

# 4

## LA TEORÍA DE LA EVALUACIÓN DE PROGRAMAS

Catalina Martínez Mediano

*Departamento de Métodos de Investigación y Diagnóstico en Educación. UNED*

**E**n este trabajo sobre *La teoría de la evaluación de programas* comenzamos realizando una revisión de la situación actual de la *evaluación como disciplina de conocimiento*, mencionando sus elementos constitutivos, para centrarnos, a continuación, en la evolución que ha seguido la disciplina de la Evaluación de programas, volcada en un primer momento en los métodos de investigación contribuyendo a su desarrollo adaptado a la problemática evaluativa y, posteriormente, en el desarrollo de la teoría de la evaluación de programas, fundamental en toda disciplina científica, que orienta la elaboración de los programas sociales y educativos, clarifica los procesos y las metas de la evaluación, comprueba su realización práctica y resultados, asesora en la mejora de los programas y contribuye, en definitiva, a la construcción del conocimiento.

### 1. LA EVALUACIÓN COMO DISCIPLINA DE CONOCIMIENTO

Una disciplina es un constructo intelectual organizado. La esencia de una disciplina es su trabajo y la clave para su existencia es su concepto general y su metateoría. Podemos distinguir cuatro elementos en la emergencia y desarrollo de una nueva disciplina (Scriven, 1991:25). Los cuatro elementos son independientes.

- El primer elemento es la consciencia y reconocimiento explícito de que algo nuevo está emergiendo y existe una definición de ese algo nuevo. Lograr su definición o esbozar sus límites conlleva varios intentos parcialmente prescriptivos, siendo éste un elemento crucial porque en él descansan las normas o las pautas de desarrollo para que esa nueva disciplina crezca. Es el primer paso hacia la metateoría: la consciencia de un nuevo campo y su definición, dibujando su forma y sus límites, y ello se da en esta nueva disciplina.
- El segundo elemento consiste en identificar y desarrollar una metodología apropiada, entendida como un grupo de procedimientos e instrumentos para generar resultados útiles en el nuevo campo. La metodología puede ser altamente específica, pero puede abarcar un amplio rango o niveles de métodos generales y modelos de análisis; así ocurre en esta nueva disciplina.
- El tercer elemento requiere un desarrollo del campo consistente en una base de datos, de producción de investigaciones, de informes, de descripciones, de ilustraciones y de principios y teorías generales. Los elementos de su generalidad deben presentarse explícita o implícitamente al menos en forma de nuevos esquemas conceptuales, con la terminología específica asociada. Ello está representado por los múltiples modelos evaluativos existentes, cada vez con un mayor grado de elaboración en sus concepciones teóricas y procedimentales.
- El cuarto elemento está relacionado con el desarrollo de una metateoría de la disciplina que proporcione un método de trabajo para la práctica. La metateoría es casi siempre descriptiva y prescriptiva, derivando la última de las aportaciones de la primera. Sin una metateoría no puede identificarse la nueva área de estudio como una nueva disciplina al no tener identificado y definido su territorio, no poder explicitar sus procedimientos, su metodología y sus límites. Y es lo que pretende lograr la reflexión sobre la existencia de la disciplina y sobre los elementos comunes y diferenciales que se pueden extraer de la diversidad de modelos evaluativos.

Los límites de la evaluación como disciplina son definidos por la mayoría de los teóricos de la evaluación en términos que la consideran un área de evaluación aplicada, como es la evaluación de programas. Scriven defiende que la disciplina debe considerarse desde una concepción amplia que recibiría el nombre de Evaluación, y que estaría integrada por las disciplinas de Evaluación de programas, de productos, de personal, de rendimiento, de proyectos y de políticas, que serían las grandes seis subáreas de la Evaluación.

La Evaluación de programas utiliza métodos y técnicas de investigación que proceden de las Ciencias Sociales y de otras disciplinas como el Derecho,

la Lógica, la Ética, así como del propio desarrollo de la teoría para la evaluación de programas.

La Evaluación es una nueva disciplina pero una antigua práctica. La defensa de la Evaluación como disciplina descansa en sus características científicas en contraste con las evaluaciones subjetivas o no sistemáticas. La evaluación es considerada como ciencia por estar comprometida con la producción de conocimiento, y no sólo conocimiento práctico; la evaluación, desde el punto de vista de la práctica, se interesa por conocer el mérito relativo de diferentes métodos de enseñanza, o de diferentes teorías de aprendizaje, o por los cambios en la salud como consecuencia de nuevas concepciones en la práctica sanitaria. Pero también la evaluación es esencial para distinguir la ciencia de la pseudociencia al evaluar la calidad de la evidencia científica, la calidad de los diseños de investigación, de los instrumentos de medida, de sus conclusiones, la calidad de los procedimientos en ingeniería, en las matemáticas, en los procesos judiciales. En el sentido de proceso intelectual, la evaluación forma parte de todas las disciplinas, tecnologías, manualidades y de pensamiento racional en general (Scriven, 1991:4).

La evaluación, asimismo, es considerada una transdisciplina, definiendo este término, creado por Scriven, como una disciplina instrumental que sirve a otras disciplinas –como es el caso de la estadística o la lógica.

El proceso de la evaluación no es simple, siendo su complejidad el principal motivo de que en más de dos milenios haya eludido un análisis satisfactorio sobre su naturaleza y sus límites disciplinarios. Y es desde finales de la década de los 60 cuando se viene realizando un gran esfuerzo por analizar la naturaleza de la evaluación.

