

La utilización de las evaluaciones de programas públicos. El caso del Sistema Nacional de Investigadores

*The Use of Evaluation of Public Programs.
The case of National System of Researchers*

Myriam Cardozo Brum*

ORCID: 0000-0002-5790-6749

Recibido el 27 de noviembre de 2019; aceptado el 5 de agosto de 2020

RESUMEN

La evaluación de programas muestra paradójicamente niveles internacionales de utilización limitados, incluido México. El objetivo del presente artículo es analizar las recomendaciones efectuadas a programas federales mexicanos evaluados en 2007-2017, sus compromisos y cumplimientos, aportar explicaciones de esta situación y realizar propuestas en el caso del Sistema Nacional de Investigadores. La metodología utilizada es el análisis de información documental y entrevistas. Los resultados muestran recomendaciones que no se transforman en compromisos institucionales y temas recurrentes no resueltos, especialmente, la dificultad para llevar a cabo evaluaciones de impacto. Se concluye en la necesidad de efectuar cambios en las organizaciones vinculadas a la evaluación que favorezcan la utilización analizada.

Palabras clave: evaluación de programas públicos, organizaciones públicas, utilización de la evaluación.

Código JEL: D73.

*Universidad Autónoma Metropolitana, Unidad Xochimilco.

ABSTRACT

The international utilization of the evaluation of programs is limited, in Mexico too. The objective of this paper is to analyze the recommendations made by the evaluation of federal programs in Mexico during 2007-2017, their commitments and the actual fulfilment of them; to review possible explanations and to make suggestions, in the case of National System of Researchers. The methodology used is based on the analysis of documental information and interviews. The results present recommendations that were not transformed into institutional commitments and recurrent issues that have not been fully resolved, particularly, the difficulty to undertake high impact evaluations. It concludes the need to make changes in the organizations of evaluation that favor the analyzed use.

Keywords: evaluation of public programs, public organizations, use of evaluation.

JEL code: D73.

Introducción

La evaluación se propone expresamente rendir cuentas a la ciudadanía y mejorar el desempeño de los programas públicos, pero como cualquier proceso, puede conllevar otras finalidades implícitas. De ahí que, el objetivo este trabajo sea rescatar la utilización de los resultados de evaluaciones, cuya calidad y utilidad ha sido valorada positivamente, para mejorar o cancelar programas, asignar recursos y rendir cuentas a la ciudadanía. Dicha utilización puede realizarse de muchas maneras y no hay una tipología universalmente aceptada. Diversos investigadores como Leviton y Hughes (1981) han identificado tres tipos de utilización: instrumental, conceptual y persuasiva. La primera ha resultado en una utilización limitada y las otras dos, han sido difíciles de medir.

En nuestro contexto nacional, es la utilización instrumental la que más se ha desarrollado y cuenta con información. Se dispone de un informe anual sobre el tema que, aunque muestra avances en los últimos años, aún requiere mayor esfuerzo. También existen investigaciones que recogen la percepción de académicos y de funcionarios internacionales sobre la necesidad de que se logre mayor utilización de los resultados de las evaluaciones (Székely, 2013; Cardozo, 2015).

En este artículo, se comienza por analizar la diversidad de factores explicativos de este fenómeno que las investigaciones teóricas y empíricas internacionales muestran, las que van desde el clima político dominante hasta aspectos organizacionales más concretos como la falta de involucramiento de los operadores y usuarios del programa o la ausencia de incentivos. Frente a la situación planteada, el presente trabajo intenta responder dos preguntas de investigación interrelacionadas: ¿cuáles han sido los procesos de evaluación desarrollados y los principales problemas que limitan la utilización de sus resultados en México? y ¿cuáles los cambios que se pueden sugerir para superarlos a las organizaciones que lideran dichos procesos?

En la imposibilidad de analizar las más de 200 evaluaciones federales que el Consejo Nacional de Evaluación de la Política de Desarrollo Social (Coneval) y la Secretaría de Hacienda y Crédito Público (SHCP), como responsables de coordinar los procesos de evaluación a nivel federal, realizan anualmente en la actualidad, se optó por analizar el caso de un programa muy cercano al medio académico (obviamente no representativo de todos los demás): el Sistema Nacional de Investigadores (SNI). En consecuencia, este artículo contextualiza brevemente a la institución responsable del mismo (Consejo Nacional de Ciencia y Tecnología, Conacyt) y profundiza en la utilización instrumental de los resultados de 13 evaluaciones realizadas al mencionado programa, en el período 2007-2017¹ con objeto de responder en este caso las preguntas antes formuladas. Esta situación es analizada a la luz de la literatura internacional sobre el tema, lo que permite arribar a conclusiones y sugerir recomendaciones.

