

LA EVALUACIÓN DE PROGRAMAS SOCIALES: FUNDAMENTOS Y ENFOQUES TEÓRICOS

Mario de Miguel Díaz

Universidad de Oviedo

I. INTRODUCCIÓN

Frecuentemente constatamos que en la literatura relativa a evaluación se utilizan algunos conceptos de forma poco precisa lo que genera indudablemente expresiones confusas. Por ello, antes de abordar el tema que nos ocupa, consideramos procedente efectuar ciertas precisiones terminológicas que permitan clarificar el punto de vista que asumimos al realizar este trabajo, así como los parámetros dentro de los cuales debe ser analizada nuestra aportación ya que, de lo contrario, su lectura puede dar lugar a interpretaciones muy distintas a nuestros planteamientos.

Un ejemplo de lo que acabamos de señalar sucede con los términos investigación evaluativa y evaluación de programas. Aunque inicialmente aluden a un mismo campo de trabajo, en sentido estricto, su significado es distinto. Desde nuestro punto de vista, la investigación evaluativa (*evaluation research*) constituye la aplicación de los principios y procedimientos de la investigación social para comprobar sistemáticamente la eficacia de un programa de intervención social. Cuando hablamos de evaluación de programas (*program evaluation*) aludimos al conjunto de principios, estrategias y procesos que fundamentan la evaluación de toda acción o conjunto de acciones desarrolladas de forma sistemática en un determinado contexto con el fin de tomar las decisiones pertinentes que contribuyan a mejorar las estrategias de intervención social. Si el objetivo de un programa es mejorar las condiciones sociales, el propósito de la evaluación es mejorar los propios programas (Rossi, Freeman y Lipsey, 1999). De lo cual se deduce que esta tarea conlleva «algo más» que la mera utilización de metodologías científicas orientadas a evaluar teorías e impactos.

Para percibir con mayor claridad los matices que distinguen ambos términos basta recordar la definición comúnmente aceptada sobre el concepto de evaluación como

«un proceso sistemático de búsqueda de evidencias para formular juicios de valor que orienten la toma de decisiones». De ella se desprende la existencia de tres elementos claves que delimitan esta actividad: *el proceso metodológico que se lleva a cabo para recoger e interpretar las evidencias relacionadas con el programa, los criterios que vamos a utilizar para emitir los juicios de valor y, finalmente, la utilización que posteriormente se va hacer de los resultados*. Por ello, toda aproximación conceptual al campo de la evaluación de programas debe contemplar necesariamente los tres pilares sobre los que se fundamenta esta disciplina —evidencias, valores y decisiones— ya que, de lo contrario, efectuaríamos un reduccionismo injustificado.

Partiendo de este planteamiento, la distinción entre investigación evaluativa y evaluación de programas se percibe de forma más nítida. *La investigación evaluativa constituye esencialmente una estrategia metodológica orientada a la búsqueda de evidencias respecto a un programa, pero la evaluación de un programa requiere, además, abordar otros problemas relacionados con los criterios a utilizar para emitir los juicios de valor y con los procesos relativos a la toma de decisiones*. Se diría, pues, que la primera constituye una de las herramientas metodológicas básicas a utilizar en una de las fases del proceso evaluativo. Ciertamente una fase muy importante —en la medida que nos facilita el soporte a partir del cual vamos a emitir las valoraciones— pero en modo alguno la única.

Hemos considerado oportuno iniciar el presente trabajo matizando esta distinción porque en la bibliografía al respecto es bastante común que a la hora de hablar de paradigmas, modelos o enfoques sobre evaluación de programas se utilicen —más o menos adaptados— aquellos que son propios de la investigación científica social como si ambos términos fueran equivalentes. Desde nuestro punto de vista esta extrapolación adolece de sentido; dado que supone una utilización indiscriminada de conceptos que aluden a dominios diferentes. *No cabe, por tanto, trasladar a la evaluación de programas los supuestos teóricos de los paradigmas clásicos de investigación propios de las ciencias sociales ya que la fundamentación de esta disciplina requiere necesariamente contemplar el mundo de los valores y de las consecuencias*. Por ello, necesitamos abordar este tema desde una óptica más comprensiva.

2. DELIMITACIÓN DEL CONTEXTO DEL TRABAJO

Una de las características dominantes en el campo de la evaluación de programas durante los últimos años ha sido la multitud y heterogeneidad de los conceptos asociados con este término. Al igual que sucede en toda disciplina emergente, constantemente están apareciendo aproximaciones, enfoques, modelos y métodos que acuñan nuevos términos y tipologías lo que hace patente la escasa consolidación epistemológica que presenta actualmente este campo. Presionados por la realidad práctica, los evaluadores se dedican a resolver cuestiones de carácter metodológico y relegan a un segundo plano las cuestiones teóricas relativas a los fundamentos de la evaluación. De ahí que *la escasa atención que se presta a la «teoría» se considere como una de las debilidades más importantes que tiene esta actividad y que impide su consolidación como disciplina científica* (Bickman 1990, Rebien, 1997).

Ante esta carencia, resulta necesario revisar los supuestos teóricos de esta disciplina con el fin de contribuir a consolidar su estatuto epistemológico y ayudar a los evaluadores a que puedan explicitar los criterios que implícitamente utilizan para seleccionar, organizar y llevar a cabo sus métodos de trabajo. La tarea del evaluador no puede limitarse a una simple elección de metodologías y métodos para obtener evidencias. Necesitamos de manera imprescindible una teoría que justifique y explique los procedimientos y las prácticas que en cada caso se llevan a cabo. *La evaluación de programas es algo más que una metodología aplicada y, al igual que sucede en todas las disciplinas, requiere sus propias bases teóricas que justifiquen y den sentido a lo que hace un evaluador, cómo lo hace y, sobre todo, para qué.*

Como ya hemos avanzado, en este campo no abundan trabajos de carácter epistemológico. Se habla mucho de métodos y metodologías pero casi nada de epistemología; es decir, sobre la teoría del conocimiento vinculada a la evaluación de programas. Más aún, en muchos textos no se matizan las diferencias entre fundamentos científicos de la evaluación de programas (*foundations*) y enfoques teóricos aplicables a la evaluación de programas (*approaches*). Evidentemente estos términos no significan lo mismo ya que, mientras el primero hace referencia a las bases que fundamentan la evaluación, el segundo alude claramente a las distintas metodologías que podemos utilizar en la práctica para abordar los procesos. Aunque, lógicamente, ambos conceptos están relacionados.

Por ello, aunque explícitamente no se manifieste la justificación teórica, cada opción metodológica ha sido elegida en función de las concepciones ontológicas y epistemológicas que el evaluador tiene sobre los conceptos de «programa» y «evaluación», lo que significa que existe un posicionamiento implícito. Detrás de cada aproximación metodológica siempre hay un enfoque teórico tácito. Ello permite que, ante un problema concreto, sea posible plantear estrategias de intervención distintas y que evaluaciones con enfoques diferentes sobre una misma intervención puedan llegar a resultados contradictorios (Dennis y Boruch, 1989). Lo cual significa que *todas las aproximaciones tienen sus ventajas e inconvenientes y que es necesario que el evaluador precise, en cada caso, el marco conceptual que asume a la hora de justificar y planificar su trabajo ya que las decisiones metodológicas deben estar siempre supeditadas a los planteamientos teóricos* (Shadish, Cook y Leviton, 1991). Necesitamos abordar la reflexión epistemológica sobre el tipo de conocimiento que nos aporta un proceso evaluativo y, en consecuencia, sobre el marco que justifica el trabajo de un evaluador. Necesitamos más y mejor teoría (Cook, 1997).

Situados en este contexto, el abanico de temas que pueden ser objeto de la reflexión es tan amplio que necesariamente procede cierta delimitación con fines didácticos. Desde nuestro punto de vista, siguiendo pautas de otros autores que se ocupan del tema (Bickman 1990, Shadish, Cook y Leviton 1991, Rebien 1997), la fundamentación teórica de la evaluación de un programa exige una toma de postura, al menos, sobre las siguientes cuestiones básicas: *el concepto de programa como estrategia de intervención social, las bases epistemológicas que determinan la construcción de las evidencias, los criterios a utilizar para formular los juicios de valor, la utilidad o usos que van a tener los informes y el enfoque metodológico a seguir para realizar el proceso evaluativo.* A partir de las opciones y

decisiones que el evaluador asuma en torno a estas cuestiones podremos entender y justificar las razones por las que elige un determinado enfoque teórico y metodológico para evaluar cada programa concreto.

3. CONCEPTO DE INTERVENCIÓN SOCIAL

Toda aproximación al estatuto científico de una disciplina conlleva identificar previamente cual es el objeto del conocimiento. Necesitamos, pues, una definición sobre el concepto «programa». Habitualmente se entiende por programa todo «*plan, proyecto o servicio mediante el cual se diseña, organiza y pone en práctica un conjunto de acciones dirigidas a la consecución de una metas*» (De Miguel, 2000b). Este término aglutina, por tanto, cualquier tipo de estrategia de intervención social orientada a promover cambios. De ahí que la primera reflexión que se debe plantear el evaluador sea sobre cómo deben ser conceptualizadas, diseñadas y aplicadas estas estrategias de intervención para que produzcan cambios sociales efectivos. En definitiva, necesitamos una teoría que fundamente la elaboración y aplicación de las estrategias de intervención social y, en consecuencia, que pueda ser utilizada como marco de referencia para tomar decisiones a la hora de planificar los procesos evaluativos (Bickman, 1987).