Scriven hace el siguiente análisis sobre la naturaleza de la evaluación:

*La evaluación no es sólo una mera acumulación y síntesis de datos relevantes para la toma de decisiones. Digamos que éste es uno de los componentes de la evaluación y uno de sus propósitos.*

*El otro componente, y previo al anterior, es el de las premisas o criterios evaluativos que verifican los valores relevantes y las normas elegidas para la realización de la evaluación.*

La evaluación como disciplina académica interesada en temas relativos a la evaluación de programas, de personas, de productos, etc., no ha empezado a ser considerada como tal sino en el último tercio del siglo XX. La discusión metodológica acaecida en los últimos años, desde diversos campos, ha sido fructífera para abrir a la evaluación su propio campo. Pero el primer paso para que la evaluación se constituya en una disciplina científica va más allá de la simple práctica evaluativa elaborando guías que orienten dicha prácti-

ca. Comienza identificando normas que gobiernen, prescriban, las evaluaciones, encontrándose suficientemente desarrollado en la actualidad, en el subcampo de la Evaluación de programas, mediante guías para realizar evaluaciones como las de Brinkerhoff et al., 1983; el Joint Committee, 1981; Mertens, 1989; Patton, 1980, 1990; Rossi y Freeman 1989; Smith, 1989; Worthen y Sanders, 1987; Worthen y White 1987, y muchas más.

Para lograr el desarrollo de la disciplina debemos considerar el concepto de Evaluación en toda su amplitud, incluyendo todos los ámbitos a los que se aplica, tales como la Evaluación de programas, productos, personas, control de calidad, políticas, etc. Además, debemos ir más allá de la discusión metodológica para entrar en temas sobre los límites y contenidos de la Evaluación, las diferencias y relaciones de la Evaluación según se desarrolle en ámbitos científicos, humanísticos, jurídicos, artísticos, etc. Explicaciones sobre la adecuación de una u otra metodología según el ámbito y propósito de la evaluación (metametodología), el desarrollo de los procesos lógicos de las evaluaciones, así como el desarrollo de las teorías evaluativas y la metateoría de la Evaluación.

La metateoría es la teoría sobre la naturaleza de un campo de investigación. Está interesada por la definición del campo que estudia, por sus límites, por las diferencias entre el campo y los campos colaterales, por las razones por las que ciertos métodos trabajan bien en la disciplina y las razones por las que otros son inadecuados. Debe ser capaz de desarrollar argumentos en favor de la Evaluación como disciplina, así como sugerir y justificar líneas para la investigación y su desarrollo.

Hasta hace poco, la elaboración de los distintos modelos sobre la evaluación de programas es lo que formaba parte de la metateoría de la evaluación al constituir un nivel que orienta las realizaciones prácticas y discusiones metodológicas. La metateoría se va elaborando desde todos los niveles e incidiendo en ellos; por ejemplo, en el desarrollo de los instrumentos de medida, en el de los métodos, las técnicas y las teorías y su discusión; desde todos los ámbitos donde se utiliza la evaluación nos preguntamos, explicamos y justificamos el uso de una técnica, de un instrumento, de un método, de una teoría, de un plan o proyecto en su totalidad y del paradigma que lo fundamenta. Continuamente estamos recurriendo a la metateoría y estamos contribuyendo al desarrollo de la misma. La metateoría puede estar implícita en la misma práctica y enseñar mediante el ejemplo de esa práctica. Y, aunque los modelos de Evaluación de programas han contribuido al desarrollo de la disciplina de la Evaluación, hay que seguir trabajando para lograr esa aportación fundamental que dote a la evaluación de una teoría general, una metodología general, unas guías generales que orienten la práctica y una metateoría general para llegar a la esencia de la disciplina.

La Evaluación es el nombre de una disciplina autónoma con su propia clasificación en la Biblioteca del Congreso de los Estados Unidos (Scriven,

1991:141). Esta disciplina se refiere al estudio y aplicación de procedimientos para hacer evaluaciones objetivas y sistemáticas. Dentro de esta disciplina autónoma, hay áreas semiautónomas como son las disciplinas de evaluación de programas, de productos, de personal, de rendimiento, de proyectos y de políticas. Otros campos aplicados autónomos incluidos son la valoración tecnológica, la evaluación médica y psicológica, el control de calidad. Otras aplicaciones se hacen en disciplinas tales como el curriculum, la crítica literaria, etc.

Scriven hace dos manifestaciones diferentes en torno a la Evaluación como disciplina:

1. Hay una sola disciplina de evaluación, que es una transdisciplina, consistente en un amplio rango de aplicaciones prácticas de la evaluación en varios campos. Uno de ellos, el más desarrollado, integrado por su propio concepto, teoría y metateoría es la evaluación de programas. Una naciente disciplina central dedicada al desarrollo de una lógica válida y distintiva, métodos y teorías generales para la evaluación.
2. Los distintos campos de la Evaluación como son la Evaluación de programas, productos, políticas, control de calidad, personal, etc., y que han desarrollado gran cantidad de modelos propios, pueden ser mejorados por el reconocimiento de sus conexiones con la disciplina central de la Evaluación.

## 2. DEL DESARROLLO METODOLÓGICO A LA TEORÍA DE LA EVALUACIÓN DE PROGRAMAS

La teoría proporciona un marco de referencia para ayudar a las personas a comprender su mundo y a actuar en él. La teoría es crucial para la investigación; no sólo proporciona orientaciones para analizar un fenómeno, sino que proporciona un esquema para comprender el significado de los resultados de la investigación.

Sin embargo, hasta hace poco tiempo, la literatura evaluativa apenas mencionaba la importancia de la teoría de la evaluación de programas ni la incorporación de dicha teoría en los procesos de evaluación. La tendencia ha consistido en que las evaluaciones estén orientadas por los métodos, y, aunque esto ha supuesto importantes contribuciones para el desarrollo de la Evaluación de programas, de otro lado, ha ocasionado que se descuidara el desarrollo del marco conceptual que guiara la práctica evaluativa. Shadish y Reichardt (1987) justifican tal hecho argumentando que la acción y la práctica preceden al desarrollo de la teoría en cualquier disciplina, y a la evaluación de programas le ha ocurrido lo mismo.

Esto puede explicarse, parcialmente, por la preocupación por realizar mejores investigaciones evaluativas que tuvieran en cuenta todos los elementos que inciden en la consecución de lo pretendido por un programa, lo que llevó a la necesidad de preocuparse por ampliar los enfoques y metodologías de investigación aplicados a la evaluación de programas. De modo que, en los primeros momentos del desarrollo de la Evaluación de programas, los métodos de investigación científica han jugado un papel muy importante en muchos de los trabajos de evaluación realizados en un intento por definir y conceptualizar la evaluación de programas. Así, Suchman (1967) vio la evaluación como la utilización de los métodos y técnicas de investigación científica utilizados con el propósito de evaluar. La evaluación de programas no se veía como una disciplina de conocimiento, sino como una modalidad de investigación aplicada dentro de la Educación y de las Ciencias Sociales. Y la utilización de los métodos científicos más prestigiosos, en su momento, era crucial para poder reclamar el estatus científico de esta nueva disciplina.