1.- Literatura internacional y nacional sobre el tema

Como ya se mencionó, la utilización de los resultados obtenidos es el objetivo que le otorga sentido a la evaluación en el proceso de mejora de las políticas y los programas públicos, una vez verificada la calidad de la misma, así como la conveniencia y viabilidad de aplicar sus recomendaciones. Dicha verificación se denomina metaevaluación y, aunque no es infalible, permite reducir los riesgos de aplicar sugerencias incorrectas formuladas por el evaluador del programa. Fue desarrollada en varios trabajos de Scriven, Stufflebeam y Shinkfield, Schwandt y Halpern y House (Cardozo, 2015). Se trata de “un tipo particular de evaluación que enjuicia la calidad y el mérito de lo que hacen los evaluadores profesionales, persigue el perfeccionamiento de sus trabajos y promueve una mayor comprensión de los mismos” (Cardozo, 2012: 187). Analiza y valora el contenido de los informes finales de evaluación, pero también la organización y funcionamiento del sistema de evaluación y el papel de otras instancias y actores que afectan la calidad final del trabajo del evaluador.

La literatura internacional sobre utilización de resultados de la evaluación coincide en la relevancia que el tema ha tomado en las últimas décadas debido a su limitado uso en la toma de decisiones (Weiss, 1988 y 1998; Hofstetter y Alkin, 2003) y en los procesos de influencia (Kirkhard, 2000; Henry y Mark, 2003; Mark y Henry, 2004). Siguiendo el trabajo de Leviton y Hughes (1981), se identifican tres tipos de utilización: a) instrumental, que provoca un cambio específico decidido por las autoridades del programa como resultado de la evaluación; b) conceptual, que implica influir en el pensamiento o actitud de otros actores del proceso (ej. diseñador de políticas); y c) persuasivo, que consiste en aumentar la contundencia de un argumento con la intención de convencer a otros.

Considerando el uso instrumental en un sentido amplio es posible incluir en él los cambios en materia de recursos para el programa, pero en la medida que son asignados por un órgano diferente al responsable del mismo², se considera más adecuado agregar un cuarto tipo de utilización: presupuestal (Cardozo, 2015). Ésta fue la que brindó uno de los primeros argumentos, presentados por la Cámara de Diputados a inicios de este siglo, para institucionalizar la evaluación.

1 En 2018-2019 no se ha publicado ninguna nueva por parte de Coneval.

2 En el caso mexicano la Secretaría de Hacienda y Crédito Público prepara el proyecto de presupuesto que será discutido, modificado y finalmente aprobado por la Cámara de Diputados.

Autores de contextos nacionales diferentes y variadas formas de organización de la evaluación, han investigado la complejidad de la utilización de sus resultados. En torno de los factores con mayor poder explicativo del nivel de utilización reina una fuerte diversidad y falta de consenso (Preskill y Caracelli, 1997; Flischer y Christie, 2009), pero, en términos generales, nos permite visualizar la mayor concentración de referencias teóricas y empíricas a los temas de cultura de la evaluación como prerrequisito fundamental para la utilización efectiva de sus resultados.

Así también como a la importancia del involucramiento de sus usuarios durante todo el proceso para mejorar su apropiación de los hallazgos (Patton, Bonbright, Bustelo, Hojlund, Laublie y Mayne citados en Cardozo, 2015). En segundo lugar, se ubican las menciones vinculadas a la credibilidad de la evaluación que, si bien se relacionan con su calidad, parecen depender más del prestigio de los evaluadores (Appleton-Dyer et al, Contandriopoulos y Brouselle, Laublie y Mayne citados en Cardozo, 2015).

En una posición muy cercana a la anterior se ubican el contexto político y normativo, el diseño de interdependencias gubernamentales y la estabilidad del programa, que pueden favorecer u obstaculizar la utilización de las recomendaciones (Perret, Cejudo y Maldonado, Appleton-Dyer et al, Contandriopoulos y Brouselle, Hojlund citados en Cardozo, 2015), que retoman trabajos clásicos como Weiss, 1988 y Mayne, 1994).

Lo mismo sucede con los aspectos organizativos: necesidad percibida de información sobre el programa, nivel jerárquico solicitante de la evaluación, apoyo de autoridades superiores y planeación del proceso, incluyendo la difusión y utilización efectiva de los resultados (los autores mencionados tienden a repetirse, siguiendo el camino iniciado por Leviton y Hughes, 1981).

Si bien muchos de los trabajos revisados mencionan la vinculación entre una elevada utilización de resultados y los procesos de asignación de recursos presupuestales, existen muy pocas referencias expresas al diseño de un sistema de incentivos para favorecer la adopción de las recomendaciones (Zapico, Cejudo y Maldonado, Bonbright citados en Cardozo, 2015), útil mientras avanza el lento proceso de cambio cultural en la materia.