El término «teoría» aplicado a la evaluación de programas alude a «*un conjunto de principios interrelacionados que explican y predicen la conducta de una persona, grupo u organización sobre la que hemos llevado a cabo una acción intencional y sistemática*» (Chen, 1990). Lógicamente, existen distintas formas de entender las relaciones entre los principios, las acciones y las conductas en el campo social, por lo que es muy importante conocer cuál es el punto de vista que considera el evaluador en cada caso concreto, ya que constituye el determinante inmediato que justifica la perspectiva o enfoque teórico que adopta para planificar y evaluar una intervención. De ahí que una de las primeras cuestiones que el evaluador debe explicitar sea la concepción teórica desde la que aborda el «programa» en tanto que estrategia de intervención social.

3.1. El programa como verificación de teorías

Una de las concepciones clásicas sobre los programas es considerarlos como estrategias que permiten la verificación de teorías que son útiles para resolver problemas sociales. Quienes adoptan esta postura enfatizan el papel de la teoría como fundamento de toda intervención social por lo que el objetivo prioritario de toda evaluación se orienta a probar las relaciones causa-efecto que se establecen entre unos tratamientos y unos resultados. Ello exige que tanto el diseño como la ejecución de los programas se lleve a cabo de acuerdo con los principios de la metodología científica más rigurosa dado que su objetivo es muy similar a la verificación de hipótesis conceptuales.

Quienes adoptan esta postura —como ya proponía el conocido trabajo de Campbell (1969), *Reforms as Experiments*—, consideran que los cambios sociales pueden ser planificados a modo de experimentos. Los programas constituyen estrategias o artefactos a través de las cuales se someten a prueba si determinadas teorías son eficaces para la resolución de problemas sociales prácticos. Al igual que sucede en todo experi-

mento, el evaluador formula una teoría explicativa sobre una realidad o problema social sobre la cual construye y aplica una estrategia de intervención siguiendo los cánones de la metodología científica (aleatorización y manipulación) con el fin de poder probar su adecuación y eficacia.

Para quienes defienden este concepto, el objetivo principal de una evaluación es lograr identificar los efectos netos de un programa —es decir, aquellos que se deben exclusivamente a las variables explicativas que justifican la intervención— al margen de los efectos brutos. Otros autores prefieren hablar de resultados e impactos, entendiendo por los primeros todos los efectos de un programa y acotando el segundo término solamente para aquellos sobre los que se tenga indicios razonables de causalidad (Schalock 1995, Boruch 1997, Apodaca 1999). Precisamente porque su finalidad primordial es la identificación de los efectos netos o impactos —a corto, medio o largo plazo— asumir este enfoque implica introducir en la evaluación de un programa toda la problemática de la experimentación científica en el campo social, con toda la carga de dificultades y limitaciones que este hecho conlleva. Las características implícitas en toda intervención social impiden la aplicación de diseños metodológicos rigurosos que nos permitan obtener indicios razonables de validez interna y externa de forma satisfactoria (Cronbach, 1982).

3.2. El programa como tecnología sistemática

La realidad práctica nos hace ver que la resolución de problemas sociales constituye más un problema de contingencias que de teorías causales. Generalmente no existen grandes discusiones sobre las teorías o principios que deben inspirar las políticas públicas. Todos sabemos qué procede hacer cuando se quieren abordar problemas relativos a la pobreza, la marginación, la discriminación social, tercera edad, etc... Las diferencias no están tanto en el terreno de las ideas como en la instrumentación práctica de las políticas que se llevan a cabo. Por ello, la evaluación de los programas, antes que los problemas de causalidad de las teorías, se deberá centrar sobre los procesos de instrumentación de las estrategias de intervención ya que son las contingencias que se establecen en cada caso las que —en última instancia— determinan las diferencias.

Este hecho ha supuesto que los enfoques relativos a la evaluación de programas desplazaran el foco de la reflexión hacia los problemas prácticos que conlleva toda estrategia de intervención ya que son los procesos los que, esencialmente, determinan la mayor o menor eficacia de los programas (Wholey, Hatry y Newcomer, 1994). De ahí que a la hora de evaluar sea muy importante distinguir entre los efectos que se deben a al desarrollo de los programas (*action hipotesis*) y los que están determinados por las teorías (*conceptual hipotesis*). Asumir esta distinción implica un cambio de enfoque sobre el concepto de programa ya que puede ser tipificado como una tecnología instrumental que debe ser diseñada y modificada de forma constante en función de su utilidad en relación con los objetivos establecidos.

Desde esta perspectiva los objetivos de la evaluación no se focalizan sobre la verificación de la inferencia causal de las teorías sino sobre las distintas hipótesis prácticas que se pueden formular en cada una de las fases que implica la conceptualización, el

diseño, la aplicación y la utilidad de una estrategia de intervención social. Por ello, un enfoque evaluativo global deberá abarcar objetivos relativos a cada una de estas fases, a saber: a) la estimación del mérito de las metas que se propone cada estrategia, b) las calidades del diseño y planificación de las actuaciones a realizar, c) el grado de cumplimiento y adecuación con que dicha planificación ha sido llevada a cabo (implementación) y d) la calidad y la utilidad de los resultados e impactos generados. La mayoría de los enfoques prácticos recientes coinciden en señalar estos cuatro objetivos como las metas a tener en cuenta en toda evaluación de un programa ya que nos permiten distinguir entre los fallos de la teoría y los fallos que se producen en la cadena de secuencias que implica el diseño y la aplicación de una estrategia de intervención social (Brinkerhoff, Brethower, Hluchyj y Nowakowski, 1985).

3.3. El programa como compromiso con el cambio social

En un tercer enfoque situaríamos aquellos que consideran que la justificación teórica de todo programa es promover un cambio social por lo que éste debe ser, igualmente, el objetivo de toda evaluación. Más allá de los problemas prácticos que conlleva toda estrategia de intervención, debemos asumir que su finalidad primordial es promover la transformación social. Por ello, en la medida que nuestra búsqueda persigue identificar las políticas, programas y acciones más adecuadas a este fin, el evaluador no puede realizar su trabajo sin cuestionar los valores y las consecuencias que se ponen en juego en cada estrategia de intervención. Lo cual significa que la calidad de su trabajo estará mediatizada no sólo por su capacidad técnica sino también por su compromiso en la resolución de los problemas sociales (pobreza, marginación, desigualdad, etc..) y en la lucha a favor de una sociedad más justa e igualitaria (Patton 1994, Fetterman 1994, De Miguel 1999).

La característica fundamental de este enfoque es que el evaluador no se sitúa al margen del programa sino que constituye uno de los elementos del mismo y, en esa medida, asume un papel crítico respecto al cambio social. Frente a los enfoques anteriores en los que desempeñaba el rol de experimentador y programador social, en este caso su trabajo se identifica con la misma finalidad que promueve el programa; es decir, necesita concebir su función dentro del conjunto de acciones que desarrolla el programa para promover el cambio y la mejora social (Nevo 1990, Sonnichsen 1994). Ahora bien, para asumir este papel es muy importante que el evaluador esté comprometido teórica y vitalmente en la resolución de los problemas sociales, situación que no siempre es un hecho.

Independientemente del nivel de aceptación de esta alternativa, lo que sí debemos reconocer es que ha contribuido decisivamente a resaltar dos aspectos realmente importantes en todo proceso evaluativo. De una parte, la utilidad de las evaluaciones debe estimarse en función de las mejoras sociales y la mayoría de las veces éstas no se producen porque no se toman las decisiones pertinentes que se derivan de los informes evaluativos. De otra, para que las evaluaciones tengan consecuencias, es preciso que los evaluadores se preocupen por maximizar el uso de los resultados de sus trabajos analizando y neutralizando, en la medida de lo posible, los factores políticos y con-

textuales que impiden que los sectores implicados tomen las decisiones oportunas. Cuando se asume el rol de agente de cambio, el evaluador no puede eludir ambos problemas, lo cual significa que no puede dar por concluido su trabajo mientras no sean tangibles las consecuencias; es decir, la mejoras generadas a partir de la evaluación.

Las diferencias entre las bases teóricas implícitas en estas tres formas de conceptualizar un programa social son evidentes. De ahí que existan muchas taxonomías que utilizan este criterio para clasificar los distintos enfoques teóricos y metodológicos que se aplican en los procesos evaluativos. Entre estas clasificaciones queremos resaltar dos que nos parecen muy pertinentes. Por un lado, la propuesta de Stufflebeam y Shinkfield (1987) que agrupa los modelos en *seudoevaluaciones*, *cuasievaluaciones* y *evaluaciones verdaderas* en función de la concepción y propósitos hacia los que se dirige el proceso evaluador. De otro, la que recientemente propone Chelimsky (1997) que clasifica los distintos enfoques teóricos en tres grandes perspectivas de acuerdo con la finalidad principal que orienta la evaluación: a) estimar los resultados y la eficiencia de los programas (*accountability perspective*), b) obtener conocimientos específicos sobre el programa como estrategia de intervención (*knowledge perspective*) y c) contribuir a mejorar el programa para incrementar la mejora y el desarrollo social (*development perspective*). Utilizando este criterio de clasificación, este último autor establece un contraste entre los posicionamientos de estas tres perspectivas respecto a nueve categorías o dimensiones que habitualmente utilizan para delimitar los supuestos de un enfoque teórico.

