

EVALUACIÓN: DEFINICIÓN, PROPÓSITO Y USOS

Ángel Mundo López¹

¹ Profesor-investigador de la Universidad Autónoma Metropolitana Unidad Xochimilco, adscrito al Departamento de Política y Cultura.

*Author for correspondence: amundo@correo.xoc.uam.mx

RESUMEN

El documento abordó la evaluación de políticas públicas, su relevancia y cómo optimizar su implementación para mejorar programas sociales. Se utilizó una metodología basada en teorías de autores como Weiss y Cohen, se analizaron modelos de evaluación cualitativos y cuantitativos, incluye también técnicas experimentales y cuasiexperimentales. Se definió la evaluación como una actividad destinada a emitir juicios basados en criterios específicos, con el objetivo de identificar logros y áreas de mejora en programas sociales. Además, se discute la “brecha de implementación” y analizamos el impacto de la evaluación en la rendición de cuentas y el control presupuestal. Las evaluaciones de impacto favorecen métodos experimentales, pero se observó cómo los métodos cualitativos, en combinación, ayudan a comprender mejor las razones detrás de los resultados. El estudio concluye que, aunque la evaluación debe contribuir a la toma de decisiones, su éxito depende de un contexto político adecuado y de un enfoque flexible que considere las realidades de cada programa. Se enfatiza que la evaluación no debe usarse como un fin en sí misma, sino como una herramienta estratégica que permite optimizar políticas y prácticas, reconociendo la importancia de adaptar sus enfoques a la situación concreta y social de cada intervención. Con base en este análisis, se propone que la evaluación de políticas, cuando se usa de manera adecuada y reflexiva, es una herramienta fundamental para fomentar la transparencia y eficacia en las intervenciones gubernamentales, contribuyendo al ejercicio de derechos ciudadanos y al mejoramiento del bienestar social.

Palabras clave: Rendición de cuentas, Eficacia programática, Toma de decisiones.

INTRODUCCIÓN

La evaluación de políticas adquiere, con el paso del tiempo, cada vez mayor relevancia; nadie niega en la actualidad, al menos en público (Thoenig, 2018), las ventajas que ofrece evaluar políticas; sin embargo, ese enaltecimiento público no guarda correlación en asuntos como la promoción y el financiamiento, por no hablar de la

Cita:
Mundo-López A. 2025.
Evaluación: Definición,
propósitos y usos.
REMEVAL 1(1): 46-57.
[https://doi.org/10.63121/
fw29fs20](https://doi.org/10.63121/fw29fs20)

Recibido:
12 Septiembre, 2024

Aceptado:
18 Octubre, 2024

Publicado:
12 Febrero, 2025

This work is licensed
under a Creative
Commons Attribution-
Non- Commercial 4.0
International license.



utilización, que son asuntos que todavía quedan pendientes². Pese al discurso laudatorio, persisten ambigüedades en torno a su significado, usos y aprovechamiento, por ello, sin intención de esclarecer un asunto que lleva décadas en la palestra, realizaremos nuestro aporte al debate de forma sintética.

Primeramente, nos enfocaremos en la definición del término evaluación, pues, aunque parece un tema de sentido común, hablar de evaluación de políticas o programas requiere de una especificación mayor; continuaremos con las características que se espera que tenga una evaluación, y para eso nos apoyaremos de la propuesta de Weiss (2018) que establece cinco características que se deben cumplir, tratando de ahondar en sus explicaciones adaptándolas al contexto mexicano. Por último, arribaremos a algunas breves conclusiones.

MARCO TEORICO

A decir de Weiss (2018), a lo largo de toda nuestra vida llevamos a cabo actos de evaluación: como cuando calificamos algún servicio que recibimos, cuando juzgamos la utilidad de algún producto o como cuando emitimos un juicio sobre la sazón de algún alimento que degustamos. Algo similar comentan Meny y Thoenig (1992), quienes dice que: “La evaluación es un gesto diario [...] A su manera, todos los ciudadanos emiten juicios sobre lo que hace el gobierno de un Estado o el ayuntamiento”.

La forma más común en la que desde niños nos vinculamos con el tema de la evaluación es con la aplicación de cuestionarios o pruebas en las que se “juzga” el nivel de aprovechamiento alcanzado en las lecciones escolares, es decir, tras fijar un estándar (un conjunto de aprendizajes esperados), se lleva a cabo una serie de lecciones que, se espera, sean asimiladas por los estudiantes, y para estimar si el objetivo se alcanzó, se recurre a aquellos instrumentos. No obstante, este tipo de “evaluaciones”, generalmente, adquieren un carácter punitivo, en el que aquella persona que no alcanza los mínimos requeridos es obligada a retomar el curso hasta que se “cumpla” con lo esperado; y aunque desde hace ya algunos años se aboga por el establecimiento de mecanismos evaluadores formativos (y por ende, no punitivos), todavía hace falta trabajar en su diseño e implementación.