Así, Tyler, que concebía la evaluación como la determinación del grado en que los objetivos de un programa eran realmente conseguidos, establecía las metas generales y los objetivos conductuales y comparaba los resultados con las metas, utilizando la metodología de los diseños experimentales y cuasiexperimentales, pruebas estandarizadas de medida y el informe clásico de investigación sobre la conveniencia de mantener o modificar los programas.

Stufflebeam considera la evaluación como el proceso de identificar, obtener y proporcionar información útil y descriptiva del valor y mérito de las metas, la planificación, la realización y el impacto de un programa, con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y comprender los fenómenos implicados. Utiliza una metodología diversa en función de cada tipo de evaluación contexto, entrada, proceso o producto, y sus informes son frecuentes para proporcionar una guía para la toma de decisiones de mejora del programa, y la generalización posible es la que permita el muestreo y las metodologías utilizadas.

Para Scriven la evaluación, como ciencia de la valoración, consiste en determinar el mérito y valor del objeto evaluado. Para él la evaluación es la recogida y síntesis de información basada en criterios evaluativos que verifiquen los valores relevantes y las normas elegidas para la realización, con la finalidad de ser utilizada en la toma de decisiones. Sirve a las funciones formativa, proporcionando información para mejorar un programa durante su realización, y sumativa, proporcionando información para implantar, modificar o retirar un programa. Utiliza una metodología neopositivista, relativista y de la lógica probativa, utilizando un tipo de informe descriptivo. La generalización de los resultados es la que permite la validez externa dificultada por la relatividad de los criterios de valor.

Stake considera que la evaluación sirve para descubrir el mérito y las deficiencias de un programa. En su "modelo de evaluación respondiente", el evaluador debe asegurarse de prestar atención a aquello por lo que le piden que haga la evaluación, pero también debe prestar atención a lo que está ocurriendo en el programa y elegir las cuestiones y criterios de valor. Utiliza metodologías cualitativas, fundamentalmente el estudio de caso. La generalización es la propia de las investigaciones naturalista, apoyándose en la experiencia vicaria. Es útil si descubre el mérito y las deficiencias de un programa para ayudar a sus participantes a incrementar su comprensión y responder a sus necesidades.

Para Guba y Lincoln, la evaluación, en un contexto o situación particular, es un proceso de recogida de datos, en continua e inseparable valoración e interpretación basada en el consenso, en donde los hechos y valores van unidos. El evaluador utiliza metodologías naturalistas, trabaja en equipo, con todo el grupo implicado en la evaluación. Establecen círculos hermenéuticos para lograr construcciones e identificar cuestiones. El informe es descriptivo, facilitando la experiencia vicaria para la generalización de los hallazgos.

Cronbach define la evaluación como la recogida y uso de la información para tomar decisiones sobre un programa educativo; es una actividad diversificada y plural, y ni un solo grupo de principios puede ser suficiente para todas las situaciones. Ve la evaluación integrada en la investigación política. Defiende el uso de metodologías cuantitativas y cualitativas. Sus informes son narrativos para facilitar la experiencia vicaria. Considera que la generalización, el generar conocimiento extrapolable desde las personas, situaciones, tratamientos y observaciones a otros, es lo más importante de la investigación evaluativa. Su finalidad debe servir para la mejora de los programas y para la formación de la comunidad política.

Pérez Juste define la evaluación como un proceso sistemático de recogida de información rigurosa, valiosa, válida y fiable, orientado a valorar la calidad y los logros de un programa para la toma de decisiones de mejora del programa, del personal implicado y del cuerpo social en el que se encuentra inmerso. Utiliza metodologías cuantitativas y cualitativas. Su informe es descriptivo, basado en los datos y en los juicios. La generalización, aquella que permite el diseño, y su utilidad para la mejora del personal, del programa y de la comunidad.

Como podemos ver, el "enfoque de evaluación orientada por el método" se vio claramente apoyada por el hecho de que la mayoría de las perspectivas de evaluación han estado orientadas principalmente por el método. Por ejemplo, para Campbell y Stanley (1963) y Cook (y Campbell, 1969), defensores del paradigma experimental, una evaluación estaba mejor realizada si seguía la rigurosidad de los diseños experimentales. Por otro lado, los defensores del enfoque naturalista argumentan que los métodos cualitativos servían

mejor a la evaluación. E incluso investigadores naturalistas como Parlett y Hamilton (1978) y Guba y Lincoln (1981) negaban cualquier mérito de los métodos experimentales o cuasi-experimentales aplicados a la evaluación. Cronbach, en 1982, argumentaba que el uso de los métodos rigurosos lleva a que una evaluación sea rígida y estrecha en su alcance. Estaba interesado en incorporar la evaluación de programas a los procesos en los contextos y en las organizaciones para facilitar una toma de decisiones plural.

Las evaluaciones orientadas por los métodos han cobrado aún más fuerza por ese largo debate entre los métodos de investigación cuantitativos y cualitativos más adaptables a la evaluación. Estas diferentes perspectivas de la evaluación han supuesto una contribución muy importante al desarrollo de la Evaluación de programas, lográndose un gran consenso entre los evaluadores en torno a la idea de que una evaluación debe comprometerse con múltiples valores y temas. Como consecuencia, la perspectiva tradicional, experimentalista, ha constatado sus limitaciones en su alcance y ha evidenciado sus dificultades en las situaciones en las cuales actúa la evaluación y que requiere fijarse en múltiples aspectos. Sobre esto, Cook y Reichardt (1986) advierten que el uso de múltiples metodologías conlleva mayor tiempo y costo, y, por supuesto, una preparación adecuada. Cada uno de los enfoques metodológicos tiene sus propias concepciones, su lógica y sus procedimientos. Sin un esfuerzo hacia una fundamentación de por qué es positiva la complementariedad metodológica, la defensa del uso de múltiples metodologías en la evaluación puede no ser apoyada. Sin embargo, los modelos de evaluación de programas que defienden la complementariedad metodológica como los de Cronbach, Pérez Juste, Rossi y Freeman o Stufflebeam, aportan criterios suficientes que muestran la conveniencia de esta complementariedad metodológica dentro de sus propios modelos.