4. LA CONSTRUCCIÓN DEL CONOCIMIENTO

Independientemente del concepto que el evaluador tenga sobre el programa, la primera fase del proceso en toda evaluación es la *«búsqueda de información técnicamente apropiada para determinar el mérito o valor del objeto que vamos a evaluar»*; es decir, la búsqueda de *«evidencias»* a partir de las cuales podemos justificar nuestras valoraciones. Así pues, cualquiera que sea la finalidad que oriente el proceso, el primer objetivo de una evaluación es obtener informaciones pertinentes, válidas y fiables que actúen como soporte para nuestros juicios. Ello implica que, en esta búsqueda, debemos conjugar dos tipos de criterios —utilidad y rigor— que no siempre se tienen en cuenta. Para ello, el evaluador deberá reflexionar sobre el tipo de conocimientos que necesita sobre el programa y cómo puede proceder para su construcción; es decir, concretar a qué denomina *«información técnicamente apropiada»* y, en consecuencia, cómo va a proceder para su obtención e interpretación. Ambas cuestiones están estrechamente relacionadas por lo que las ideas previas que tiene el evaluador sobre las cuestiones epistemológicas serán determinantes a la hora de tomar estas decisiones.

4.1. Bases teóricas del conocimiento

Dado que existen diversos tipos de conocimientos (causación, generalización, predicción, descripción, comprensión, etc.), el evaluador se verá obligado a justificar cuál

les van a constituir el objeto de su búsqueda ya que, en cada trabajo, unos pueden ser más adecuados que otros. La predilección por un tipo específico de metodología sólo puede ser justificada a partir de los supuestos que fundamentan toda teoría del conocimiento; es decir, las bases ontológicas, epistemológicas y metodológicas que sustentan la evaluación de programas como una disciplina inserta en el marco de la investigación social empírica. En la práctica es habitual especificar los planteamientos metodológicos con que se aborda la evaluación de un programa pero rara vez se hace mención alguna a las concepciones sobre la naturaleza de la realidad que es objeto de análisis (bases ontológicas) o sobre los fundamentos para la construcción del conocimiento científico (bases epistemológicas). Los evaluadores normalmente consideran que abordar estas cuestiones es algo tangencial a su trabajo por lo que suelen evitarlas.

Esta postura, aunque es comprensible, es realmente engañosa ya que —como señalan Shadish, Cook y Leviton (1991)— *la opción metodológica elegida por un evaluador implica implícitamente una toma de postura ontológica y epistemológica*. Por ello, a efectos de que no se produzcan graves disfunciones entre lo que se piensa sobre una realidad (programa) y lo que se hace en la práctica para conocerla, es necesario que los evaluadores se habitúen a explicitar las bases teóricas que justifican sus planteamientos metodológicos; es decir, a conectar los métodos que utilizan con los parámetros teóricos que consideran pertinentes para la construcción del tipo de conocimiento —explícito o tácito— que esperan obtener a través del proceso evaluativo.

En la práctica esta situación se resuelve acudiendo a los paradigmas clásicos que se emplean en la investigación social, por lo que habitualmente se utilizan los distintos supuestos o principios que definen cada uno de ellos para justificar los distintos enfoques u modelos evaluativos. Ello determina que con frecuencia se hable de los paradigmas positivista, naturalista y crítico —u otras de las denominaciones al uso— aplicados al campo de la evaluación de programas sin previamente reconceptualizar la idoneidad de los mismos para el tipo de conocimiento que se persigue en cada caso y, en consecuencia, sobre los métodos a utilizar. Entendemos que esta extrapolación constituye un apriorismo injustificado, ya que, de una parte, no tiene en cuenta que en este campo no existe ningún paradigma que resulte intachable —ya que los problemas y dificultades son cuestiones inherentes a todos ellos— y, de otra, supone abordar el tema de la diversidad paradigmática desde las tesis de la incompatibilidad o complementariedad, al margen de las tendencias actuales sobre la unidad (Walker y Evers 1988, Howe 1992).

Precisamente, debido a que ninguno de los paradigmas clásicos ofrece unas bases epistemológicas plenamente satisfactorias al tipo de conocimiento que requiere esta realidad, diversos autores tratan de promover un marco teórico adecuado que justifique esta disciplina que cada día adquiere más relieve. Entre los esfuerzos realizados en esta línea, consideramos muy interesante la propuesta reciente de un «nuevo marco paradigmático» orientado hacia la explicación social realista (Pawson y Tilley 1997) cuyos supuestos teóricos se enmarcan en la tradición que sitúa la explicación científica entre el positivismo y el relativismo (Hesse 1974, Blaskar 1975, Sayer 1984). Avanzar en esta línea nos puede aportar un marco epistemológico más adecuado que acudir a los paradigmas clásicos.

4.2. Naturaleza de las evidencias: datos y opiniones

Ubicados en el plano del realismo, hemos apuntado que el primer objetivo de todo proceso evaluador es la obtención de «evidencias» sobre las que podamos apoyar nuestros juicios. En sentido etimológico se dice que se tiene evidencia de algo cuando se tiene certeza; es decir, «cuando la información que disponemos nos ofrece certidumbre manifiesta y perceptible». Ahora bien, en los procesos evaluativos recogemos informaciones de distinta naturaleza y no todas nos ofrecen el mismo grado de certidumbre. En unos casos son datos —cuantitativos y cualitativos— obtenidos a través de procedimientos sistemáticos —más o menos rigurosos— que nos aportan informaciones precisas sobre diversos aspectos del programa. En otros recabamos las opiniones (percepciones, expectativas, valoraciones, etc...) que nos manifiestan el concepto o juicio que formulan sobre diversos aspectos del programa las audiencias implicadas en el mismo.

El problema radica en a qué tipo de información le atribuimos la categoría de conocimiento cierto, válido y fiable. Algunos evaluadores, inmersos en la lógica positivista, sólo aceptan como certezas aquellos datos «objetivos» que han sido obtenidos mediante la utilización de una metodología científica rigurosa. Este apriorismo les lleva lógicamente a centrar toda búsqueda de información sobre procedimientos metodológicos «duros» en los que se garantiza la validez interna y externa de los datos obtenidos. Digamos que prevalecen los criterios formales a los prácticos, independientemente de su adecuación para la evaluación del diseño, desarrollo y mejora de un programa y de su utilidad para las audiencias implicadas. Con frecuencia un tratamiento y análisis sofisticado de los datos impide que los resultados y los informes de una evaluación sean comprensibles para aquellos que deben tomar decisiones (Cook y Reichardt 1986, Fetterman y Pitman 1986).

Frente a ellos situaríamos a quienes presentan una actitud más pragmática al considerar que el rango de un conocimiento viene determinado por su utilidad, lo cual significa ante todo preguntarse por el tipo de información que se precisa para evaluar cada fase de un programa y, especialmente, la que necesitan los clientes y las audiencias implicadas para tomar decisiones. Cuando se adopta este punto de vista resulta imprescindible recabar «otro tipo de conocimientos» contruidos fundamentalmente a partir de datos «subjetivos» y hacerlo, igualmente, con la adecuada metodología sistemática para que puedan ser considerados como evidencias contrastadas. *De hecho muchos de los trabajos evaluativos a realizar tienen esta finalidad: estimar el parecer o juicio sobre el programa que comparten la generalidad de personas implicadas en el mismo (opinión pública).*

En definitiva, el evaluador deberá precisar a qué tipo de conocimientos les atribuye la categoría de «evidencias». Todos sabemos que existen muchas clases diferentes de conocimientos que pueden ser objeto de nuestra búsqueda sólo que, en cada estudio y fase del proceso evaluativo, se pone el énfasis sobre un determinado tipo o categoría en función de su pertinencia y oportunidad. Ello no significa que unos sean mejores que otros, ya que todos tienen sus ventajas e inconvenientes. Lo aconsejable es abordar la evaluación de un programa integrando distintas informaciones a partir de estudios

y métodos diversos que nos permitan obtener conocimientos más completos, fiables y válidos, porque *«la calidad de un conocimiento se incrementa en la medida que existe un mayor escrutinio público sobre el mismo»* (Shadish, Cook y Leviton, 1991).