Como puede verse, la evaluación es un término que nos acompaña desde nuestras primeras etapas de la vida; no obstante, en el caso de la evaluación de políticas y programas, particularmente con un objetivo social, la definición de esta actividad adquiere un carácter un poco más técnico. Cohen y Franco (1992) definen a la evaluación como:

² Las metaevaluaciones mexicanas también identifican importantes problemas en la utilización de las recomendaciones surgidas de los procesos de evaluación (Cardozo, 2012).

La pretensión de comparar un patrón de deseabilidad (imagen-objetivo hacia la cual se orienta la acción) con la realidad (la medida potencial en la cual ésta va a ser modificada, o lo que realmente sucedió como consecuencia de la actividad desplegada) y, por otro lado, la preocupación por alcanzar eficazmente los objetivos planteados [...] Evaluar es fijar el valor de una cosa; para hacerlo se requiere un procedimiento mediante el cual se compara aquello a evaluar respecto de un criterio o patrón determinado” (p. 73).

Con base en lo anterior, podemos proponer una definición como la siguiente: evaluación es la emisión de un juicio sobre alguna intervención pública, que se emite después de haber aplicado distintas técnicas de investigación, tanto cualitativas como cuantitativas, que permiten analizar el alcance de los objetivos del programa, para arribar a un conjunto de conclusiones y emitir recomendaciones que, eventualmente pueden ayudar a mejorar la eficacia y la operación de dicha intervención.

En esta propuesta de definición intentamos recuperar algunos de los puntos que Weiss (2018) postula que debe cumplir la evaluación: 1) Realización de una valoración sistemática, que se emite, ya sea, 2) sobre la operación o, 3) sobre los impactos del programa con base en, 4) el establecimiento de determinados estándares que sirven para realizar la comparación entre lo buscado y lo que se desea alcanzar, por último, 5) identifica el propósito de llevar la evaluación a cabo. Desarrollaremos a continuación cada uno de estos puntos.

Emisión de un juicio de valor al que se arriba gracias a la aplicación de un determinado método. Si bien ya desde finales del siglo pasado, Patton reflexionaba en torno a la “guerra de los paradigmas” que, desde su perspectiva, no ayudaba al desarrollo de la evaluación, y por ello, abogaba por el establecimiento de enfoques combinados, la realidad es que las evaluaciones de impacto tienen un sesgo fuertemente alineado con los métodos cuantitativos, específicamente mediante la aplicación de procesos experimentales (pese a las complicaciones existentes, como por ejemplo, el mayor tiempo requerido para la aplicación de las técnicas de observación así como de una mayor cantidad de recursos económicos; los conflictos éticos -dejar deliberadamente sin atención a sectores que la necesitan-; garantizar la validez, tanto externa como interna, etc.). Ante esto, dicen Campbell y Stanley (1973):

Son muchas las situaciones sociales en que el investigador puede introducir algo similar al diseño experimental en su programación de procedimientos (p. ej. El cuándo y el a quién de la medición), aunque carezca de control total acerca de la programación de estímulos experimentales (el cuándo y el a quién de la exposición y la capacidad de aleatorizarla), que permite realizar un auténtico experimento [...] porque se carece de control experimental total, es imprescindible que el investigador tenga un conocimiento a fondo de cuáles son las variables específicas que su diseño particular no controla.

Gracias a la flexibilidad de los métodos cuasiexperimentales se reduce la rigurosidad que se debe tener para controlar las fuentes de invalidación tanto internas (historia, maduración, administración de test, etc.) como externas (efecto reactivo o de interacción de las pruebas, efectos de interacción de sesgos de selección y la variable experimental, etc.)³.

Lo importante, en todo caso, es identificar, qué es lo que se ha logrado (o por qué no se han obtenido los objetivos buscados) mediante la aplicación de una metodología que, idóneamente, pueda conjugar las técnicas de investigación tanto cualitativas como cuantitativas. Siguiendo a Deleau *et al.* (1986), se busca responder a preguntas del tipo: ¿En qué grado la acción pública ha provocado que se alcanzaran los objetivos? ¿Cuáles de sus elementos han contribuido? ¿Otra política podría haber sido más eficaz? ¿Estos resultados son generalizables a políticas de la misma naturaleza, pero con poblaciones objetivo o aplicadas en regiones diferentes? ¿La política ha tenido efectos positivos o negativos? ¿Qué se puede decir *ex-post* de la pertinencia de sus objetivos o qué otros objetivos deberían sustituir a los actuales?

