone: una función de producción que convierte insumos públicos en productos a través de un proceso de actividades claramente concatenadas, y que se espera que dicho producto, al ser entregado a la población beneficiaria, sea capaz de generar procesos sociales específicos (los resultados esperados, comúnmente fraseados como la contribución a la reversión del problema que da origen a la intervención). Además de una conceptualización lógica, denominada cadenas de resultados, estos programas se distinguen de otras formas de gasto público por su vocación externa, orientada a la población. Por tanto, deben distinguirse de las actividades, acciones y funciones que se realizan como parte de la gestión interna de la administración pública como la coordinación, la modernización, el manejo patrimonial del Estado, la capacitación interna, la ampliación de oficinas gubernamentales y los procesos de compras y vinculación consustanciales a la operación del gobierno. Estas acciones, por supuesto, son evaluables y monitoreables, pero con las herramientas de valoración del desempeño organizacional y/o institucional, tales como el tablero de control y otras herramientas de la administración por objetivos. En este tipo de intervenciones, las de mantenimiento organizacional, no hay un vínculo directo con la implementación de un proyecto o la entrega de bienes y servicios a la población, por lo que muchos conceptos asociados a la MML, como los niveles Propósito y Fin, o la necesidad de identificar poblaciones potencial y objetivo, así como mantener un padrón de beneficiarios (elementos de la evaluación), resultan poco adecuados para capturar el desempeño de estos programas. Para la selección adecuada de estas alternativas, es necesario analizar la viabilidad de la revisión de la estructura programática, la revisión del diseño de estos programas o, en su defecto, la identificación de los indicadores que mejor permitan capturar su desempeño de acuerdo a la lógica de la intervención.<sup>1</sup> Cuando se busca evaluar este tipo de intervenciones como si fueran programas públicos directamente orientados a la población, el uso de la metodología es muy forzado, lo que puede generar distorsiones importantes en la lógica de la acción (por ejemplo, pedirle a una acción interna que resuelva un problema público sin que medie una conexión causal clara entre los

---

<sup>1</sup> Este análisis rebasa los alcances y propósitos de este documento.

productos y resultados) y/o forzar de manera incremental la adaptación artificial de una función interna a la lógica de un programa público (de *delivery*).

## **CRITERIOS DE NO EVALUABILIDAD CON ESTAS HERRAMIENTAS**

En seguimiento a los estándares de relevancia y utilidad planteados como estándares de calidad para la evaluación para América Latina y el Caribe (Rodríguez Bilella et al 2017), los criterios adoptados para clasificar a un programa presupuestario como no idóneamente evaluable con esta metodología reflejan un componente teórico y otro práctico. En relación con este último, se descartaron dos programas presupuestarios que fueron descartados en el Decreto de Presupuesto de Egresos por diversos motivos, por lo que no figurarán en el ejercicio presupuestal correspondiente a 2018. Tal es el caso del programa Vigilancia y Evaluación de la Auditoría Superior (Programa #128) y el Programa Impulso al Desarrollo Económico (Programa #186). Con independencia de los factores que pudieran haber influido en la decisión de no continuar con esos programas, no es ético ni idóneo realizar este tipo de evaluación (diagnóstica) en condiciones en las que claramente será imposible contribuir a la mejora o producir información pertinente para la toma de decisiones sobre estos programas.

El resto de los programas descartados para la evaluación diagnóstica corresponde a un subconjunto de 10 programas presupuestarios que, por sus características, no tienen una ruta de salida clara a la población y corresponden más propiamente a la gestión interna de la administración pública o a una serie de actividades que no están claramente vinculadas con la prestación de bienes y servicios entregados a la ciudadanía para promover un resultado social determinado. En estos casos, el grupo evaluador consideró contraproducente la aplicación de estas herramientas de evaluación diagnóstica por tratarse de un ejercicio forzado que puede tener dos efectos no deseados: la valoración excesivamente baja de estos programas, por la difícil adecuación de sus lógicas de gestión a la valoración de resultados, de acuerdo con la metodología de marco lógico y la naturaleza de los 12 reactivos de la evaluación diagnóstica, basados en dicha metodología. En estas condiciones, se considera que la aplicación del instrumento produciría una valoración muy baja del desempeño de los programas, pero que este resultado sería un artefacto del instrumento de evaluación aplicado y no necesariamente un reflejo adecuado del desempeño de dichas funciones. Igualmente, el tipo de recomendación y sugerencias derivados de la aplicación de dicha metodología conduciría a una