4.3. La controversia entre deducción e inducción

Partiendo de las consideraciones anteriores fácilmente se desprende que tan oportunos y adecuados son los procedimientos metodológicos deductivos como los inductivos para la obtención de «evidencias». En los casos en los que buscamos comprobar relaciones entre inputs y outputs u obtener conocimientos generalizables es lógico que procedamos con una metodología deductiva rigurosa que nos facilite una información conclusiva al respecto. Por el contrario, si el propósito no es generalizar una estrategia sino analizar y comprender lo que sucede en cada intervención para contribuir a mejorar el programa, entonces lógicamente resultará más adecuado llegar al conocimiento por vía inductiva a partir de las distintas informaciones e interpretaciones que nos aporten los implicados. No cabe, por tanto, contraponer dos tipos de procedimientos —verificación frente a descubrimiento— porque buscan cosas diferentes.

De ahí que los problemas teóricos que conlleva la construcción del conocimiento en un campo tan complejo como es la evaluación no se resuelvan señalando que se opta por tal o cual metodología cuantitativa o cualitativa. La metodología a utilizar es la resultante de una serie de decisiones teóricas que previamente hay que definir y precisar de forma explícita si queremos justificar nuestros procedimientos. Por ello, una vez que hemos determinado qué clase de conocimientos se necesitan y por qué, será el momento de proceder a concretar las bases teóricas que justifican nuestros métodos. Ello nos permitirá discernir que no todos los métodos son aconsejables para un mismo objetivo y que es muy importante considerar las ventajas e inconvenientes de cada uno de ellos en función de los diferentes propósitos que pretendemos alcanzar con una evaluación. Lo cual significa que, desde el punto de vista epistemológico, no existe ninguna primacía de unos métodos sobre otros. De hecho, *una de las principales características de la evaluación de programas es que sus métodos cubren toda la gama de paradigmas que dominan en la investigación social* (Rossi, Freeman y Lipsey, 1999).

Si se asume esta postura, la consecuencia lógica es que no cabe enfrentar o contraponer las metodologías cuantitativas y cualitativas como más o menos idóneas para la evaluación de programas ya que todas son necesarias. Las características específicas de los programas y de los procesos que conlleva su conceptualización, diseño, implementación y evaluación nos obliga a considerar fenómenos de naturaleza distinta y en contextos diversos para cuyo conocimiento se requieren métodos diferentes en cada caso, por lo que la polémica respecto a los paradigmas metodológicos queda en un segundo plano (Cook y Reichardt 1986, Fetterman, Kaftarian y Wandersman, 1996). De ahí que, *la evaluación de programas constituya actualmente el ámbito de la investigación social aplicada en el que se hace más patente la superación del debate entre metodología cuantitativa y cualitativa, lo cual significa que clasificar los modelos evaluativos según este criterio carece de sentido* (Reichart y Rallis 1994, Cook 1997, De Miguel 1999).

5. EL PROBLEMA DE LOS CRITERIOS DE VALOR

La evaluación de un programa no se limita a una mera recolección, análisis e interpretación de los datos para obtener «evidencias». Una vez que disponemos de la información necesaria comienza la siguiente fase del proceso: formular los juicios de valor. A la vista de los datos obtenidos deberemos emitir nuestra valoración de forma clara, precisando el rango y matices de los juicios (ausencia o presencia de calidad, puntos fuertes y débiles, recomendaciones y propuestas, etc.). Ahora bien, *para proceder a su formulación el evaluador se encuentra con otro problema: concretar los criterios o marcos de referencia que va a utilizar para emitir sus juicios*. Los datos obtenidos sobre un programa son simplemente datos —más o menos probados— pero su valoración depende del criterio elegido. Precisamente, debido a que en la práctica no existen evaluaciones libres de valores, el evaluador deberá especificar en cada caso los criterios que utiliza como punto de referencia para formular sus juicios respecto a los distintos aspectos del programa.

Habitualmente en la literatura se hace la distinción entre los conceptos de mérito y valor como dos marcos de referencia distintos —intrínseco y extrínseco— a la hora de formular los juicios (Guba Lincoln, 1981, 1989, Scriven 1991), términos que a su vez nos remiten a otra clasificación de los criterios de evaluación en dos grandes bloques: criterios descriptivos y criterios prescriptivos (Alkin y Ellet, 1990). Realizamos una valoración descriptiva cuando se elige o asume «valores de otros» como criterio de referencia para emitir los juicios sobre el programa. Por el contrario, denominamos una valoración como prescriptiva cuando es el propio evaluador el que define los criterios que considera adecuados y correctos para emitir los juicios, al margen de que éstos sean o no coincidentes con los que defienden otras audiencias.

Las diferencias entre ambos enfoques se perciben fácilmente. Cuando realiza una valoración descriptiva nuestro papel se limita a preguntarles a los demás los criterios a utilizar y aplicarlos con rigor y corrección. En cambio, optar por una valoración prescriptiva exige decantarse por unos criterios frente a otros lo que supone previamente adoptar una postura personal ante todas las cuestiones relacionadas con los valores que se ponen en juego cada programa. En el primer caso, al aceptar los objetivos de los demás como criterios de valoración, el evaluador puede conservar su independencia y mantenerse al margen de intereses sectoriales. En el segundo, dado que no existe una teoría general que pueda considerarse correcta sobre la valoración prescriptiva, es muy problemático que un evaluador —especialmente cuando actúa como externo— utilice sus propios criterios como marcos de referencia para evaluar un programa y que estos sean aceptados por todas las partes (Vedung, 1997).

Por esta razón, en la práctica de las evaluaciones relativas a programas sociales es más común utilizar criterios descriptivos que prescriptivos. Aunque estos últimos, en tanto que defienden la primacía de valores absolutos (igualdad, justicia, calidad intrínseca, etc...) cumplen una función muy importante a lo largo de todo el proceso evaluativo, son realmente las teorías descriptivas las que nos proporcionan criterios más legítimos y útiles tanto desde el punto de vista social como técnico (Shadish, Cook y Leviton, 1991). De ahí que la mayoría de los evaluadores defiendan la utilización de este tipo de criterios para emitir los juicios de valor sobre los programas.

No obstante, el problema no concluye al señalar que se asume como referencia una valoración descriptiva ya que los «otros» no son un conjunto homogéneo sino que tienen, o pueden tener, intereses diversos y, en ocasiones, contrapuestos. Necesitamos, por tanto, señalar cuál es la perspectiva valorativa externa que utilizamos para emitir los juicios; es decir, especificar el criterio que elegimos para juzgar que un programa es «bueno». En este sentido cabe considerar que existen tres grandes opciones para el evaluador según utilice preferentemente criterios políticos, democráticos o técnicos.

5.1. Criterios políticos

Todos conocemos que la mayoría de los planes y programas de intervención social están patrocinados por organismos y servicios públicos. Ello significa que tanto su concepción como el diseño y la aplicación normalmente responden a determinados supuestos ideológicos e intereses de la política que los sustenta. Los programas constituyen formalmente la concreción de unas decisiones políticas respecto a unas prioridades en materia de actuación social, lo que indudablemente supone apostar por un tipo de valores. Esta apuesta no sólo se traduce en promover un determinado tipo de programas sino también en todas las decisiones que implica la selección y priorización de estas estrategias, especialmente las relativas a la redistribución y asignación de los recursos.

Cabe, por tanto, hablar de un tipo de evaluación que asume como criterios de valoración los objetivos —políticos, económicos, sociales, etc.— que se proponen alcanzar los organismos y agencias que patrocinan los diversos programas. El evaluador, en estos casos, acepta los valores del patrocinador (gestores, administradores, empresarios, etc...) como punto de referencia para emitir sus juicios y así poder facilitarle la información que necesita para que pueda tomar las decisiones pertinentes. Este hecho no implica en modo alguno que el evaluador, tanto si centra su tarea sobre los objetivos cubiertos como sobre los recursos empleados, pierda su objetividad convirtiendo su trabajo en una tarea burocrática al servicio de un estamento político. Al contrario, sus análisis y recomendaciones pueden cuestionar tanto los criterios que inspiran las políticas públicas como las acciones desarrolladas y, en consecuencia, promover cambios internos importantes.

5.2. Criterios democráticos

Frente a la utilización exclusiva de los valores de los «estamentos políticos», cabe pensar que quienes mejor conocen la oportunidad y efectividad de toda estrategia de intervención social son la personas directamente implicadas en la misma por lo que sus criterios deben constituir un marco de referencia preferente para formular los juicios de valor. Al margen de los planteamientos de los gestores, debemos considerar que todas las personas que constituyen las audiencias de un programa generan —en menor o mayor grado— una serie de expectativas, percepciones y valoraciones intrínsecas que es necesario tener en cuenta a la hora de emitir las valoraciones sobre su oportunidad y eficacia. Más aún, en tanto que es algo que les afecta directamente, no

sólo tienen derecho a ello sino que no podemos prescindir de sus criterios para valorar el programa (De Miguel, 1995). Cuando el evaluador asume este marco de referencia decimos que realiza una valoración descriptiva pluralista basada en criterios democráticos.