A continuación, Weiss se refiere a la forma en la que la política se está poniendo en práctica, es decir, la operación (instrumentación o implementación). Desde el estudio señero de Pressman y Wildavski (1998), se reconoce que casi nunca una política se instrumenta de la forma en la que fue diseñada, y en parte depende de la claridad con la que se especifican los distintos pasos que hay que seguir (sin embargo, debemos tener en cuenta que, en ciertas experiencias, como lo atestigua muy bien la mexicana, muchas veces los programas públicos comienzan a instrumentarse y, *ex post facto*, se plasma el diseño en lineamientos o términos de referencia). Pero, independientemente de la minuciosidad con la que se detalle el derrotero, los instrumentadores reinterpretarán y adaptarán la política a su entorno, lo que nos lleva a plantear preguntas del tipo: ¿cómo luchar contra una práctica consuetudinaria de la administración pública mexicana en la que, por ejemplo, el manual de procedimientos es sustituido por la capacitación que realiza el burócrata con mayor experiencia? De la misma manera, en su metafórico estudio *The Implementation Game*, Bardach (1978) identificó las distintas estrategias que los burócratas ponen en marcha al momento de poder poner en práctica una política, que, gracias tanto al monitoreo como a la evaluación de los procesos podremos identificar y reducir sus efectos perniciosos sobre la política en cuestión. Todo esto hace alusión a la denominada brecha de implementación, esa que puede ser identificada mediante la evaluación de procesos o de operación.

El tercer punto al que hace referencia Weiss se refiere a los productos (llámense resultados, o bien, impacto) que genera una política, es decir, en qué medida la acción pública fue pertinente y suficiente y mitigó el problema para el cual fue originalmente

³ Para una mejor comprensión de las fuentes de invalidación internas y externas se recomienda consultar los textos de Hernández, *et al.*, (2014) y, Campbell y Stanley (1973)

diseñada y puesta operación. Ya anteriormente habíamos destacado que las evaluaciones de impacto tienen predilección por la aplicación de métodos experimentales o cuasiexperimentales que ayudan a conocer, con la mayor certeza posible, cuál ha sido el cambio generado en la población objetivo, no obstante, como también ya lo comentamos, los métodos cualitativos, en los últimos años han permitido refrendar su aporte a este propósito, sirva como ejemplo el texto de Bamberger *et al.* (2018), que detalla la forma en la que la aplicación de métodos mixtos ayuda a identificar desde una perspectiva más amplia la modificación de la situación de base que dio origen a la política y por qué. En términos concretos, los métodos cuantitativos nos permiten identificar el cuánto y los métodos cualitativos el cómo y por qué.

Pasando al cuarto punto, la politóloga estadounidense hace referencia a los propios objetivos que la intervención se planteó. En sus propias palabras: “una vez recopiladas las evidencias sobre el proceso y los resultados, la evaluación determina el mérito del programa al comparar las evidencias con una serie de expectativas, además siempre existirá un elemento de juicio. En ocasiones el criterio aplicado para emitir juicios proviene de la declaración oficial de los objetivos del programa [...] (Weiss, 2018). En este punto, tendremos que especificar dos cosas: Por un lado, en la primera parte del enunciado se puede apreciar que es necesario que la evaluación se diseñe pensando en lo que el propio programa postula, es decir, ¿qué es lo que se pretende lograr?, lo que, implícitamente, termina siendo una crítica a la estandarización de las evaluaciones, ya que no podemos pedir que un proceso realizado para un determinado contexto sirva para aplicarse a cualquier programa sin importar los objetivos⁴. Si bien es cierto que el establecimiento de los términos de referencia, como los que elabora el Consejo Nacional de Evaluación de la Política de Desarrollo Social (Coneval), pueden servir de apoyo para que otras entidades puedan aprovecharlos para evaluar sus programas, es necesario que dichos términos puedan ser adaptados al programa específico. De igual forma se debe considerar que, independientemente de las posturas normativas del evaluador, éste debe considerar el diseño de la política y no juzgar la intervención desde su marco axiológico, pues, de incurrir en este error, el programa puede ser completamente invalidado, aunque, en los hechos, esté cumpliendo con lo que se esperaba de él. Por otro lado, como acabamos de ver, Weiss especifica que, independientemente del objetivo que persigue la evaluación, existe la obligación de que el personal que la conduce emita un juicio de valor para determinar en qué medida se ha cumplido o no, con los objetivos establecidos por los tomadores de decisión.