El propio Cook (1985) ha adoptado una perspectiva postexperimental, de "multiplicidad crítica", ya que defiende que la perspectiva del paradigma experimental debe transformarse en otra más comprehensiva que se adapte mejor a los aspectos políticos y organizativos de la evaluación de programas. Cook, junto a sus colaboradores Shadish y otros, afirma que los evaluadores deberían planificar sus evaluaciones previendo el uso de múltiples metodologías, investigar múltiples aspectos de los programas y considerar el punto de vista de múltiples participantes. Defiende las ventajas de esta nueva perspectiva sobre el paradigma experimental y considera que reduce las posibilidades de interpretación errónea, a la vez que proporciona una información más comprehensiva, global e integradora para los procesos políticos, además de favorecer una evaluación más racional y consciente de sus valores. En los primeros tiempos de su desarrollo, la Evaluación de programas tuvo dificultades que no podían resolverse metodológicamente, debido a que una evaluación no puede realizarse sin limitaciones de tiempo y de recursos, y no puede abarcar todas las opciones, y ello supone llegar a acuerdos que requieren no sólo de unas metodologías, sino de más de un grupo de principios y de



teorías que guíen el proceso de evaluación de programas. Shadish y Cols. (1986) dicen que el reto más importante que tiene planteada esa multiplicidad crítica es el desarrollo de la teoría evaluativa.

El uso de una metodología u otra para la evaluación de un programa dependerá de los propósitos de la evaluación, del estado de madurez del programa, de los recursos disponibles, de los intereses y necesidades de los patrocinadores y de los evaluadores y del medio político y organizativo del programa. De modo que, cuando un programa está desarrollándose, el estudiar su eficacia mediante un diseño experimental carece de sentido puesto que pasaría de puntillas sobre los procesos y los contextos; y utilizar los métodos cualitativos, en exclusividad, puede no ser lo más adecuado para valorar la eficacia de un programa de amplio alcance, elaborado para aplicarlo a toda la población de un país.

Tantas discusiones sobre los métodos puede llevar a pensar que la solución a estos problemas se resolvería tras conseguir una mayor sofisticación en los métodos de investigación. El desarrollo de los métodos de investigación para la evaluación de programas ha sido fundamental, pero se necesita una conceptualización y un esfuerzo teórico para sistematizar e integrar los factores conceptuales y los métodos de investigación para que avance la evaluación de programas y se consolide, aún más, como disciplina. (Chen, 1990:17). La teoría de programas debe proporcionar orientaciones para identificar qué temas son los más importantes en una evaluación, qué método o métodos son los más relevantes, y sugerir cómo aplicar el método para alcanzar los objetivos propuestos en la evaluación.

Chen y Rossi argumentan que las evaluaciones orientadas por el método tienden a maximizar un tipo de validez en perjuicio de otras. Advierten de estos problemas y señalan la importancia de la teoría de programas para lograr varios tipos de validez. Y tratan de resolverlo con su enfoque de *Evaluaciones conducidas por la teoría* (1981, 1983, 1987, 1989, 1992). Para Rossi y Freeman (1993), creadores del enfoque de *Evaluación comprehensiva, adaptada y conducida por la teoría*, toda evaluación debe centrarse en tres tipos de actividades: 1) análisis relacionados con la conceptualización y diseño de las intervenciones, 2) asesoramiento para la realización del programa, 3) valoración de la utilidad del programa, sus efectos y la relación entre costos-eficacia. Debe recoger información *conceptual*, de las necesidades de los programas sociales para planificar, redefinir, realizar y evaluar, *procesual* y de impacto, utilizando metodologías cuantitativas y cualitativas. El informe se orienta a la utilidad, a dar respuesta a las necesidades sociales. Y la generalización es aquella que permitan los diseños.

Entre alguna de las ventajas de la realización de *evaluaciones guiadas por la teoría de programas* está la de que la teoría permite comprobar si los resultados del programa se deben a cómo se ejecuta el programa o se debe a la

conceptualización propia del mismo. La teoría de programas ayuda a clarificar si un programa está siendo desarrollado siguiendo los criterios explicitados en el mismo y en las condiciones adecuadas. La teoría de programas puede usarse para especificar los efectos intermedios de un programa que podría llegar a evidenciar los aspectos positivos y negativos del mismo antes de que los resultados finales puedan manifestarse, lo que proporcionaría posibilidades de mejora del programa antes de que se haya concluido el programa.

La incorporación de la teoría de programas en una evaluación fortalece la información de los evaluadores para clarificar las metas del programa, para observar los procesos y para atender las necesidades de los patrocinadores y de los participantes y beneficiarios, y para orientarles en la realización de las diferentes decisiones alternativas de mejora. Como disciplina, la evaluación de programas debe desarrollar su propia teoría, y los métodos, en vez de ser tratados como un fin en sí mismos, deben ser considerados como medios para favorecer el desarrollo del conocimiento. Como dice Bickman (1987) la evaluación es, a menudo, considerada como una ciencia práctica, pero es ambas cosas, una práctica y una ciencia que requiere de su propia teoría.

Actualmente, cuando el debate metodológico ha contribuido al desarrollo de los métodos, se necesita un marco comprehensivo, general, teórico, que vincule e integre las actividades de la evaluación a un nivel más general que el que consiguen los distintos modelos de evaluación de programas. Porque, a medida que se han desarrollado los modelos de evaluación de programas, se ha ido tomando conciencia de que las intervenciones de los programas sociales o educativos no trabajan, normalmente, de un modo predecible, y que los problemas a resolver no son sencillos. Los que toman decisiones también advierten que prefieren solucionar o reducir los problemas más que iniciar cambios globales precipitados. Por ello, las evaluaciones que se centran sólo en los resultados de los programas no son útiles, puesto que no informan para tomar decisiones al no indicar en qué dirección, en qué aspectos y de qué manera se debe incidir para conseguir mejores resultados.