adaptación forzada, por tanto, poco útil y transparente, de las lógicas de estas actividades y funciones bajo el lenguaje y los conceptos derivados de la metodología. En estos casos, el ciclo virtuoso en el que la evaluación produce hallazgos tanto para fortalecer la evaluabilidad como para armonizar y fortalecer la orientación a resultados de un programa no se produce, por la incompatibilidad de las herramientas utilizadas y la naturaleza de la función, o implica la construcción de teoría causales sumamente complejas, con poca trazabilidad, para justificar la orientación externa y el resultado observable en la población, que en realidad están muy lejos de los impactos directos de dichas gestiones.

Los criterios antes expuestos para la selección se utilizaron como tipos ideales, y con un sesgo hacia la evaluabilidad. Así, por ejemplo, se incluyeron programas con al menos un componente orientado a la ciudadanía o aquellos dedicados a la implementación de procesos, instituciones u organizaciones nuevas, como si se tratase de proyectos discretos.

A continuación, se detalla la lista de programas clasificados como no evaluables de acuerdo con los criterios antes descritos.

NÚMERO DE PROGRAMA	NOMBRE DEL PROGRAMA	JUSTIFICACIÓN
107	Certeza Jurídica para la Sociedad	La realización de funciones notariales por parte de la Consejería jurídica del Estado describe la operación de un trámite, pero no caracteriza la provisión de un bien o servicios a la población que asegure la gobernabilidad mediante la atención de la agenda pública del estado.
117	Programa Integral de Servicios y Soluciones de TIC e Innovación Tecnológica Gubernamental	Estrategia interna de fortalecimiento de la administración mediante arquitectura TIC.
128	Vigilancia y Evaluación de la Auditoría Superior del Estado	No continúa en 2018.
130	Fortalecimiento, conservación y administración de los bienes y servicios del estado	Función interna de mantenimiento institucional.
136	Representación jurídica del Estado	Prestación de servicios jurídicos internos, gestión organizacional interna.

NUMERO DE PROGRAMA	NOMBRE DEL PROGRAMA	JUSTIFICACIÓN
142	Fortalecimiento legislativo del Estado	La aprobación de leyes es una función constitucional primaria, no la prestación de un bien o servicios directo a la población.
148	Eficiencia del Gasto Público para Resultados	La implementación de la orientación a resultados es un programa híbrido, ya que produce bienes públicos indirectos. Aunque no se aconseja equiparlo a un programa público, hay razones simbólicas por las que la orientación a resultados de esta implementación es relevante como ejemplo o <i>benchmark</i> . Esta posibilidad se discutirá en el taller participativo.
179	Fortalecimiento de las actividades del poder Ejecutivo	Las gestiones de coordinación interinstitucional son parte de la gestión interna, no atienden directamente a la población de manera sistemática.
186	Impulso al Desarrollo Económico	No continúa en 2018.
204	Evaluación de Programas de Seguridad Ciudadana y Procuración de Justicia	La realización de evaluaciones no ofrece un servicio directo a la población. No obstante, al igual que en el caso del programa 148, hay condiciones híbridas que se discutirán en el taller participativo.
205	Combate a la corrupción y actuación de los servidores públicos	La instalación de Comité de Ética, en sí misma, solo es una actividad interna, no hay teoría de cambio externa.

## **BIBLIOGRAFÍA**

Aldunate, Eduardo y Julio Córdoba, Formulación de programas con la metodología de marco lógico, Comisión Económica para América Latina, Serie Manuales # 68, Santiago de Chile, 2011.

Rodríguez Bilella, Pablo D. et al. Estándares de Evaluación para América Latina y el Caribe / Rodríguez Bilella, Pablo D.; Sergio Martinic Valencia; Luis Soberón Álvarez; Sarah D. Klier; Ana L. Guzmán Hernández; Esteban Tapella. 1a ed. – Ciudad Autónoma de Buenos Aires, 2017.