⁴ Podríamos decir que, aunque actualmente en México, existe una fuerte tendencia a la aplicación de una misma solución (transferencias monetarias) a distintos problemas (llámense problemas de acceso a satisfactores básicos, atención de problemas de desnutrición, sustitución de servicios de cuidados, etc.), no podemos esperar que un mismo tipo de evaluación se aplique a diferentes contextos, dice Cardozo (2012) que Thoenig estaba preocupado desde el 2002 por “la tendencia a aplicar las mismas recetas metodológicas en diferentes contextos obteniendo conclusiones a partir del mero tratamiento de los datos sin tomar en cuenta una teoría explicativa que permita su interpretación”

El último punto al que hace referencia Weiss (2018) es al del propósito de la evaluación, y en este caso se debe ser muy claro: el propósito de toda evaluación de políticas públicas, y específicamente de las sociales, es: “contribuir al mejoramiento del programa y de la política”. Pese a ello, en este punto deberemos añadir algunas cosas, por ejemplo:

Para que se pueda alcanzar la mejora del programa, es necesario que, con base en las brechas identificadas (de pertinencia -es decir, que las acciones sean adecuadas para atender el problema identificado-, de congruencia -que implica que los objetivos puedan ser alcanzados con los medios diseñados- o de implementación -si se realizan los procesos tal y como fueron diseñados- (Cardozo, 2012), se emitan recomendaciones que posibiliten su adopción, pues, sólo señalar las fallas sin indicar la ruta a seguir para su corrección puede ser insuficiente para cumplir con el propósito de la evaluación. Además, es necesario precisar que sería conveniente que el equipo evaluador estableciera los periodos que considere adecuados para su cumplimiento, siempre y cuando tome en cuenta los márgenes de tiempo a los que se encuentran sujetas las dependencias, ya sea por cuestiones administrativas⁵ o normativas⁶. En este caso es importante tomar en consideración el debate, precisamente entre la misma Weiss y Patton acerca de la utilización de las evaluaciones. Patton, por ejemplo, afirma que, dado que las evaluaciones aportan mucho más de lo que cuestan, deben apoyarse y financiarse, además de que, para él, la utilización dependerá del “conocimiento [que] incluye la capacidad de proponer un enfoque adecuado para interactuar con los usuarios objetivo y las partes interesadas clave con el fin de educarlos como usuarios de la información, y trabajar con ellos en un compromiso mutuo para utilizar el proceso de evaluación y sus resultados” (en Delahais, *et al.* 2021: 38). Por su parte, Weiss reconoce que “Los responsables de programas y políticas públicas no utilizan sistemáticamente la evaluación «como base para la toma de decisiones» [además de que] Mike Patton no menciona la palabra «política». En su mundo todos se comportan de forma racional” (Delahais, *et al.* 2021: 54), es decir, Patton soslaya el peso que la política, o los intereses inter e intra organizacionales ejercen sobre la evaluación. En general, la utilización dependerá de diversos factores, v.g., la existencia de un marco normativo vinculante (¿se cuenta con un mecanismo que obligue a los instrumentadores a retomar las recomendaciones? -en este caso conviene ver el procedimiento existente en la Ciudad de

⁵ No se podría, por ejemplo, emitir una recomendación que implicara el incremento presupuestal para ampliar la cobertura del programa en un plazo inmediato, sin tomar en consideración los plazos de programación presupuestal que dependen de los calendarios fiscales en cada administración.

⁶ En este caso, tampoco podría establecerse una recomendación de rediseño de la operación sin considerar los plazos establecidos en caso de requerirse una reforma que implique a otros actores más allá del poder Ejecutivo; pongamos como ejemplo lo sucedido en México recientemente, en la que varios programas de atención a grupos vulnerables han sido establecidos en la Carta Magna, por lo que, impulsar una reforma a alguna de esas intervenciones deberá pasar por un proceso de cabildeo tanto a nivel de comisiones y partidos en el Congreso.

México-⁷), el nivel de complejidad de la recomendación, etcétera; en otras palabras, se debe tener en consideración el contexto social de la evaluación al que hace referencia Thoenig (2018).