### **3. LA NATURALEZA DE LA TEORÍA DE LA EVALUACIÓN DE PROGRAMAS**

Un programa de intervención educativa o un programa de intervención social supone un esfuerzo intencionado y organizado para intervenir en un proceso social con el propósito de resolver un problema y proporcionar un servicio. La estructuración de los esfuerzos organizados adecuadamente y dirigidos a la consecución de las metas deseadas supone que el programa opera bajo una teoría, aunque, a veces, esta teoría sea implícita. La teoría está siempre implícita, o explícita, en la mente de los investigadores; y de los

evaluadores. La teoría explícita proporciona una guía general para la elaboración del programa y explica cómo el programa organiza su funcionamiento, aunque puede verse influida por los prejuicios. La teoría, en la evaluación de programas, debe estar explícita.

La evaluación de programas es algo más que una metodología aplicada; la evaluación es diversa en muchos modos y su potencialidad para la unidad intelectual, que es lo que Scriven llama la lógica de la evaluación, podría crear puentes entre la diversidad de evaluadores con el fin de desarrollar la teoría de la evaluación de programas.

Los intentos por mejorar *la teoría de la evaluación de programas* son antiguos. Weiss (1972) sugería que la utilización de los resultados de una evaluación podría verse reforzada si se incluyeran estos elementos básicos:

1. Un análisis de las premisas teóricas del programa.
2. Una especificación de los procesos del programa para informar acerca de los vínculos entre los objetivos y las metas del programa.
3. Un análisis de las realizaciones del programa y de sus resultados.
4. Un informe detallado de cuáles son las alternativas más efectivas que pueden ayudar a mejorar el programa, y cómo realizarlas.

Otra de las razones del movimiento para el desarrollo de *la teoría de la evaluación de programas* se relaciona con la necesidad de comprender cómo se realizan los tratamientos. Los estudios centrados en la realización de los programas y en la intervención de los mismos proporcionan evidencias de que la realización es algo muy complicado dentro del ámbito humano. Las dificultades que se han revelado en la puesta en marcha y desarrollo de los programas han señalado la necesidad de que los patrocinadores, los promotores de un programa, presten atención a estos aspectos para mejorar la implementación del programa. Es importante, por ello, que, dentro del desarrollo del marco teórico, se integre el desarrollo de los programas en la acción, como parte importante de la actividad evaluativa. Se trataría con ello de evitar lo que a veces ocurre cuando se evalúan los programas, que quienes realizan dichas evaluaciones tienden a proceder más según sus puntos de vista y su preparación metodológica que en función de las necesidades reales de ese programa, y esto no siempre es consistente con los objetivos de quienes elaboran o diseñan un programa adaptado a unas necesidades para resolverlas.

Por otra parte, las dificultades implícitas en la puesta en práctica de los programas proporciona a los evaluadores grandes dificultades para desarrollar sus estrategias y sus marcos conceptuales de modo que les permitieran

hacer frente a estos problemas, y ciertamente en las Ciencias Sociales, y en Educación, no pueden utilizarse los métodos de una manera estricta como, por otra parte, se pueden utilizar en las Ciencias Físicas. En Educación, y en Ciencias Sociales, los métodos de investigación no se pueden aplicar si no es atendiendo a los propósitos, las circunstancias y contextos en los que se desarrollan los programas.

Desde el punto de vista conceptual los evaluadores no sólo deben informar de la realización de sus evaluaciones, sino que deben trabajar con los patrocinadores del programa para conocer y mejorar los modos de implementación de la realización de los programas, y, posteriormente, deben colaborar para mejorar el programa y su realización. Deben informar sobre los cambios que se requiere hacer en el programa, qué criterios los fundamentan, qué estrategias se deben planificar para lograrlos, quiénes deben realizar los cambios y qué consecuencias, previsibles, y no previsibles, producirán las modificaciones para la mejora (De la Orden, 1996).

Chen, para argumentar su Teoría de evaluación de programas, toma una idea de Weber. El desarrollo de la evaluación de programas y su emergencia como disciplina se ha visto beneficiada por el desarrollo de la sociedad moderna, racionalista y contable. Sin embargo, Weber advertía que es importante ver la diferencia entre dos tipos de racionalidad: la racionalidad formal y la racionalidad sustantiva, o los valores. La racionalidad formal tiene que ver con la eficiencia y la eficacia en la consecución de una meta específica en un tiempo dado, mientras que la racionalidad de valor o sustantiva se refiere a los propósitos sustantivos y fines a largo plazo de los individuos, de los grupos y de las sociedades.

Cuando la racionalidad formal se dirige a una explicación parcial, estrecha, de las actividades sociales, los propósitos generales y las metas a largo plazo se pierden de vista. El argumento de Weber es relevante para la conceptualización de la evaluación de programas. Cuando su conceptualización es estrecha y simplemente mide las metas conseguidas, el énfasis se pone en la racionalidad formal, los evaluadores se interesan por la eficiencia y por la eficacia del tratamiento, descuidando algo tan esencial como la cuestión sobre la adecuación de las metas a las necesidades y las consecuencias intencionadas y no intencionadas a medio y largo plazo. Para evitar tales problemas se necesita ampliar el marco conceptual de la evaluación de programas, implicada más en la racionalidad sustantiva o de valor, que proporcione una información más rica y real a los propósitos de los programas y a sus implicaciones en intereses sociales más amplios.