Por otra parte, la evaluación provoca emociones, afecta sentimientos y genera temores y estrés que pueden estar en la base de la resistencia al cambio organizacional, que requiere liderazgos comprometidos con las estrategias y decisiones de utilización de resultados, mecanismos de comunicación más abiertos, participación de los interesados, negociación, mayor motivación y desarrollo de una cultura pertinente al cambio propuesto (Schein, 2004; Robbins y Judge, 2009; Chiavenato, 2009).

Henry y Mark (2003) presentan un enfoque novedoso que retoma aspectos culturales y elementos de psicología. Además, es importante mencionar el esfuerzo de Johnson (1998) que revisa los modelos (variables y relaciones entre ellas) de utilización, explícitos e implícitos, de los autores más relevantes de su momento para proponer un meta-modelo que los recupera y articula, incluyendo también variables organizacionales, socio-psicológicas e interaccionales.

En síntesis, la revisión teórica efectuada permite concluir que la tendencia mundial muestra poca utilización de los resultados de la evaluación con algunas excepciones consolidadas en los Estados Unidos de América (EUA), Canadá y Reino Unido y otras incipientes en países latinoamericanos como Chile. Los autores revisados señalan un amplio y disperso conjunto

de posibles factores explicativos de este fenómeno; en cambio, la calidad del informe y la rigurosidad metodológica del trabajo, la utilización de incentivos y los aspectos emotivos sólo aparecen mencionados en pocos casos. Por el contrario, en el trabajo que nos ocupa, la calidad de la evaluación realizada, como ya se mencionó, es un prerrequisito para ocuparnos de que sus resultados sean tomados en cuenta. En el nivel federal mexicano, corresponde al Coneval o a la SHCP hacer su valoración —metaevaluación, aunque ésta no se exige legalmente en México— o solicitarla a un experto externo.

2. Métodos

Se empezó por definir los indicadores a emplear para la medición de la variable “utilización instrumental”: porcentaje de recomendaciones formuladas por los evaluadores que se transforman en lo que Coneval denomina Aspectos Susceptibles de Mejora (ASM), número de ASM cancelados o dados de baja, y porcentaje de ASM cumplidos. En el trabajo se da seguimiento especialmente a los ASM de tipo específico, que dependen sólo de la unidad responsable del programa.

Con base en el informe de Coneval 2013-2014 se identificó que Conacyt mostraba un buen desempeño en los tres indicadores elegidos y se resolvió analizar uno de los programas a su cargo, el de mayor antigüedad (SNI) y coincidente con nuestro interés académico, y se revisaron todas sus evaluaciones en 2007-2017. Las inquietudes muchas veces manifestadas por los investigadores participantes en el SNI hacían suponer que se identificarían algunos problemas vinculados con la organización y el funcionamiento del sistema de evaluación encabezado por Coneval, aunque no exclusivas de este programa.

Se hizo un seguimiento pormenorizado de las recomendaciones recibidas en cada evaluación y las reacciones institucionales ante ellas, señalando las que serían asumidas como ASM, así como su posterior cumplimiento. La información utilizada provino de la página web de Coneval y de la SHCP, en tanto instancias coordinadoras de evaluaciones, y del Conacyt, la que fue complementada mediante consultas formuladas por correo electrónico al funcionario responsable.

3. Análisis de la utilización de resultados de las evaluaciones 2007-2017 del Programa Sistema Nacional de Investigadores (S191) a cargo de CONACYT

El SNI fue creado en 1984 con el objetivo de otorgar distinciones y estímulos económicos a los investigadores que generen nuevos conocimientos. Las evaluaciones coordinadas por Coneval y disponibles para este trabajo fueron: cinco Evaluaciones Específicas de Desempeño (EED 2008 a 2010 y 2012 y 2014), tres evaluaciones de Consistencia y Resultados (ECyR 2007, 2011 y 2017) y cinco Específicas o Fichas de Monitoreo y Evaluación (2013 a 2017). Se utilizó información del Mecanismo para el Seguimiento de ASM de Evaluaciones Externas de Programas Federales que se comenzó a aplicar en 2008, tanto a partir de la página de Coneval como de las de Conacyt y Transparencia Presupuestaria. Salvo las ECyR de 2007 y 2017, todas las demás cuentan con un documento de Posición Institucional y de seguimiento de recomendaciones.