Además de la mejora de la política, la evaluación puede servir, eventualmente, para el control presupuestal que el Poder Legislativo realiza en un régimen con un sistema de pesos y contrapesos, es lo que O'Donnell denomina rendición de cuentas horizontal. No obstante, en entornos como el mexicano, en el que la aprobación presupuestal no necesita de una mayoría calificada, es altamente factible que el presupuesto se defina, prácticamente, en el sentido que defina el titular del Poder Ejecutivo (si es que su partido cuenta con mayoría relativa en el Congreso, tal y como sucedió durante varias décadas durante la hegemonía priista, o como ha acontecido en las últimas dos legislaturas en las que el partido del Presidente ha contado con esa ventaja en ambas cámaras), es decir, que se realice una asignación de los recursos basada preponderantemente en los intereses del partido en el poder y no en los resultados identificados por la evaluación.

De manera paralela, la evaluación también podría ser utilizada como un medio para la rendición de cuentas vertical (más allá de la perspectiva limitada de O'Donnell (1997))⁸. La rendición de cuentas vertical a la que aludimos es la que Schedler (2008) definió asociándola a tres componentes: información, justificación y sanción; es decir, las evaluaciones funcionarían como el primer componente de la tríada schedleriana,

⁷ Por ejemplo, en la Ciudad de México, la Ley de Desarrollo Social (LDS), establece en su artículo 42D que será el consejo de Evaluación (Evalúa CdMx), quien, con base en los informes finales de las evaluaciones, analizará las recomendaciones y propondrá a su vez las propias, juzgando su viabilidad tanto normativa como política y presupuestaria, acordando un calendario de cumplimiento con el ente evaluado, y en caso de que no suceda de esta manera, se iniciará un proceso de controversia que será dirimido por la Comisión intersecretarial de Desarrollo Social y, “[...] Una vez agotado el procedimiento ante la Comisión el cumplimiento de las recomendaciones será obligatorio, estableciéndose entre el Consejo de Evaluación y el evaluado un programa y calendario para su cumplimiento. La omisión en el cumplimiento de esta obligación será hecha del conocimiento de la Contraloría General del Distrito Federal [sic] y sancionada en términos de la Ley Federal de Responsabilidades de los servidores públicos”. Pese a las aparentes ventajas de este modelo, también se debe considerar que la conformación del calendario y el llamado a la conformación de la Comisión Intersecretarial es una potestad del titular de la Jefatura de Gobierno, es decir, la propia autoridad que instrumenta las acciones de desarrollo social, por lo que lo convierte en juez y parte, lo que puede entorpecer los procedimientos para dirimir las controversias. Por ello, la promulgación de la Constitución Política de la Ciudad de México en el 2017 establece en el artículo 47-3 que, “El Consejo de Evaluación de la Ciudad de México, con base en la ley de la materia, determinará mediante acuerdos generales el número de comités encargados de evaluar respectivamente las políticas, programas y acciones en materia de desarrollo económico, desarrollo social, desarrollo urbano y rural, seguridad ciudadana y medio ambiente. Las recomendaciones que emitan los comités serán vinculantes para orientar el mejoramiento de las políticas, programas y acciones”.

⁸ La rendición de cuentas horizontal, de acuerdo con O'Donnell (1997), es la que se realiza cuando el poder vigila al poder, es decir, cuando hay un efectivo sistema de pesos y contrapesos entre los poderes Ejecutivo, Legislativo y Judicial, en tanto que la rendición de cuentas vertical es identificada, casi exclusivamente por el politólogo argentino con los procesos electorales, en donde la ciudadanía tiene la potestad (cada determinado tiempo) de premiar o sancionar a las autoridades.

proporcionando información a los ciudadanos sobre la efectividad de las políticas; en un segundo momento, los ciudadanos podrían interpelar a las autoridades sobre los resultados obtenidos y ellas tendrían la obligación de argumentar, fundar y motivar el porqué de dichos resultados; y por último, en caso de no ser satisfactoria o justificada la respuesta, se podría llegar al tercer y último componente: el de la aplicación de sanciones, una situación que, no obstante se debe matizar, algo que haremos en el siguiente párrafo.