Shadish, Cook y Leviton (1991: 31) defienden la necesidad de fundamentar las teorías evaluativas en cinco componentes teóricos básicos. Son los siguientes:

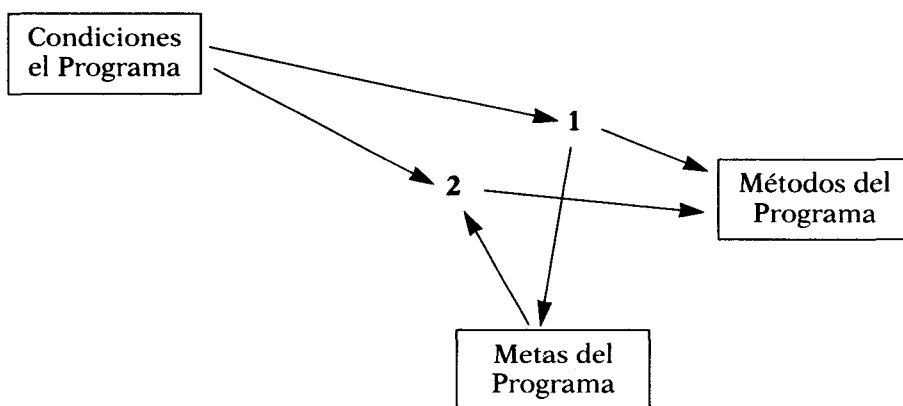
- *Programación social.* El componente relativo a la programación social tiene que ver con la naturaleza de los programas sociales y sus funciones en la solución de los problemas sociales. Se relaciona con la estructura interna y con el funcionamiento de los programas, con sus relaciones con las instituciones políticas y sus actividades y con los procesos a través de los cuales los programas y sus componentes pueden cambiarse para mejorar el rendimiento de programa.
- *La construcción del conocimiento.* Este componente se interesa por la comprensión de evidencia aceptable y los métodos para producir tal evidencia sobre el objeto que va a ser evaluado.
- *Los valores.* El componente del valor está relacionado con el papel que los valores y los procesos de valoración juegan en la evaluación. Está relacionado con la elección de los valores que deben estar representados en una evaluación y con la construcción de juicios de valor sobre los programas sociales.
- *La utilidad del conocimiento.* El componente del uso, o utilidad, tiene que ver con el modo en que el conocimiento de las Ciencias Sociales y de la Educación puede ser utilizada en la política social y en los programas. Está relacionado con los posibles modos de incrementar la utilidad de los resultados de las evaluaciones.
- *La práctica evaluativa.* El componente relacionado con la práctica evaluativa tiene que ver con las actividades que el evaluador hace en su profesión. Estudia el papel de los evaluadores en relación con los que realizan el programa, cómo decidir las cuestiones o preguntas claves, de dónde conseguir u obtener la información para elaborar y responder a tales cuestiones, y qué métodos se deben utilizar para priorizar las cuestiones más relevantes y los temas sobre los cuales las dudas o las incertidumbres son mayores, así como conocer las posibilidades y limitaciones de tiempo, los recursos financieros, las destrezas del personal y las normas para proceder a la realización de la evaluación (Cook y Shadish, 1986, Shadish y Reichardt, 1987 en Shadish, Cook y Leviton, 1991:36).

En una teoría de evaluación de programas, los criterios seleccionados para juzgar la realización de un programa claramente implican juicios de valor. Temas claves relacionados con cómo diseñar un tratamiento o cómo llevarlo a cabo, también implican juicios de valor. Podemos decir que la teoría, los valores, las propias preferencias, e incluso los prejuicios, deben hacerse explícitos. No cabe duda de que cualquier opción estará influida por la teoría –y por los prejuicios.

No hay un único significado del término *teoría* universalmente aceptado. Normalmente connota un cuerpo de conocimientos que organiza, describe,

prescribe, predice, explica con el fin de ayudar a la comprensión y al control de los temas que estudia, aunque las teorías hacen esto de diferentes maneras. No obstante, el significado del término teoría más aceptado es el que la define como un grupo de concepciones interrelacionadas, principios o proposiciones para explicar o guiar acciones sociales.

Las teorías son de dos tipos: descriptivas y prescriptivas. El propósito de las teorías llamadas descriptivas es describir y explicar un fenómeno, un hecho. La teoría prescriptiva se relaciona con lo que se debe hacer o cómo algo se puede hacer mejor. Las teorías descriptivas toman un grupo de condiciones y métodos, o procedimientos, como datos, constantes, y tratan de describir y explicar la probabilidad de los resultados, las metas, como variables de interés. En contraste, las teorías prescriptivas, y los principios prescriptivos se apoyan en unas condiciones y metas y prescriben los mejores métodos, o procedimientos, como variables de interés, para conseguirlas. (Reigeluth, 1983:20).



1. Teorías descriptivas.
2. Teorías prescriptivas

*La teoría de la evaluación de programas integra las dos naturalezas de la teoría: la descriptiva y la prescriptiva. Las teorías prescriptivas incluyen las siguientes características:*

1. Orientación a la acción,
2. diseño del tratamiento y su realización, y
3. un amplio rango de acciones para decidir sobre la elección de los criterios de valoración de los logros.

La teoría de la evaluación de programas es definida por Chen (1990) como *una especificación de lo que se debe hacer para conseguir las metas deseadas, qué otros impactos o logros importantes pueden ser también anticipados, y cómo estas metas y estos impactos pueden generarse.*

La primera parte de esta definición se relaciona con la teoría prescriptiva, también llamada normativa –¿qué se debe hacer?– y la segunda parte se relaciona con la teoría descriptiva, también llamada causativa, en lo que subyace a los mecanismos causales que relacionan los tratamientos del programa, los procesos de realización, los programas de tratamiento, los procesos de implementación y las metas. Chen prefiere utilizar la terminología de *teoría normativa*, en lugar de teoría prescriptiva, y teoría causativa, en vez de *teoría descriptiva*. Ambas subteorías forman parte de la teoría de la evaluación de programas.

El objetivo de las *evaluaciones normativas* es proporcionar información para mejorar la estructura de los programas, los tratamientos, el medio en el que se desarrollan y sus metas. Por otra parte, el principal objetivo de las *evaluaciones causativas* es valorar las relaciones entre los tratamientos, los medios en donde se realizan y los logros.