Hasta 2013 fueron formuladas 26 recomendaciones —pocas reiteradas—, 11 de las cuales se transformaron en ASM que se han cumplido en forma plena. Se refieren a gestionar una evaluación de impacto, mejorar los indicadores de la Matriz de Indicadores para Resultados (MIR), revisar la definición de sus poblaciones, eliminar divergencias entre plataformas informáticas, manejar expedientes en forma electrónica, realizar estudios entre otros. También existen algunos cumplimientos parciales —algunas acciones, pero no la evaluación sugerida, o la evaluación de impacto, pero no con enfoque experimental o cuasiexperimental—.

Respecto de las demás recomendaciones se argumentó que: ya se habían cumplido —replantear definición de poblaciones y cobertura, mejorar el portal, aclarar criterios de fijación de metas—, se estaba en proceso de discutir las (descentralización del SNI), no era atribución del programa (desagregación de gastos), no se consideraba adecuado —uso del índice de Hirsch— o suficientemente fundamentado —reducir la subjetividad de las Comisiones Dictaminadoras— entre otros.

Posteriormente a 2013, se reiteraron las recomendaciones en torno a la definición de poblaciones, la MIR y la evaluación de impacto, así como las respuestas de Conacyt argumentando sus desacuerdos, avances ya efectuados y agregando la solicitud de que el tema de la MIR se tratara conjuntamente con la SHCP y Coneval. Estas posiciones institucionales llevaron a que fueran muy pocos los ASM acordados: en 2013-2014 no se comprometió ninguno, al siguiente año se aceptó uno: analizar la pertinencia de realizar un estudio sobre la trayectoria y productividad de sus beneficiarios. En 2016-2017 se asumieron dos ASM: uno sobre la alineación de los documentos normativos del programa y otro en el sentido de uniformar las definiciones poblacionales. Resalta la ECyR de 2017 que realizó 20 recomendaciones, pero el SNI no asumió ninguna como ASM.

En cuanto al cumplimiento de los ASM acordados, éste siguió siendo alto. Se destaca que se realizó el análisis de factibilidad de la evaluación de impacto, una revisión y modificación del Reglamento del SNI y se avanzó en el estudio sobre la trayectoria y productividad de los beneficiarios, que se terminaría en 2018.

En general, puede señalarse que en 2007-2017 hubo un número importante de recomendaciones no utilizadas, en buena parte con argumentación convincente de Conacyt, pero resultó preocupante el rechazo de todas las propuestas en 2013 y 2017. Es importante recordar que, según informa Coneval (2017-2018), en marzo de 2014, Conacyt había acumulado un 53% de incumplimiento, razón que lo impulsó a diseñar un plan general para 2016-2018, que le permitió lograr un pleno cumplimiento en marzo de 2018. Dicho cumplimiento de ASM ha permitido eliminar divergencias entre distintas plataformas informáticas, modernizar el sistema de cómputo o modificar el Reglamento del SNI, pero subsisten problemas sustantivos reiterados en las recomendaciones desde el inicio de las evaluaciones:

a. Evaluar el impacto del programa

Coneval exige que la evaluación sea de tipo experimental o cuasiexperimental con diseño del contrafactual, lo que tiene ventajas y desventajas que no es el momento de analizar (Cardozo, 2012) y que implica enfrentar problemas técnicos y éticos importantes, por lo que en muchos casos de alta complejidad de otros contextos nacionales se ha optado por sustituir el análisis atributivo por el contributivo (Forss et al, edit., 2011, García y Cardozo, 2017) y combinar

indicadores cuantitativos y cualitativos.

Se trata de una dificultad que ha estado claramente planteada desde la primera evaluación del SNI. Esta recomendación fue transformada en ASM y dio lugar a una evaluación que el programa considera útil, pero Coneval no aceptó por considerar que su metodología no era rigurosa. A partir de allí, todas las evaluaciones realizadas insisten en la necesidad de llevar a cabo una de impacto, incluso en 2016-2017 se agrega la necesidad de planearlas en su forma ex ante de manera sistemática.

Por su parte Conacyt argumentó rigidez de parte de Coneval, poca comprensión del tipo de programas desarrollado por este Consejo, diferente a los dirigidos al combate a la pobreza, participación en el Programa de Incubadoras de Evaluaciones de Impacto de Coneval y señaló con buen fundamento que existen diferencias significativas entre quienes ingresan al SNI y quienes son rechazados, por lo que estos últimos no pueden funcionar como grupo de control en una medición de impacto de tipo cuasiexperimental. Agregaba que se tenía planeada una evaluación de impacto cuantitativo-cualitativa, a efectuarse en 2018 por prestigiados investigadores, pero su posterior informe muestra que no se trata de una evaluación global del programa sino dirigida exclusivamente a “la productividad académica posterior del investigador” (Campos, 2018).

b. Mejorar el diseño de indicadores de la MIR

Esto ha estado señalado en todas las EED realizadas y se planteó como ASM por lo menos en 2008 y 2009. Al respecto puede decirse que, si bien se han realizado reuniones de trabajo con SHCP y Coneval, y efectuado continuos ajustes, los evaluadores no se han mostrado satisfechos incluso hasta sus últimos trabajos, realizados en 2017. En 2015-2016 Conacyt señaló que las recomendaciones recibidas de SHCP y la Secretaría de la Función Pública no coinciden, lo que consideró demostrativo de la poca comunicación existente entre estas dos instancias.