Dentro de la evaluación podemos encontrar dos enfoques: el punitivo y el formativo. El punitivo es aquel en el que la evaluación sirve para aplicar sanciones, ya sea de tipo político-electoral, presupuestal, administrativas o penales, en contra de las instancias o funcionarios públicos, ya sea por incumplimiento de sus obligaciones o porque no se alcanzan los resultados esperados; esto ocasionará, en parte, reticencias y desconfianza hacia la evaluación, precisamente, por pensar en las consecuencias negativas que ella trae consigo. Por otro lado, el enfoque formativo o de aprendizaje, de acuerdo con Nioche, representa la búsqueda de que el evaluador recoja y juzgue las evidencias, y funja como “facilitador del proceso de aprendizaje, lo que implica un cambio de rol y de las características requeridas para cumplirlo (interdisciplinariedad, habilidades de negociación, didácticas, etc.)” (citado en Cardozo y Mundo, 2012: 28). En México, desde finales de los años 90, se introdujo la evaluación de desempeño de la mano de los postulados de la NGP, y ésta tenía un fuerte carácter punitivo, es decir, castigar a los funcionarios que no alcanzaran los objetivos que tenían encomendados, no obstante, desde hace algunos años, tanto el Coneval como el Evalúa, por poner un ejemplo, han ido adoptando una perspectiva que adopta más por el aprendizaje que por la sanción. Sin embargo, debemos dejar en claro que las políticas que no obtienen resultados adecuados deben ser modificadas o sustituidas puesto que se trata de intervenciones que no logran establecer los cambios esperados en la población que atienden.

Adicionalmente a lo espuesto, se deben considerar un par de asuntos adicionales. Primeramente, si la evaluación tiene un solo propósito (la mejora de las intervenciones públicas, llámense políticas, programas o proyectos) y distintos usos (rendición de cuentas horizontal por medio del control presupuestal, ejercicio democrático, rendición de cuentas vertical, etc.), se debe tener presente que la evaluación no es un fin en sí mismo. Si por un momento adoptamos la visión romántica de **Patton**, en donde pareciera que la información proporcionada por la evaluación sirve, sí o sí, para la mejora de las políticas independientemente de los procesos políticos y técnicos, es porque la evaluación forma parte de todo un proceso, el proceso de toda política pública y, por ende, debemos considerar a la evaluación dentro de todo un esquema de planeación.

Yehezkel Dror (1980), durante las últimas décadas del siglo pasado propuso el establecimiento de un modelo general de planificación en el que, desde la fase diagnóstica se debía pensar en llevar a cabo las actividades de evaluación a lo largo de todas

las etapas de las políticas, por eso, las evaluaciones *ex ante*, *ex post* y *ex tempore* debían acompañar tanto la formulación como la implementación de la política. También Cohen y Franco (1992) establecían un marco general en el que no se separaba a la evaluación del resto de las actividades de los programas públicos, sino que iban de la mano, sirviendo como un proceso de retroalimentación.

La palabra planificación comenzó a padecer de un estigma que la asociaba, en el contexto de la Guerra Fría, al bloque socialista (por la planificación quinquenal realizada en la antigua Unión de Repúblicas Socialistas Soviéticas). Dice Luis Aguilar que este tipo de planificación se encontraba asociada al modelo racional exhaustivo⁹, propio de las sociedades totalitarias, que implicaban una mirada omnicompreensiva, así como una alta capacidad de procesamiento y análisis de la información que permitía conocer no sólo las causas, sino las consecuencias de cada alternativa de acción (2000, 2000a). Más adelante, la adopción del modelo de la Nueva Gerencia Pública, que buscó instaurar el modelo de administración gerencial privada en las organizaciones gubernamentales sin considerar las enormes diferencias entre ambos campos, continuó en su ocultamiento de la importancia de la planeación, y no ha sido sino hasta hace muy poco tiempo cuando se ha reafirmado la importancia de apostar por la planeación.

Por último a este respecto debemos decir que, sin la pretensión de realizar un cansado resumen de las distintas tipologías que existen en torno a la evaluación (*ex ante*, *ex post*, *ex tempore*, interna, externa, mixta, participativa, de diseño o formulación, de procesos o de implementación, de resultados, de impacto), puesto que ya Cardozo y Mundo (2012), Cardozo (2012), Franco y Cohen (1992), han abonado a este objetivo, es necesario decir que, se debe diferenciar entre la evaluación y la investigación evaluativa, pues, esta última, sólo se preocupa por la acumulación del conocimiento, mientras que la primera tiene fines “aplicados e instrumentales” (Cardozo y Mundo, 2012), y aunque Weiss se refiere de forma indistinta a la evaluación y a la investigación evaluativa, a aquella la caracteriza su carácter práctico, mientras que la segunda se conforma con realizar hallazgos; la primera se realiza a pedido de los instrumentadores que buscan la mejora de la acción pública en tanto que la segunda es iniciada por una iniciativa personal por parte de los investigadores, generalmente asociados a los centros de investigación de las universidades. En síntesis, sólo la investigación evaluativa representa un fin en sí mismo, en tanto que la evaluación amplía sus horizontes. La evaluación tiene un valor sólo en la medida en la que permite el rediseño de las políticas y la mejora en la toma de decisiones.