La teoría normativa de la evaluación de programas debe proporcionar guías sobre qué metas, o qué logros, deben intentar conseguirse y deben ser examinados, y cómo debe diseñarse el tratamiento y planificarse su realización. La teoría causativa especifica cómo trabaja el programa mediante la identificación de las condiciones bajo las cuales ciertos procesos emergerán y cuáles serán las consecuencias (Chen, 1990: 43).

La teoría normativa procede examinando las premisas, las concepciones, los procedimientos, o el conocimiento previo. Además, proporciona una justificación para la estructura y actividades del programa, guía la planificación, la formulación y la realización del programa. De otro lado, la teoría causativa para la evaluación de programas se basa en la experiencia; representa el conocimiento empírico, busca las relaciones causales entre los tratamientos y los resultados. Una teoría causativa especifica los mecanismos causales que vinculan, median o condicionan las relaciones de causa-efecto entre las variables de tratamiento (variables independientes) y los logros del programa (variables dependientes).

Al interesarse y preocuparse por los problemas del día a día del programa, el evaluador se sensibiliza hacia las necesidades de la gente que aplica el programa y potencia la utilización de los resultados evaluativos al no perder el objetivo humano de las necesidades de quienes trabajan en el programa. Además, los patrocinadores y promotores de los programas necesitan la información a tiempo para actuar y tomar decisiones eficaces.

La evaluación desde la teoría normativa requiere de la consistencia entre la estructura teórica del programa y la estructura de funcionamiento del programa para ser valorada. Esto puede realizarse en un estadio temprano del desarrollo del programa, que proporcionará información para diagnosticar pronto los problemas a tiempo y realizar acciones para tratar de que estos problemas se subsanen, fortaleciendo la estructura del programa y/o los procesos de realización.

Cuando el tratamiento de un programa es construido y ejecutado adecuadamente, la credibilidad del programa se potencia. No obstante, esto no lleva implícito que el programa sea eficaz. La evaluación del programa, desde la teoría normativa, es insuficiente, por sí misma, para comprender si el programa consigue las metas para las cuales ha sido diseñado o para conocer las consecuencias no intencionadas del programa. Para comprender la eficacia del programa, la teoría causativa es importante y necesaria.

Una evaluación que se rige por la teoría causativa proporciona información sobre el impacto generado por el programa y sobre cómo dichos impactos fueron generados. El valor de un programa es muy difícil de juzgar sin tener información sobre factores de contexto o de intervención que ayuden a conocer por qué un programa tiene éxito o fracasa.

La especificación de la teoría causativa en una evaluación es especialmente útil para la mejora futura del programa. Cuando la teoría causativa es proyectada en detalle, la evaluación puede conocer las debilidades de los mecanismos causales que subyacen en el programa, identificar los factores contextuales e intervinientes que dificultan o facilitan los procesos del programa y sus resultados, y sugerir posibles estrategias para ocuparse de mejorar sus programas e introducir los cambios necesarios para ello. Ambas teorías, normativa y causativa, son necesarias y constituyen las dos subteorías de la teoría de la evaluación de programas.



## BIBLIOGRAFÍA

- ALKIN, M.C. (1969). "Evaluation theory development", *Evaluation Comment*, 2, 1.
- BICKMAN, L. (Ed.) (1987). *Using program theory in evaluation*. San Francisco: Jossey-Bass.
- CAMPBELL, D.T. y STANLEY, J.C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, ILL.: Rand-McNally (1 ed. 1963). (Traducido al castellano: *Diseños experimentales y cuasi-experimentales en la investigación social*. Buenos Aires: Amorrortu. 1970.
- CHELINSKY, E. y SHADISH, W.R. (1997). *Evaluation for the 21<sup>st</sup> Century*. Thousand Oaks, CA.: Sage.
- CHEN, H-T. (1990). *Theory-driven evaluation*. Newbury Park: Sage.
- CHEN, H-T. y ROSSI, P.H. (1981). *The multi-goal, theory-driven approach to evaluation: a model linking basic and applied social science*. *Social Forces*, 59: 106-112.
- CHEN, H-T. y ROSSI, P.H. (1983). *Evaluating with sense: The Theory-driven approach*. *Evaluation Review*, 7, 106-122.
- CHEN, H-T. y ROSSI, P.H. (1989). *Issues in theory-driven perspective*. *Evaluation and Program Planning*, 12, 299-306.
- CHEN, H-T. y ROSSI, P.H. (Eds.) (1992). *Using theory to improve program and policy evaluation*. New York, London: Greenwood Press.
- COOK, T.D. y CAMPBELL, D.T. (1969). *Quasi-experimentation: Designs and analysis issues for field settings*. Chicago, ILL.: Rand-McNally.
- COOK, T.D. y REICHARDT, CH. S. (1986). *Métodos cualitativos y cuantitativos en investigación educativa*. Madrid: Morata.
- CRONBACH, L.J. (1973). "Remarks to the new society". *Evaluation Research Society Newsletter*, 1, 1, (1-3).
- CRONBACH, L.J. (1982). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- CRONBACH, L.J., AMBRON, S.R., DORN-BUSCH, S.M., HESS, R.D. HORNIK, R.C. PHILLIPS, D.C., WALKER, D.F. y WEINER, S.S. (1980). *Toward reform of program evaluation*. San Francisco, CA: Jossey-Bass.
- CRONBACH, L.J. y SUPPES, P. (Eds.) (1969). *Research for tomorrow's schools: disciplined inquiry for education*. New York: Macmillan.
- DE LA ORDEN, A. (1996). "Evaluación, innovación y calidad educativa". XI Congreso Nacional de Pedagogía: Innovación pedagógica y políticas educativas. San Sebastián: Sociedad Española de Pedagogía y Universidad del País Vasco.
- FETTERMAN, D. M. (1992). "Theory in evaluation: we think, therefore we theorize (An ethnographer's perspective)". En H-t Chen y P. H. Rossi: O.c.
- FULLAN, M. (1994). *Change forces. Probing the depths of Evaluational Reform*. Bristol: Falmer Press.
- GUBA, E.G. y LINCOLN, I.S. (1981). *Effective evaluation*. San Francisco, Washington, London: Jossey Bass.
- GUBA, E.G. y LINCOLN, I.S. (1989). *Fourth generation evaluation*. London: Sage.
- HOUSE, E.R. (Edit.) (1986). *New directions in educational evaluation*. London: Falmer Press.
- LINCOLN, Y.S. (1985). *Organizational theory and inquiry*. London: Sage.
- LINCOLN, I.S. y GUBA, E.G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- LINCOLN, Y.S. y GUBA, E.G. (1986). "But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation", (73-78), En D.D. Williams (Ed.): *Naturalistic evaluation*. New directions for program evaluation, 30, San Francisco: Jossey-Bass.