Finalmente, en 2016-2017, después de recibir críticas al diseño de su MIR por la Comisión Económica para América Latina y el Caribe (CEPAL) y Coneval, de redactar un documento metodológico y de reformularla, Conacyt solicitó tratar el tema del indicador adecuado de Fin con SHCP y Coneval, ya que no puede cambiarlo porque viene predefinido en la MIR que se debe completar cada año. Si bien es fundamental que esta matriz del programa muestre su congruencia interna y diseñe indicadores pertinentes que se apliquen periódicamente y permitan la rendición de cuentas, casi no se informa de estos resultados, dejando la impresión que su perfeccionamiento se persigue para cumplir una formalidad, no por la utilidad de su aplicación.

c. Definir claramente sus poblaciones y medir su cobertura

Se trata nuevamente de una recomendación que aparece en todas las EED revisadas, pero no se había planteado como ASM hasta 2017. En 2010 la institución respondió que plantearía la metodología de su cálculo y en 2011-2012 explicitó las modificaciones realizadas que, nuevamente en 2012-2013, volvieron a ser cuestionadas por el evaluador y así ha continuado sucediendo hasta el último trabajo revisado (Seguimiento de 2016-2017). Se entiende que Conacyt ha trabajado siempre con los investigadores que demandan un apoyo y los que

efectivamente lo reciben, pero eso no significa que no pueda hacer un intento por identificar la población potencial que cumple sus requisitos y determinar una población objetivo a la que el Consejo planee apoyar.

Se concluye que se ha hecho un esfuerzo por mejorar el problema como la preparación de un documento metodológico, pero no se considera un asunto plenamente resuelto. Si bien el Modelo Sintético de Evaluación del Desempeño (MSD) de la SHCP califica la gestión de ASM del SNI como 3 (medio) en 2013 y 2014 y 4.3 (alto) en 2017³, en este trabajo consideramos que, si bien el SNI ha mostrado interés y ha realizado acciones para resolver los problemas planteados, incluso alguno no acordado como ASM, los tres puntos previos son relevantes y siguen pendientes.

A partir de la Posición Institucional de 2013 que rechazó todas las recomendaciones, se han seguido presentando algunas con argumentos bastante críticos, fundamentalmente que: 1) el programa no es el único responsable sino que sólo contribuye a promover y fortalecer la calidad de la investigación científica y tecnológica, puesto que la misma es afectada por otros programas del Conacyt y la SEP, 2) los formatos de evaluación de Coneval son muy rígidos, 3) los evaluadores no se contactan con el Consejo para evaluar sus programas, 4) Conacyt trabaja en función de la demanda recibida y no pueden calcular una población potencial, y 5) los resultados de las evaluaciones son superficiales e inoportunos. En consecuencia, en 2016-2017, solicitan otro tipo de evaluaciones más apegadas al funcionamiento real del programa.

En síntesis, las autoridades de Conacyt apoyan las evaluaciones y valoran su contribución, pero consideran que no se aplican metodologías adecuadas al programa y que sus recomendaciones no son factibles, claras y oportunas, por lo que en su mayoría no han mostrado evidencias de una utilización que contribuya sustantivamente a mejorar el programa. Para conocer rigurosamente su efecto sería necesaria la evaluación de impacto que no se ha efectuado.

4. Actitudes de las organizaciones involucradas en el proceso

Para lograr mejores resultados en la utilización de la evaluación se requieren cambios en las organizaciones involucradas: Conacyt, en tanto organización responsable del programa evaluado, y Coneval y SHCP, como órganos coordinadores de los procesos de evaluación, que se relacionan con el enfoque teórico elegido y la organización, planeación y difusión del proceso evaluativo. El análisis realizado muestra la insistencia de los evaluadores del SNI en repetir recomendaciones muy similares en todo el período de análisis, más vinculadas a asuntos técnicos (definición de poblaciones, alineación de documentos normativos, MIR, etc.) que, a la revisión integral y profunda del programa, en parte debido a que no se trata de evaluaciones de impacto. Es en este tipo de evaluaciones donde se debería superar la tendencia al análisis cuantitativo centrado en la productividad y analizar el conjunto de efectos generados, incluyendo los que han distorsionado el ambiente académico de las universidades y la calidad de sus productos.