Ahora bien, para que el evaluador asuma este tipo de perspectiva valorativa es necesario que previamente haya enfocado todo el proceso evaluativo de forma participativa; es decir, teniendo en cuenta que quienes están directamente implicados en un programa son los que deben ser los actores principales de su evaluación. Los argumentos para ello están claros. De una parte, son los que conocen más de cerca la realidad del mismo y, por tanto, pueden facilitarnos informaciones más pertinentes e interesantes. De otra, son los más implicados en las consecuencias y difícilmente se podrán poner en marcha mejoras si ellos no participan activamente en su valoración y elaboración. Todas las propuestas y postulados actuales sobre evaluación pluralista insisten en esta línea: *la oportunidad y conveniencia de utilizar criterios democráticos a la hora de diseñar, aplicar y evaluar cualquier estrategia de intervención social* (Fetterman, Kaffarian y Wandersman 1996, Floc'Hlay y Plottu 1998, Apodaca 1999).

5.3. Criterios técnicos

En una posición intermedia entre la política y la ciudadanía situaríamos la valoración técnica; es decir, aquella que lleva a cabo el evaluador mediante la aplicación de procedimientos e instrumentos científicos apropiados. Los procesos de evaluación conllevan la aplicación de una serie de reglas y principios sobre estrategias de intervención social cuya correcta utilización contribuye a clarificar tanto la adecuación de los programas como la calidad de sus resultados (Martin 1993, Martin y Kettner, 1996). Cuando el evaluador asume este enfoque utiliza la propia teoría normativa sobre evaluación de programas como criterio o marco de referencia a la hora de formular sus juicios de valor; es decir, la calidad de las metas propuestas, la coherencia del diseño efectuado, la adecuación de los procesos de implementación, la efectividad de los resultados e impactos generados y la oportunidad y viabilidad de las decisiones tomadas en relación con los procesos de mejora.

No obstante, el problema no se resuelve señalando que se van a utilizar criterios técnicos ya que supone dejar los marcos de referencia demasiado abiertos. De ahí que el evaluador deba precisar también las normas, estándares e indicadores que utilizará en cada trabajo para valorar los diferentes aspectos del programa. En unos casos para establecer estas normas podrá acudir a referencias nacionales e internacionales al uso (estándares de calidad, normas internacionales, protocolos sobre buenas prácticas, etc.). En otros se verá obligado a establecer criterios y elaborar indicadores que le permitan objetivar el proceso de valoración y justificar ante las audiencias los juicios emitidos. La dificultad que entraña esta tarea determina que sea una de las cuestiones técnicas más problemáticas y controvertidas en todos los procesos evaluativos (De Miguel, 1991).

Dado que cabe utilizar diferentes perspectivas valorativas como marco de referencia para formular los juicios, se pueden igualmente establecer taxonomías evaluativas

utilizando este criterio de clasificación. Así, algunos autores hablan de *modelos descriptivos y modelos prescriptivos* (Alkin y Ellet, 1990), y otros de *evaluación burocrática, auto-crática y democrática* (Mac Donald, 1977). También se distingue entre *modelos preordenados o convencionales y modelos pluralistas o participativos* según el grado de implicación de las audiencias en los procesos de valoración (Lincoln y Guba 1985, Guba y Lincoln 1989, Rebien 1997). Ahora bien, independientemente de la oportunidad y adecuación de estas clasificaciones, lo que sí debemos constatar es que el cuestionamiento actual de criterios políticos y técnicos ha generado una fuerte tendencia a favor de los criterios democráticos que se manifiesta en todos los enfoques evaluativos actuales —*evaluación pluralista, evaluación participativa, evaluación respondente, evaluación democrática, cluster evaluation, empowerment evaluation, etc.*— cuyo denominador común se caracteriza por concebir el trabajo del evaluador como un compromiso con los derechos de todas las partes con el fin de contribuir a la educación cívica, el desarrollo humano y la emancipación social (Fetterman 1994, Smith 1998).

6. UTILIDAD Y CREDIBILIDAD DE LOS INFORMES

Toda evaluación debe concluir con la toma de decisiones que se deriven de las valoraciones formuladas. Aunque el objetivo inmediato de todo proceso es ofrecer juicios razonados para la toma de decisiones, la finalidad última de toda evaluación es lograr que tenga consecuencias. Precisamente, porque todos esperamos que nuestro trabajo sea útil para resolver problemas sociales, el evaluador no pueda prescindir del análisis y explicación de las relaciones que existen entre los resultados y recomendaciones que emite en sus informes, las decisiones que se toman al respecto y las consecuencias que éstas tienen socialmente. No preocuparse por estas cuestiones pone en tela de juicio la credibilidad de nuestro trabajo. Por ello, la reflexión teórica que conlleva el diseño y aplicación de una estrategia evaluativa debe abarcar también el análisis de los factores —cómo, cuándo, dónde, para qué—, que condicionan la utilización posterior de los resultados que se aportan en los informes (Rebien 1997, Chelimsky 1997).

Ahora bien, abordar el tema de cómo se puede maximizar el uso de los informes evaluativos (*knowledge use*) no es una tarea fácil ya que son múltiples los problemas y factores implicados. No obstante, cabría agrupar esta problemática en torno a cuatro preguntas o interrogantes fundamentales: ¿Qué clases o tipos de usos pueden tener nuestros informes?, ¿Quiénes son los usuarios potenciales de los mismos?, ¿Qué papel puede jugar el informe para estos usuarios?, ¿Cuáles son los factores políticos y contextuales que pueden dificultar la utilización de los resultados? Efectuaremos algunos comentarios sobre estas cuatro cuestiones.

6.1. Posibles usos del Informe

Abordar el tema de la utilidad o usos del informe nos remite nuevamente a la finalidad que promueve y orienta todo el trabajo evaluativo, ya que entre ambas cuestiones existe una mutua implicación. Como ya hemos comentado, existen propuestas

diferentes en relación a los fines de una evaluación aunque, tomando como referencia los usos, se podrían agrupar las diversas alternativas en dos bloques: a) los que entienden que la finalidad primordial de una evaluación es la búsqueda de «observaciones y explicaciones generalizables» mediante la utilización de procedimientos científicos rigurosos y b) los que consideran que lo verdaderamente importante es facilitar a todos los implicados la información que necesitan para tomar decisiones (stakeholders services). En el primer caso se conciben el informe evaluativo de forma similar al de una investigación por lo que los problemas de difusión, diseminación y utilización de los conocimientos son muy afines a los que suceden en otras disciplinas sociales y que se conocen a través de la literatura especializada (Keeves, 1988). En el segundo, dado que cada evaluación de un programa es una tarea singular, el evaluador se ve obligado plantearse en cada proyecto cuál puede ser la utilización posterior o usos posibles que puede tener la información que se aporta en los informes. Al respecto, las demandas y usos más frecuentes que se hacen de los informes evaluativos son los siguientes: *para tomar decisiones, identificar problemas, inducir acciones y justificar decisiones* (Braskamp y Brown 1980, Ewell 1989).

6.2. Identificar los usuarios potenciales de la información

Precisamente porque entendemos que la finalidad prioritaria en toda evaluación de programas es facilitar a las audiencias la información que necesitan, consideramos que, inicialmente, el evaluador deberá identificar quiénes son los usuarios potenciales de su trabajo ya que de ello dependen los tipos de usos que pueden tener sus informes. Deberá tener en cuenta, además, que aunque los trabajos sean objeto de contrato no significa en modo alguno que los informes tengan utilidad solamente para los clientes. De ahí que una de las primeras cuestiones para enfatizar la utilización de los resultados sea la identificación de los usuarios potenciales de nuestro trabajo con el fin de ofrecerles la información que necesitan, lo que implica, previamente, analizar los problemas específicos que se dan en cada contexto y programa: audiencias implicadas, las posibilidades que tienen de acceso a la información, tipos de informes a elaborar, etc.. En toda evaluación existen muchos «implicados» que, en mayor o menor medida, pueden hacer uso de los informes siempre que estos les ofrezcan informaciones oportunas y asequibles Aunque algunos estén más comprometidos que otros en los procesos de toma de decisiones ello no significa que los trabajos no les sean útiles. En la medida que hacemos llegar informes adecuados a todas las partes implicadas en un programa —gestores, administradores, agencias, instituciones, científicos, técnicos, implicados, ciudadanos, etc.— incrementamos las posibilidades de que nuestro trabajo tenga mas consecuencias.

6.3. Clarificar el nivel de operatividad del informe

Además de identificar los usuarios potenciales de nuestro trabajo también tenemos que plantearnos el papel que va jugar el propio informe en los procesos de toma de decisiones. Aunque existen muchas propuestas al respecto, siguiendo a Shadish, Cook

y Leviton (1991), *cabe establecer tres tipos de planteamientos donde los informes se utilizan con criterios diferentes: instrumental, conceptual y persuasivo*. Se considera que un informe es instrumental cuando los resultados obtenidos se utilizan de forma directa para tomar decisiones sobre el programa. Se considera como informe conceptual aquel cuyas valoraciones no suelen ser utilizadas como soporte para decidir cambios pero que, indirectamente, influyen sobre cómo perciben y valoran las personas las intervenciones sociales. Finalmente se entiende que el informe cumple una función persuasiva cuando las evaluaciones son utilizadas para convencer a las audiencias que tienen estimaciones apriorísticas sobre los programas de que la evaluación es el proceso correcto para valorar cualquier estrategia de intervención. Precisar estas cuestiones previamente no sólo nos ayudará a elaborar informes más operativos y útiles desde el punto de vista práctico sino también a conceptualizar la función o papel que le otorgan a nuestro trabajo las partes implicadas en una evaluación. Con el fin de evitar confusiones, algunas de estas cuestiones deben quedar explícitas en el protocolo que se elabore en la fase de negociación de proyecto.