El segundo aspecto a considerar, y con el que pretendemos concluir este trabajo, se encuentra íntimamente relacionado con el anterior. Si asumimos nuevamente

⁹ A la letra, Aguilar menciona lo siguiente: “En un frente los racionalistas, analistas sinópticos, comprehensivos, planificadores, maximizadores...; en el otro frente, los incrementalistas, racionalistas limitados, pluralistas, los politólogos de las [ciencias de las] políticas” (2000: 59)

que algunos sectores consideran a la evaluación como un fin en sí mismo, sin considerar que la utilidad de la evaluación va más allá de identificar los resultados de la política, en los últimos años también se ha dado un proceso metonímico, en el que una parte asume la posición de un todo, en este caso nos referimos a la reificación que ha experimentado la palabra evidencia. En los últimos años una corriente de pensamiento en la que podemos mencionar los textos de Head (2010) y de Merino *et al.* (2021), así como la realización de múltiples seminarios, como el del Instituto 512 (2022) o el del Conevalvideo (2023), vemos que se antepone a la evidencia (y no a la evaluación) como la base para la toma de decisiones.

De entrada, debemos decir que la palabra evidencia es un calco del inglés *evidence*, y el Diccionario Panhispánico de la Lengua Española, señala la traducción como *certera*, tanto en su calidad de manifestación concreta como de cualidad abstracta, pero identifica un enorme abuso en el uso de esta palabra, y añade que no se justifica

El empleo indiscriminado en español de la voz evidencia como sinónimo de prueba o indicio, calco censurable del inglés *evidence*: en inglés, *evidence* es toda prueba (circunstancial, testimonial, material, documental, etc.) que se alega en un proceso judicial; en español, solo sería aceptable como sinónimo de prueba evidente, esto es, prueba clara y manifiesta; así, no resultan apropiados usos como los siguientes: “Las evidencias que se han aportado no parecen en todo caso muy convincentes” (Ninyoles Idiomas [Esp. 1977]); “Las circunstancias y las evidencias eran claras en contra del Dr. Sittón» (Siglo [Pan.] 12.5.1997)” (RAE).

Por otro lado, Hausmann (2016) menciona que, “es sensato exigir que las políticas se basen en evidencia y que esta sea lo mejor posible, dentro de límites razonables de tiempo y de presupuesto. Sin embargo, la forma en que ahora se implementa este enfoque puede estar causando mucho daño al debilitar nuestra capacidad de aprender y de mejorar lo que hacemos”. Este exministro de planificación venezolano, advierte que esta tendencia proviene, como decíamos anteriormente, de esa preponderancia que existe en la aplicación de métodos experimentales (con su epítome, los ensayos de control aleatorizados), no obstante, esta preponderancia de la evidencia puede conducir a errores de atribución, como por ejemplo, pensar que la aplicación de un determinado método (como la introducción de una nueva tecnología para mejorar el aprovechamiento académico -es decir, la evidencia-) puede extrapolarse a otros entornos en los que no rendirá los mismos frutos. Por eso, Hausmann aboga por un proceso de aprendizaje y retroalimentación concomitante.

Por otro lado, es necesario detallar que, en el lenguaje jurídico, no se sentencia a un delincuente con base en las “evidencias”, sino que éstas sirven para contextualizar la forma en la que se comete un determinado delito. La valoración de esas pruebas (o evidencias) por parte del juez y del jurado es la que permitirá determinar la culpabilidad o la inocencia del acusado. Trasladando este razonamiento a la evaluación, no son

las evidencias las que permiten la mejora de una determinada política, se trata de todo el proceso evaluativo en el que ellas juegan un papel importante (más no el único), ya que permiten realizar ciertas afirmaciones sobre la política, pero no serán el único que factor que permita emitir el juicio de valor. La evidencia forma parte de todo un proceso, por lo que no puede adquirir la forma de ese todo, pues, esas evidencias, como dice Hausmann, pueden ser juzgadas erróneamente y pueden no producir los mismos resultados en contextos diferentes.

CONCLUSIONES

En el presente se especificó la definición de la actividad de la evaluación, no en su sentido coloquial, como nos han ayudado a identificar los autores citados, sino en su manifestación como parte del proceso de las políticas, por lo que, la evaluación tiene su valía en la medida en la que cumple su principal propósito (la mejora en la toma de decisiones para el rediseño de la política), y puede ayudar a otro tipo de actividades que se realizan en un régimen democrático.