Conacyt ha aceptado y realizado mejoras de gestión, que no modifican el fondo del programa, y ha rechazado un número importante de otras recomendaciones, en muchos casos

3 Se considera buen desempeño por encima de 3.8, pero en 2017 se identifican problemas en la información de sus padrones de beneficiarios.

argumentando dificultades bien fundamentadas. Respecto de los tres problemas reiterados y analizados se considera que, en el tema de las poblaciones, Conacyt podría avanzar y no lo ha hecho: definir y estimar el tamaño del universo de investigadores que potencialmente podría ser beneficiado con su apoyo y el del conjunto que el programa se propone atender cada año, para que pueda ser calculada la cobertura efectivamente lograda a través de la población atendida.

En relación a los otros dos, se trata fundamentalmente de asuntos vinculados al enfoque teórico, la organización y planeación del Sistema de Evaluación del Desempeño (SED) en su conjunto, que requieren ser modificados por los órganos que lo coordinan (Coneval y SHCP). En especial, la exigencia que la evaluación sea de tipo experimental o cuasiexperimental con diseño del contrafactual que, más allá de sus limitaciones técnicas y éticas, se ha vuelto casi imposible de realizar en la forma exigida por el guión de análisis de factibilidad de Coneval, provocando que sólo se hayan realizado 18 evaluaciones de impacto desde 2007 a la fecha y desconociendo otros métodos que podrían facilitarlas (García y Cardozo, 2017).

En cuanto a la MIR, se trata de una adaptación del Marco Lógico, empleado desde hace mucho tiempo por los organismos internacionales (Banco Mundial, Banco Interamericano de Desarrollo, Organización para la cooperación y el Desarrollo Económicos, Cepal), que ha permitido al sistema responder a la crítica de falta de evaluación de diseño en el período 2001-2006. Sin embargo, se ha vuelto rígida, podría cubrirse en un formato más simple, trae precargado por la SHCP el “fin” de cada programa (a menudo no compartido por los programas, como en el caso del SNI), razón por la que se cumple con su llenado sin que haya evidencia que se utilice en la gestión (tampoco su impacto ha sido evaluado).

Los asuntos aquí revisados se vinculan muy especialmente con el tema del liderazgo y la cultura organizacional presente en las decisiones asumidas, así como también con los procesos de comunicación de recomendaciones o ASM y de motivación para cumplirlas, dirigidas hacia las organizaciones responsables de los programas evaluados.

Conclusiones

El trabajo mostró que la información correspondiente a las recomendaciones, ASM acordados y cumplidos por programa, es difícil de integrar porque surge de documentos diferentes, ubicados en también distintos apartados y páginas institucionales. Los informes anuales de seguimiento de ASM de Coneval se muestran completos en cuanto al cumplimiento por dependencia, pero no se desglosan por programa. Las posiciones institucionales señalan sólo las recomendaciones que consideran aceptables para el programa, pero no hay un documento público posterior que confirme los ASM asumidos y mucho menos que informe de su cumplimiento. Esta información sólo puede obtenerse de los posteriores informes de evaluación, cuando estos los mencionan, por lo que se recomienda elaborarlo como parte del cambio organizacional necesario para lograr una mayor comunicación y rendición de cuentas del Coneval y el Conacyt a los ciudadanos⁴.

⁴ Documento que Coneval tiene que disponer para poder hacer el acumulado por institución que presenta y que Conacyt podría elaborar dado su limitado número de programas.

En muchos casos las recomendaciones de los evaluadores al SNI han sido rechazadas por Conacyt, pero sólo se han podido analizar aquellas que cuentan con la información necesaria. Nuevamente, se requiere un mayor esfuerzo de transparencia en la comunicación. En relación con las recomendaciones efectivamente transformadas en ASM, el SNI muestra un alto cumplimiento cuantitativo de los ASM asumidos, sin proceder a realizar bajas ni cancelaciones, demostrando su compromiso con esta parte del proceso.

Los beneficios obtenidos por el programa como consecuencia de la utilización instrumental de los resultados de sus evaluaciones han sido de corte operativo, como el relativo a la plataforma informática y el sistema de cómputo o la modificación de su Reglamento, pero las reiteradas recomendaciones sustantivas como la definición de las poblaciones, la construcción de la MIR y la más importante, la evaluación de impacto, siguen pendientes y requieren mayor motivación, flexibilidad y liderazgo tanto de la organización responsable del programa evaluado como de las coordinadoras de la evaluación. Para resolver el primero de los problemas anteriores, tal vez difícil pero no imposible, se requiere un cambio en la actitud y decisión de Conacyt hacia encarar la tarea de definir y estimar la población potencial, objetivo y atendida.