6.4. Elaboración y difusión adecuada de los informes

Finalmente, para maximizar el uso de los informes también debemos considerar otros problemas que tienen que ver con el proceso de elaboración y difusión de los resultados. Respecto a la elaboración, el evaluador deberá preocuparse de facilitar a cada audiencia la información precisa, con datos que pueda fácilmente interpretar y con un lenguaje adecuado. De ahí que en este campo no sea oportuno utilizar técnicas de tratamiento y análisis de datos complejas cuyos resultados son difíciles de valorar salvo para especialistas. Los problemas de la difusión generalmente depende de factores de orden político que se generan en cada contexto de forma especial. En ocasiones, quienes encargan la evaluación se adueñan del informe y sólo lo difunden en la medida que beneficia a sus intereses. En el sentido opuesto situaríamos los planteamientos que consideran que todos los informes deben ser difundidos en términos inteligibles para todas las audiencias con el fin de que todos puedan participar y valorar tanto la utilización que se hace de los resultados como de las decisiones que se toman al respecto. Aunque las posibilidades del evaluador al respecto son bastante limitadas, ello no significa que no deba tener en cuenta estos problemas (McLaughlin y McLaughlin 1989, Wholey, Hatry y Newcomer 1994).

En resumen, aunque inicialmente la utilidad de una evaluación depende de los propósitos, audiencias e intereses en función de los cuales se establece, en la práctica son los procesos de toma de decisiones los que determinan finalmente las consecuencias que tiene cada proceso. Por ello, la tarea del evaluador no concluye con la elaboración del informe sino que debe reflexionar sobre los problemas que genera su utilización para que tenga consecuencias. Asumir esta perspectiva implica cuestionar no sólo qué información necesitan quienes deben tomar decisiones (evaluación orientada a la decisión) sino también la que deben conocer todos los implicados (evaluación orientada al usuario). No es lo mismo pensar que las decisiones dependen de los gestores y administra-

dores de los programas que involucrar también a las audiencias, clientes e implicados. Tampoco es igual que el informe aporte datos concluyentes en términos técnicos o se faciliten informaciones y recomendaciones adecuadas a las diversas audiencias. Fácilmente se advierte que la utilidad de una evaluación es mayor en la medida que tienen posibilidad de acceso a la información todos los sectores implicados.

De lo cual se deduce que el problema de la utilidad y usos de los informes es una cuestión compleja que depende de factores diversos (identificación de usuarios potenciales de los informes, claridad de los resultados, análisis de las alternativas, procesos de toma de decisiones, etc.), que necesitan ser analizados con cierto detalle para evitar obstáculos que impiden su utilización. De ahí que, durante los últimos años, hayan aparecido diversas corrientes y enfoques metodológicos que centran su preocupación sobre las consecuencias de los informes: *decision-oriented evaluation* (Stufflebeam y al., 1971), *utilization-focused evaluation* (Patton, 1978) *stakeholder-based evaluation* (Bryk, 1983), *beyond decision-oriented evaluation* (Sheerens, 1987), *outcomes-based evaluation* (Schalock, 1995), etc.. Aunque estos enfoques presentan entre ellos matices distintos, el denominador común de todos ellos es la necesidad de abordar los procesos evaluativos teniendo en cuenta que la credibilidad del evaluador se halla comprometida por la utilidad y usos que se derivan de sus informes.

7. DELIMITACIÓN DEL ENFOQUE METODOLÓGICO

Al iniciar el presente trabajo hemos distinguido entre fundamentos y enfoques evaluativos como dos términos relativos a la evaluación del programa que tienen un significado diferente. También hemos dicho que los evaluadores suelen ocuparse más de precisar los segundos que los primeros por lo que hemos considerado oportuno abordar algunos conceptos relacionados con los fundamentos porque, desde nuestro punto de vista, éstos deben ser objeto de reflexión antes de optar por un determinado enfoque metodológico. La elección de una aproximación o enfoque específico no se limita a una simple distinción entre metodología cuantitativa y cualitativa ya que supondría una simplificación absurda de una realidad compleja que tiene muchos matices y requiere ser abordada de forma comprensiva.

En la práctica, la determinación de un enfoque metodológico (*approach*) se conoce como la elección del modelo y el diseño de una evaluación. Cuando hablamos de «modelo» hacemos referencia a la forma o modo de abordar y representar la conceptualización teórica del programa que consideramos oportuna para llevar a cabo su evaluación una vez que hemos analizado los supuestos o fundamentos teóricos que justifican esta actividad: naturaleza del programa, la finalidad que nos proponemos, audiencias implicadas, información necesaria, metodología a utilizar, criterios de evaluación, utilidad o usos de los resultados, etc. Por el contrario, el término «diseño» alude al proceso que deberemos seguir para obtener las informaciones que reúnan la categoría de «certezas». El objetivo de los modelos es lograr una representación teórica sobre los fenómenos a investigar y los procesos a realizar que sea coherente y adecuada a la naturaleza del programa y los propósitos de nuestro trabajo. La finalidad de los diseños es suministrar a los investigadores una lógica científica que aporte validez tanto a los

procesos de evaluación como a los resultados. Así pues, mientras que el modelo alude al marco teórico que justifica el enfoque que hemos elegido para evaluar un programa, el diseño hace referencia a la estrategia metodológica que vamos a utilizar para que los datos obtenidos puedan ser considerados como evidencias (De Miguel, 2000).

7.1. La elección del modelo evaluativo

Al igual que sucede en toda investigación social, la elección de un modelo para evaluar un programa implica tomar decisiones sobre una serie de cuestiones teóricas relativas al tipo de estrategia —programa— que se pretende evaluar, las bases epistemológicas que justifican la construcción de los conocimientos, los criterios de valoración y la finalidad o usos que justifican la evaluación. Las preguntas claves que se debe formular un evaluador antes de decidir el marco teórico en el que va a situar su trabajo son las siguientes: ¿Qué quiero saber o qué quiere saber el cliente que me encarga la evaluación en relación con el programa? ¿De qué tipo de programa estamos hablando?, ¿Qué pretende hacer y/o decidir el cliente respecto al programa objeto de estudio? ¿Qué tipo de información necesito? ¿Cuáles son los procedimientos idóneos para su obtención? ¿Cuál es el papel que voy a desempeñar?, etc... Las respuestas a estas cuestiones son las que realmente determinan el enfoque o aproximación metodológica y, en consecuencia, las tácticas y estrategias que justifican un proceso evaluativo que debemos adoptar.

7.1.1. Finalidad del estudio evaluativo

El primer paso para elegir un modelo evaluativo es identificar los fines que se pretenden alcanzar. La identificación de los fines se puede realizar mediante dos procedimientos: según el tipo de función que persiguen los estudios y de acuerdo con los objetivos concretos que se plantean. Respecto a lo primero, actualmente se considera que los procesos evaluativos se orientan a una de estas tres funciones básicas: la rendición de cuentas (*accountability*), la mejora de los programas (*development*) y el incremento de conocimientos respecto al mismo (*knowledge*) (Chelimsky, 1997). Desde el punto de vista de los objetivos los principales propósitos que determinan la petición de procesos evaluativos se pueden agrupar en tres grandes bloques, a saber: a) Comprobar si los objetivos de los programas han sido alcanzados, b) Recabar opiniones fundamentadas sobre la valoración que hacen del programa las audiencias implicadas, y c) Obtener informaciones cualificadas a partir de las cuales se pueda tomar decisiones sobre el futuro de un programa específico (Vedung 1997). A estas tres alternativas se añade una cuarta cuando el objetivo es estimar la relación entre los recursos empleados y los resultados obtenidos (coste-efectividad, coste-beneficio) que da lugar a los modelos económicos (Lewin 1983, Drummond, Stoddart, y Torrance 1991, Nas 1996).

El evaluador deberá precisar cuál o cuáles son los propósitos concretos que realmente promueven cada trabajo con el fin de adoptar un marco teórico adecuado. Aunque ciertamente cada persona puede generar una comprensión «peculiar» de los intereses de los patrocinadores y las necesidades de los clientes y, en consecuencia,

establecer con libertad los parámetros teóricos que considera apropiados para enfocar el proceso evaluativo, procede señalar que «el modelo a utilizar depende del tipo de información que se solicita y que el evaluador debe ajustar su elección a la finalidad que justifica el encargo». En otros términos, los modelos deben ser elegidos en función de los propósitos del estudio.

7.1.2. Concreción de los parámetros del modelo

Una vez que hemos concretado los fines procede delimitar el resto de dimensiones, factores y variables que intervienen en la determinación y representación de un modelo. Como ya hemos comentado en otro momento (De Miguel, 1989), para que un modelo pueda servir de guía, además de estar fundamentado sobre una teoría, necesita concretar los fenómenos que intervienen y las relaciones que se establecen entre los mismos. Exige, además, especificar operacionalmente las variables con el fin de probar la validez empírica del modelo y la posibilidad de generalizar su aplicación. En el caso de la evaluación de un programa conlleva, además, una toma de postura sobre el papel de los valores y rol que el evaluador va a desempeñar ante todos los factores implicados en el proceso a los que hemos aludido anteriormente.