Con base en ello, se concluye que las políticas, los programas y los proyectos, pueden mejorar sus resultados en la medida en que se evalúan y, a pesar de sus debilidades, un sistema vinculante puede ayudar de mejor manera a sacarle el máximo provecho a este tipo de esfuerzos (pese a las distintas limitaciones que presenta); aunque no se puede soslayar un sistema como el que el Coneval instrumenta con su Seguimiento a Recomendaciones de Evaluaciones Externas que, implícitamente asume, que los funcionarios y tomadores de decisión harán lo que más convenga a las intervenciones públicas.

Por todo ello, la evaluación es una herramienta indispensable para ampliar el ejercicio de los derechos políticos y ciudadanos, en un régimen democrático, pero también para abonar en el cumplimiento de los derechos económicos, sociales, culturales y ambientales.

LITERATURA CITADA

- Aguilar, L. (2000). *El estudio de las políticas*. Miguel Ángel Porrúa.
- Aguilar, L. (2000a). *La hechura de las políticas*. Miguel Ángel Porrúa.
- Bamberger, M., Rao, V., y Woolcock, M. (2018). Uso de métodos mixtos en monitoreo y evaluación: Experiencias del desarrollo internacional. En G. Pérez y C. Maldonado (Eds.), *Antología sobre evaluación de impacto* (pp. 380-402). CIDE.
- Bardach, E. (1978). *The implementation game*. The MIT Press.
- Campbell, D., y Stanley, J. (1973). *Diseños experimentales y cuasiexperimentales en la investigación social*. Amorrortu.
- Cardozo, M. (2012). *Evaluación y metaevaluación en las políticas y programas públicos: Estado del arte*. UAM.

- Cardozo, M., y Mundo, Á. (2012). *Guía de orientación para la evaluación de políticas y programas de desarrollo social*. Evalúa-DF.
- Cohen, E., y Franco, R. (1992). *Evaluación de proyectos sociales*. Siglo XXI.
- Conevalvideo. (2023, [05 de septiembre]). *Foro CONEVAL-INAFED: Evidencia para la toma de decisiones en los municipios de México* [Video]. YouTube. <https://www.youtube.com/watch?v=YuknMwf7P6Y>
- Deleau, M., Nioche, J. P., Penz, P., y Poinard, R. (1986). *Évaluer les politiques publiques: méthodes, déontologie, organisation, rapport du groupe de travail*. La Documentation Française.
- Delahais, T., Devaux-Spatarakis, A., Revillard, A., y Ridde, V. (2021). *Évaluation: Fondements, controverses, perspectives*. Éditions Science et Bien Commun.
- Dror, Y. (1980). *Un modelo general de planificación*. DCPyS, UNAM.
- Hausmann, R. (2016, 25 de febrero). El problema con las políticas basadas en evidencia. *Project Syndicate*. <https://www.project-syndicate.org/commentary/evidence-based-policy-problems-by-ricardo-hausmann-2016-02/spanish>
- Head, B. (2010). Evidence-based policy: Principles and requirements. En *Strengthening evidence-based policy in the Australian federation* (pp. 13-26). Productivity Commission.
- Instituto 512. (2022, [22 de junio]). *Webinar: Toma de decisiones basada en evidencia* [Video]. YouTube. <https://www.youtube.com/watch?v=f8vQtl7OVOE>
- Hernández, R., [otros autores]. (2014). *Metodología de la investigación*. McGraw-Hill.
- Meny, Y., y Thoenig, J. C. (1992). *Políticas públicas*. Ariel.
- Merino, M. (2021). *Rendir cuentas en situaciones de emergencia: Un modelo basado en la evidencia*. Universidad de Guadalajara.
- O'Donnell, G. (1997). Rendición de cuentas horizontal y nuevas poliarquías. *Nueva Sociedad*, 152, noviembre-diciembre, pp. 143-167.
- Pressman, J., y Wildavsky, A. (1998). *Implementación*. FCE.
- Real Academia Española. (2005). *Diccionario panhispánico de dudas*. Espasa Calpe.
- Schedler, A. (2008). *¿Qué es la RC?* IFAI. http://www.iepcjalisco.org.mx/sites/default/files/que_es_la_rendicion_de_cuentas.pdf
- Thoenig, J. C. (2018). La evaluación como conocimientos utilizables para las reformas de la gestión pública. En C. Maldonado y G. Pérez (Eds.), *Antología sobre evaluación* (pp. 269-292). CIDE.
- Weiss, C. (2018). Preparando el terreno. En C. Maldonado y G. Pérez (Eds.), *Antología sobre evaluación* (pp. 43-84). CIDE.