La construcción de la MIR exige cambios fundamentalmente en SHCP y Coneval para simplificarla y permitir que los funcionarios de Conacyt, y otras dependencias, participen libremente en su diseño, de manera que la internalicen y utilicen como instrumento de gestión, no sólo como exigencia normativa. En cuanto a la falta de evaluaciones de impacto, que afectan al conjunto de programas federales, se continúa sin saber cuáles han sido los efectos logrados en 35 años de aplicación del SNI, más allá de los incrementos provocados en la productividad de los investigadores, ya que existen pocos estudios rigurosos sobre sus efectos adversos, la calidad de los trabajos impulsados, su utilidad para los usuarios de la investigación científica e innovación tecnológica y, especialmente, su contribución a la solución de los principales problemas nacionales.

En consecuencia, se propone que Coneval modifique su liderazgo provocando un cambio en la cultura evaluativa en el sentido de reconocer que no es posible seguir sin realizar las mencionadas evaluaciones debido al enfoque teórico elegido y decida autorizar evaluaciones de impacto mediante el diseño de metodologías alternativas, ad hoc a las características de cada caso (no experimentales, contributivas, cuantitativo-cualitativas, etc.), ante la dificultad enfrentada para llevar a cabo estudios experimentales con diseño del contrafactual que han resultado difíciles, caras, no explicativas y con problemas éticos.

Para reducir la opacidad en la comunicación que rodea al tema en una organización relativamente pequeña como Coneval⁵, se propone nuevamente un cambio en su estrategia, que invierta las prioridades: que sacrifique el ritmo de crecimiento cuantitativo de las evaluaciones que coordina, con el alto riesgo de que sus resultados sean poco utilizados, para dedicar más tiempo a garantizar su calidad, comunicar, dialogar y dar seguimiento a las recomendaciones, de manera que efectivamente contribuyan a la mejora continua de los programas e, indirectamente, a lograr su finalidad. Esto debería acompañarse de una evaluación de impacto de los propios ASM aplicados a lo largo de los años estudiados.

Mucho se han estudiado las resistencias de los trabajadores a asumir los cambios organizacionales impulsados por los gerentes, pero en este caso se trata de la resistencia

⁵ Coneval tiene aproximadamente 120 trabajadores.

de las autoridades de Coneval y SHCP a modificar los procesos que lideran, a pesar de las dificultades enfrentadas en casos como las evaluaciones de impacto.

Se concluye, con las limitaciones de información mencionadas, que no existe evidencia de que los resultados de evaluación hayan tenido una amplia y profunda utilización instrumental en el caso estudiado, especialmente de tipo instrumental, situación que la literatura internacional también ha venido documentando desde los años sesenta, y que muestra la urgencia de innovar en la búsqueda de mecanismos más efectivos.

Propuesta

Para incrementar el uso de los resultados de evaluaciones en general y del SNI en particular, interesa retomar de los factores señalados en la literatura mundial, aquellos que resultan aplicables a nuestro contexto y sobre todo los que pueden ser modificados en el corto plazo por el propio sistema de evaluación, de manera de incrementar las posibilidades de que las “buenas” evaluaciones sean tomadas en cuenta. Para ello se recomienda introducir o incrementar los siguientes factores:

- Utilización de un enfoque de evaluación adecuado a situaciones de complejidad.
- Participación de funcionarios y otros actores desde la planeación de los trabajos.
- Diseño de términos de referencia ad hoc a cada programa.
- Selección de evaluadores en concursos por invitación (abierta o restringida), asegurando su probidad moral, la pertinente formación y prestigio de su equipo y justificando sus costos.
- Mayor solicitud de trabajo de campo complementario del efectuado en gabinete, incluyendo el levantamiento de percepciones de beneficiarios y organizaciones sociales; presentación oportuna, accesible y clara del informe.
- Utilización presupuestal efectiva.
- Desarrollo de procesos de metaevaluación que garanticen la calidad de la evaluación.
- En los casos de calidad verificada, buscar mecanismos (diálogos, incentivos monetarios y simbólicos, colectivos o individuales, conclusiones vinculatorias, etc.) que promuevan una más amplia utilización de sus recomendaciones.
- Pago de las evaluaciones con un fondo independiente de los programas evaluados, que garantice la libertad de expresión del evaluador.
- Seguimiento continuo del uso comprometido de las recomendaciones.
- Impulso continuo de la cultura de la evaluación.
- Promoción de la reglamentación de la ley que otorga el carácter de órgano constitucional autónomo al Coneval.