Así pues, una vez clarificada la finalidad y antes de tomar una opción metodológica, el evaluador deberá reflexionar y explicitar los distintos supuestos o fundamentos que adopta en cada caso concreto y que constituyen los determinantes inmediatos en la elección de modelo. A la vista de la literatura (Shadish, Cook y Leviton 1991, Rebieñ 1997), no existe consenso entre los autores sobre cuáles deben ser los factores o dimensiones que se consideran básicas para la construcción de un modelo evaluativo, como paso previo para su especificación. Nuestra propuesta al respecto es que, al menos, se deben considerar los parámetros que hemos abordado anteriormente, a saber: *la conceptualización del programa como estrategia de intervención social, las bases epistemológicas que determinan la construcción de los conocimientos, el tipo de criterios a utilizar para la formular los juicios de valor, la utilidad o usos de que pueden tener los informes y el papel a desempeñar por el evaluador a lo largo del proceso*. Consideramos que estas dimensiones aglutinan las variables o características más importantes a partir de las cuales se puede definir un modelo evaluativo

7.1.3. Utilización de las taxonomías

Generalmente los evaluadores no construyen un modelo específico para cada evaluación sino que utilizan como marco de referencia alguno de los que ya están formalizados en algunas de las taxonomías al uso (MacDonald 1977, House 1980, Guba y Lincoln 1981, Madaus, Scriven y Stufflebeam 1983, Worthen y Sanders 1987, Stufflebeam y Shinkfield 1987, Stake y Mabry 1997, etc...). El problema que presentan estas taxonomías es que están elaboradas con distintos criterios de clasificación que impiden cotejar y contrastar las distintas categorías que establecen. La elección de una u otra como marco de referencia ya implica una toma de postura del evaluador que si no se hace de forma razonada puede, en principio, generar incoherencias al no existir

correspondencia entre los supuestos teóricos del modelo elegido y el tipo de evaluación que procede realizar en función de la finalidad.

Entre estas taxonomías nos interesa destacar de forma especial la propuesta por Stecher y Davis (1991) ya que —desde nuestro punto de vista— nos parece realmente útil pues no sólo tiene en cuenta como criterio de clasificación la finalidad de la evaluación sino que también utiliza para su elaboración dimensiones muy semejantes a las que nosotros proponemos: la naturaleza del programa, las preguntas o cuestiones básicas a resolver, la información que necesitamos, el eje o núcleo sobre el que se centra el proceso y el rol que el evaluador desempeña en el proceso. A partir de estos criterios, *estos autores agrupan los modelos evaluativos en cinco enfoques —experimental, dirigido a metas, orientado a las decisiones, centrados sobre los clientes y respondente— que pueden ser considerados como arquetipos representativos de las diversas opciones o aproximaciones metodológicas que se pueden plantear a la hora de abordar la evaluación de un programa.* Desde nuestro punto de vista, esta clasificación de los modelos nos parece realmente interesante y útil como ya hemos comentado en otro momento (De Miguel, 2000).

7.2. La elección del diseño y los métodos

Una vez que hemos decidido el enfoque o modelo que vamos a asumir para llevar a cabo la evaluación procede establecer el diseño; es decir, «la estrategia o plan que vamos a seguir donde se especifica la secuencia de decisiones acerca de cómo recoger, ordenar y analizar los datos en función de los fines de programa» (Anguera, 1992). De ahí que el término «diseño» se utilice como sinónimo de plan o estrategia metodológica que orienta el proceso evaluativo, lo cual implica realizar una planificación detallada sobre los aspectos o dimensiones del objeto a evaluar que vamos a considerar, los procedimientos que vamos a utilizar para recabar y analizar la información pertinente, y los sistemas a emplear para dar validez a las informaciones que van a sustentar los juicios de valor que se formulen.

Al igual que sucede con los modelos, también podemos utilizar diversos criterios a la hora de clasificar los diseños evaluativos (Ruttman 1984, Fitz-Gibbon y Morris 1987, Rossi y Freeman 1989, Scriven 1991). Entre estos criterios queremos destacar los tres que se utilizan con mas frecuencia:

7.2.1. En función de enfoque teórico adoptado

Utilizando este criterio los diseños se suelen clasificar de forma dicotómica en función de que el marco de la planificación se establezca tomando como referencia la evaluación de los efectos del programa o el significado que éste ha tenido para los implicados. Esta es la postura de adoptan algunos autores (Guba y Lincoln 1981, Cronbach 1982, Patton 1987, Rebien 1997) al plantear dos tipos de diseños frontalmente opuestos —*preordenado y respondente*— como representativos de dos formas diferentes de abordar los procesos evaluativos relativos a los programas sociales. Una evaluación preordenada parte de un modelo teórico formalizado que se especifica mediante un

plan en el que se detallan todos los elementos del proceso: objetivos, audiencias, información necesaria, instrumentos, fuentes, técnicas de análisis, criterios de comparación, etc. Por el contrario, la evaluación respondiente se caracteriza por partir de marco teórico flexible que busca las valoraciones de las personas implicadas en el programa para lo cual es necesario efectuar una planificación emergente que se pueda reconducir en cada etapa del proceso en función de las observaciones que se deduzcan de la etapa anterior.

7.2.2. El sistema de control que permitan las variables del programa

Desde un punto de vista metodológico estricto, el concepto de diseño alude a la «estrategia mediante la cual se pretende probar el efecto de una o más variables manipuladas»; es decir, un plan que nos permita constatar los efectos de las intervenciones promovidas mediante un programa. De ahí que cuando se utiliza el concepto de diseño con terminología positivista, su ordenación y clasificación se establece a partir de los tres requisitos o condiciones que debe reunir todo experimento, a saber: aleatorización de los sujetos, manipulación de la variable independiente y grupo de control. Si trasladamos estos requisitos a la metodología de evaluación de programas, los diseños evaluativos puede ser clasificados en cuatro grandes grupos en función del grado o nivel de garantía con que se cumplen estos requisitos. Así, de acuerdo con Anguera (1992, 1995), los estudios evaluativos se pueden clasificar en función de la metodología utilizada en los siguientes tipos: *observacionales*, *selectivos*, *cuasiexperimentales* y *experimentales*. Consideramos que a esta clasificación, aunque es muy comprensiva, se debería añadir la metodología centrada sobre *costos-beneficios*, cuyos diseños se utilizan cuando la finalidad del proceso se orienta a estimar la eficiencia del programa.

7.2.3. El sistema de control que se utiliza sobre la información recabada

Finalmente elegir un diseño implica determinar el procedimiento mediante el cual vamos a contrastar la información que recabamos sobre el programa —global o parcialmente— a fin de obtener las evidencias necesarias. Para ello no sólo cabe recurrir a los procedimientos propios de los estudios experimentales sino que también existen otros sistemas de control que —aunque no sean tan potentes— pueden ser más adecuados y oportunos para ser utilizados en determinadas fases y/o tipos de programas. En esta línea positivista, algunos autores (Rossi y Freeman, 1989) clasifican los diseños utilizando como criterios la existencia de grupo de control y el procedimiento utilizado para ello. Así hablan de diseños comparativos o de cobertura parcial mediante controles aleatorios, contruidos y estadísticos, y diseños no comparativos o cobertura total que utilizan controles genéricos, reflexivos y en la sombra. Desde enfoques participativos los sistemas de control no se fundamentan tanto en la estructura del diseño como en las técnicas (juicios de expertos, comisiones ad hoc, triangulación, etc..) que se utilizan para contrastar y consensuar la información procedente de las diversos estudios y audiencias (Pineault y Davaluy 1989, Apodaca 1999).

7.3. Correspondencia entre modelos y diseños

Hemos avanzado que entre modelos y diseños existe una interacción evidente que se hace igualmente patente en el caso de la evaluación de un programa. Establecido el modelo, la elaboración del diseño constituye, en parte, una consecuencia lógica. Cuando la evaluación se enfoca desde una aproximación positivista lo lógico es que elaborem un diseño preordenado y que utilicemos una metodología lo más coherente posible con los supuestos de este paradigma aplicando sistemas de control —a ser posible— comparativos. Por el contrario, si nuestro enfoque se establece desde posiciones más pluralistas lo normal es que la planificación del diseño se organice de forma más flexible fundamentalmente a partir de estudios observacionales y utilizando procedimientos de control participativos y no comparativos.