- LIPSEY, M.W. y POLLARD, J.A. (1989). "Driving toward theory in program evaluation: more models to choose from". *Evaluation and Program Planning*, 12, 317-328.
- MARTÍNEZ MEDIANO, C. (1993). *Factores de eficacia de los Centros Asociados de la UNED*. Madrid: ICI-AIESAD-UNED.
- MARTÍNEZ MEDIANO, C. (1996). "Investigación evaluativa sobre los factores de eficacia de los Centros Asociados de la UNED". *Revista Iberoamericana de Educación superior a Distancia*. AIESAD, Vol. VIII, 1. 25-50.
- MARTÍNEZ MEDIANO, C. (1997). *Evaluación de programas educativos. Investigación evaluativa. Modelos de evaluación de programas*. Madrid: UNED, CU. 159.
- PARLETT, M.R. y HAMILTON, D. (1978). *Evaluation as illumination: A new approach to the study of innovative programmes*. Occasional paper, 9, Centre for Research in the Educational Sciences. University of Edinburgh. 1972.
- PATTON, M.Q. (1980). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage. (ed rev.1990).
- PÉREZ JUSTE, R. (1992). Evaluación de programas de orientación. V Seminario Iberoamericano de Orientación. Actas AEOEP y Departamento MIDE. (46-66).
- PÉREZ JUSTE, R., MARTÍNEZ MEDIANO, C. y RODRÍGUEZ GÓMEZ, G. (1994). *Centros educativos de calidad*. Madrid: ITE.CECE.
- REIGELUTH, CH. M. (1983). "Instructional Design: What is it and why is it?" (3-36). En: C.M. Reigeluth (Ed.). *Instructional-Design. Theories and Models: An overview of their current status*. Hillsdale. N.J. London: Lawrence Erlbaum Associates, Publishers. 487 pp.
- ROSSI, P.H. y FREEMAN, H.E. (1993). *Evaluation: A systematic approach*. (5a ed.). Newbury Park, CA: Sage.
- SCRIVEN, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: Sage.
- SHADISH, W.R. y REICHARDT, C.S. (1987). *Evaluation Studies Review Annual*. Newbury Park, CA: Sage.
- SHADISH, W.R. COOK, T.D. y LEVITON, L.C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.
- STAKE, R.E. (1967). *The countenance of educational evaluation*. Teachers College Record. 68, (523-540).
- STAKE, R.E. (1975). *Programm evaluation, particularly responsive evaluation*. Occasional Paper Series, 5. University of Western Michigan. Evaluation Center.
- STAKE, R.E. (1980). "Program evaluation, particularly responsive evaluation". En W.B.Dickrell y D.Hamilton (Eds.), *Rethinking educational research*, (72-87), London: Hodder and Stoughton.
- STAKE, R.E. (1990). "Responsive evaluation". En H.G. Walberg y G.D. Haertel (Eds.): *The International Encyclopedia of Educational Evaluation*. Oxford, New York: Pergamon Press, (75-77).
- SUCHMAN, E.A (1967). *Evaluation research: Principles and practice in public service and social action programs*. New York: Russell Sage Foundation.
- WEISS, C.H. (1972). *Evaluation research: Methods for assessing program effectiveness*. Englewood Cliffs, NJ: Prentice-Hall.
- WEISS, C.H. (1972). "Evaluating educational and social programs: A treeful of owls". En C.H.Weiss (Ed.): *Evaluating actions programs: Readings in social action and education*. (3-27). Boston: Allyn and Bacon.
- WEISS, C.H. (1975). *Investigación evaluativa: Métodos para determinar la eficiencia de los programas en acción*. México: Trillas.
- WEISS, C.H. y BUCUVALAS, M.J. (1989). "Truth tests and utility tests: Decision-makers' frame of reference for social science research". En H.E.Freeman y M.A.Solomon (Eds.). *Evaluation studies review annual*, 6,(695-706). Beverly Hills, CA: Sage.

---

## RESUMEN

---

*En este trabajo sobre La teoría de la evaluación de programas comenzamos realizando una revisión de la situación actual de la evaluación como disciplina de conocimiento, mencionando sus elementos constitutivos, para centrarnos, a continuación, en la evolución que ha seguido la disciplina de la Evaluación de programas, volcada en un primer momento en los métodos de investigación contribuyendo a su desarrollo adaptado a la problemática evaluativa y, posteriormente, en el desarrollo de la teoría de la evaluación de programas, fundamental en toda disciplina científica, que orienta la elaboración de los programas sociales y educativos, clarifica los procesos y las metas de la evaluación, comprueba su realización práctica y resultados, asesora en la mejora de los programas y contribuye, en definitiva, a la construcción del conocimiento.*

---

**Palabras clave:** Teoría de la evaluación de programas. Evaluación como disciplina de conocimiento. Métodos de investigación adaptados a la evaluación de programas.

---

## ABSTRACT

---

*This paper is about Program Evaluation Theory. It shows, at the beginning, an overview on Evaluation as knowledge discipline, analyzing their constitutive elements. Then, we focus on the development of Program Evaluation, with one special interest, in the 60's, on research methods, with one important contribution to develop appropriate methods to the special evaluation questions and issues, and in the 80's involved on the development of Program Theory, one of the elements constitutive of the knowledge discipline, to guide the construction of social and educative programs, its goals, its implementation and outcomes, and to guide the decision making, to change and improve programs, and thus, to contribute to the construction of knowledge.*

---

**Key words:** Program Evaluation Theory. Evaluation as knowledge discipline. Research methods appropriate to program evaluation.

---