Finalmente, ante los recientes cambios introducidos por el gobierno de la Cuarta Transformación en la renovación de los liderazgos de las instituciones responsables del tema analizado, direcciones, secretarías ejecutivas, consejeros, existe optimismo de que las recomendaciones realizadas sean efectivamente incorporadas.

Referencias

- Campos, R. (2018). Informe final de diseño e implementación de una evaluación de impacto del programa presupuestal S191, Sistema nacional de Investigadores (SNI). Recuperado de <https://www.conacyt.gob.mx/index.php/el-conacyt/evaluacion-de-programas-conacyt>.
- Cardozo, M. (Comp., 2015). *Institucionalización del proceso de evaluación. Calidad y utilización de sus resultados*. México: UAM-X - CESOP.
- Cardozo, M. (2012). *Evaluación y metaevaluación en las políticas y programas públicos. El estado del arte*. México: UAM-X.
- Conacyt (Consejo Nacional de Ciencia y Tecnología) (s/f). Evaluación de Programas del Conacyt. México. Recuperado de <http://www.conacyt.mx/index.php/el-conacyt/evaluacion-de-programas-conacyt>.
- Coneval (Consejo Nacional de Evaluación de la Política Social), (varios años). Informe de Seguimiento a los Aspectos Susceptibles de Mejora de los Programas y Acciones Federales de Desarrollo Social, 2008 a 2015. México. Recuperado de http://www.coneval.gob.mx/Evaluacion/CMPE/Paginas/Seguimiento_recomendaciones_2014_2015.aspx.
- Chiavenato, Idalberto (2009), *Comportamiento organizacional. La dinámica del éxito en las organizaciones*. México: Mc Graw Hill.
- Fischer, D. y Christie, C. (2009). Evaluation Use: Results from a Survey of U.S. American Evaluation Association Members. En *American Journal of Evaluation Vol. 30(2)*, pp. 158-175.
- Forss, K., Marra, M. y Schwartz, R. (2011). Evaluating the Complex: Attribution, Contribution and beyond. En *Comparative Policy Evaluation Vol. 18*. Alemania - Reino Unido: Transaction Publishers.
- García, E. y Cardozo, M. (2017). Evaluación de impacto: más allá de la experimentación. En *Política y Cultura*, No.47: 65-91.
- Henry, G. y Mark, M. (2003). Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. En *American Journal of Evaluation No. 24 (3)*, pp. 293-314.
- Hofstetter, C. y Alkin, M. (2003). *Evaluation Use Revisited*, en Kellaghan, T. y Stufflebean, D. (eds.), *International Handbook of Educational Evaluation*. Boston: 197-222.
- Johnson, R. (1998). Toward a theoretical model of evaluation utilization. En *Evaluation and Program Planning 21(1)*. Nueva York: Elsevier Science Ltd: 93-110.
- Kirkhard, K. (2000). Reconceptualizing Evaluation Use. An Integrated Theory of Influence, en Caracelli, V. y Preskill, H. (eds.) *New Directions for Evaluation No. 88*, San Francisco: Jossey Bass: 5-23.
- Leviton, L. y Hughes, E. (1981). Research on the Utilization of Evaluations: A review and synthesis. En *Evaluation Vol. 5 No. 4*: 525-548.
- Mark, M. y Henry, G. (2004). The Mechanisms of Outcomes of Evaluation Influence. En *Evaluation No. 10*: 35-57.
- Mayne, J. (1994). Utilizing Evaluation in Organizations: The Balancing Act, en Leeuw, F., Rist, R. y Sonnichsen, R. (eds.), *Can Governments Learn? Comparatives Perspectives on Evaluation and Organizational Learning*. Nueva Brunswick y New Jersey: Transaction Publishers: 17-44.
- Preskill, H. y Caracelli, V. (1997). Current and Developing Conceptions of Use: Evaluation Use TIG Survey Results. En *Evaluation Practice, No.18 (3)*, pp. 209-225.
- Robbins, S. y Judge, T. (2009). *El comportamiento organizacional*. México: Pearson Education.
- Schein, E.H. (2004). *Organizational culture and leadership*. San Francisco: Jossey-Bass.
- Székely, M. (2013). Evaluación externa de los resultados del Consejo Nacional de Evaluación de la Política de Desarrollo Social, Instituto Tecnológico de Estudios Superiores de Monterrey, México. Recuperado de http://www.coneval.gob.mx/Informes/Coordinacion/Evaluaciones_auditorias_externas/Informe_EER.pdf.
- Weiss, C. (1998). Have we Learning Anything New about the Use of Evaluation? En *American Journal of Evaluation No.19 (1)*: 21-33.
- Weiss, C. (1988). Evaluation for Decisions: Is Anybody There? Does Anybody Care? En *Evaluation Practice Vol. 9 (1)*, pp. 5-19.