Así pues, dado el abanico de modelos y diseños posibles, el evaluador deberá concretar en cada caso la estrategia metodológica que considera más apropiada en función de su reflexión y análisis sobre los supuestos teóricos que justifican toda evaluación, pero teniendo en cuenta que entre modelo y diseño debe existir una clara coherencia, tal como hemos plasmado en otro trabajo anterior (De Miguel, 2000). Especialmente deberá valorar —a la hora de tomar una opción— las consecuencias prácticas, ya que la finalidad última en toda evaluación de programas es la toma de decisiones pertinentes que incidan sobre la mejora del programa. De ahí la necesidad de potenciar enfoques evaluativos integradores que tengan en cuenta tanto los efectos del programa como su significado para los implicados ya que ambos aspectos son esenciales a la hora de elaborar las propuestas de mejora. Ello implica, de nuestro punto de vista, que sea aconsejable abordar siempre la evaluación de un programa con un enfoque comprensivo que integre aproximaciones teóricas diversas y estrategias metodológicas complementarias. Lo cual significa optar por modelos y diseños menos radicales y más eclécticos.

8. CONSIDERACIONES FINALES

Una vez que hemos abordado «algunos» de los problemas que conciernen a la fundamentación teórica de la evaluación de programas y a las exigencias que conlleva la elección de un enfoque o aproximación metodológica, procede efectuar ciertas consideraciones finales que, a modo de resumen, contribuyan a clarificar el complejo marco en el que se mueve el desarrollo de esta disciplina y el trabajo de un evaluador. Como ya hemos dicho, la evaluación de programas es una tarea esencialmente práctica pero ello no significa que esté alejada de la teoría. Al contrario, necesitamos más y mejor teoría que ilumine lo qué hacemos, cómo lo hacemos y por qué. En definitiva, necesitamos justificar teóricamente una actividad que es esencialmente aplicada.

En este proceso de búsqueda hacia una fundamentación teórica, consideramos oportuno recapitular algunas de las lecciones que hemos aprendido desde la experiencia (Shadish, Cook y Leviton 1991, Chelimsky 1997, Cook 1977, Scriven 1997, Stake y Mabry 1997, Rossi, Freeman y Lipsey 1999, De Miguel 1999b) y presentar algunas de las connotaciones más relevantes que actualmente enmarcan la evaluación de toda

estrategia de intervención social. Señalamos, entre ellas, aquellas que nos parecen que pueden contribuir positivamente a consolidar la identidad científica de esta disciplina. A saber:

1. *Profundizar sobre las teorías relacionadas con el cambio social*

Abordar la evaluación de una estrategia de intervención social exige previamente precisar qué se entiende por cambio o mejora social y cómo debemos proceder para que esto sea posible. Ello requiere un conocimiento de las teorías sobre cómo se pueden resolver los problemas sociales y, en consecuencia, sobre las características que debe reunir toda estrategia de intervención que se oriente a promover determinadas mejoras. En esta línea algunos autores señalan la necesidad de que quienes planifican los programas formulen previamente una «teoría normativa» que defina y justifique cada una de las intervenciones y cómo se debe aplicar para que tenga las consecuencias deseadas (Wholey, Hatry y Newcomer, 1994). El evaluador que tiene un mayor conocimiento de las teorías sobre el cambio social estará en mejores condiciones de llevar a cabo procesos evaluativos más adecuados.

2. *Establecer una teoría evaluativa explícita además de implícita*

Toda evaluación requiere una reflexión teórica previa sobre el tipo de programa objeto de análisis con el fin de justificar los modelos y procedimientos que vamos a utilizar para llevar a cabo los procesos evaluativos. Generalmente esta reflexión no se lleva a cabo de forma explícita, aunque de una u otra forma todos los evaluadores tienen una teoría sobre la evaluación de forma implícita. Por ello, a efectos de que nuestra actividad sea algo más que una metodología aplicada, necesitamos reflexionar sobre los diversos supuestos que definen nuestro concepto sobre el programa y explicitar los parámetros teóricos y criterios a partir de los cuales vamos a elegir el enfoque evaluativo. Necesitamos pasar de una teoría tácita o implícita sobre el programa a una teoría articulada que oriente y justifique la aproximación metodológica elegida (Weis 1997, Rossi, Freeman y Lipsey 1999).

3. *Clarificar los fundamentos epistemológicos de toda evaluación*

Ahora bien, para llegar a la formulación de una teoría explícita es necesario que, antes de proceder a la planificación de una evaluación, el evaluador reflexione y precise el punto de vista que adopta en relación con los diferentes supuestos o fundamentos que delimitan y justifican la evaluación de un programa concreto. De modo especial deberá concretar el marco de referencia que va a utilizar en el proceso de construcción del conocimiento; es decir, delimitar qué pretende conocer, qué tipo de conocimientos necesita y cómo va a proceder para su construcción. Esta concreción lógicamente implica reflexionar antes sobre cuestiones básicas de carácter ontológico, epistemológico y metodológico que son claves para poder entender y explicar sus decisiones prácticas (Shadish, Cook y Leviton 1991, Rebien 1997). Suplir esta reflexión por la «afiliación» a un paradigma no contribuye al desarrollo del pensamiento crítico del evaluador.

4. *Abordar la evaluación como una tarea ecléctica*

Este recurso —habitualmente utilizado— de definir o explicitar los supuestos teóricos dentro de un paradigma no es recomendable en el campo de la evaluación porque,

como ya hemos avanzado, no existe un paradigma que ofrezca un marco teórico comprensivo capaz de integrar todos los interrogantes y problemas que conlleva la realización de un proceso evaluativo que por su propia naturaleza debe responder a distintas finalidades, diferentes audiencias y usos muy diversos. La evaluación es esencialmente una tarea ecléctica por lo que es muy cuestionable la utilización de los paradigmas clásicos con este fin (Cook, 1997). Este hecho nos advierte, al mismo tiempo, en la necesidad de profundizar sobre el estatuto científico de esta disciplina desde otras perspectivas más realistas (Pawson y Tilley, 1997).

5. Elegir los métodos en función de la finalidad

Dado que no existe un prototipo de paradigma para evaluar programas, procede insistir que los métodos deben ser elegidos en función de la finalidad que orienta el proceso y no al revés como frecuentemente sucede. La experiencia nos lleva a afirmar que no existen métodos de evaluación que se puedan aplicar siempre y en todas las situaciones sino que éstos deben ser elegidos en función del propósito hacia la que se orienta la evaluación y la necesidad de obtener resultados «creíbles». Así como los métodos cualitativos no son buenos para generar estimaciones cuantitativas sobre los resultados de un programa, tampoco los métodos cuantitativos son adecuados para obtener informaciones respecto a la teoría que lo fundamenta y los procesos desarrollados. No cabe, por tanto, apriorismos o planteamientos exclusivos ya que son necesarios ambos tipos de metodologías. Las dificultades se presentan en el plano de la logística, no en el de la lógica (Cook y Reichardt 1986, Cook 1997, Rossi, Freeman y Lipsey 1999).

6. Fomentar la diversidad metodológica

Cuando se asume que las evaluaciones deben responder a diversos propósitos y audiencias, lo habitual es utilizar una combinación parsimoniosa de datos, análisis y métodos en función del tipo de informaciones y conocimientos que perseguimos en cada momento y fase del proceso. La evaluación es un campo en el que se pueden aplicar todo tipo de estudios y metodologías y donde la tarea se enriquece precisamente con la pluralidad. Por ello, la única opción válida es abordar los procesos con un enfoque comprensivo y utilizar los procedimientos que en cada caso estimamos oportunos, al margen de polémicas estériles que no conducen a ninguna parte. La evaluación es un campo del conocimiento donde el debate entre metodologías está superado (Reichardt y Rallis 1994, Datta 1997, De Miguel 1999).

7. Promover trabajos de síntesis antes que individuales

Precisamente porque la mayoría de las evaluaciones exigen recabar informaciones diversas, utilizando fuentes muy distintas y métodos diferentes, parece necesario abordar esta tarea con una visión pluralista e integradora con el fin de que los juicios de valor se fundamenten sobre conocimientos contrastados por la opinión mayoritaria de todos los implicados en el programa. De ahí que la fiabilidad y validez de las observaciones y recomendaciones aumente en la medida que hayan sido elaboradas a partir de la integración de los resultados de varios estudios, lo que significa priorizar los trabajos de síntesis y metaanálisis frente a los individuales (Lipsey y Wilson 1993, Cook 1997). No debemos olvidar que la calidad de un conocimiento se incrementa en la medida que haya un escrutinio público mayor sobre el mismo.

8. *Velar por la credibilidad de los informes*

Fácilmente se comprende que la elaboración del informe sobre una evaluación no es una cuestión neutra en relación con las aplicaciones o usos de las recomendaciones sugeridas. Para que el informe o informes sean útiles deben reunir dos notas que nos parecen muy importantes. De una parte, deben responder a las necesidades de las audiencias y ser elaborados de forma objetiva que no ponga en cuestión la independencia del evaluador, lo que conlleva considerar la naturaleza de las relaciones entre el evaluador y los implicados. De otra, deben ser redactados en términos claros y con datos precisos que sean fácilmente comprensibles por las audiencias implicadas. Ambas cuestiones condicionan la credibilidad de los informes y, en consecuencia, su utilidad (Patton 1986, Wholey, Hatry y Newcomer, 1994).

9. *Analizar los factores contextuales y políticos que condicionan el uso de la información*

Además de la credibilidad que ofrecen los informes existen otros factores contextuales y políticos que condicionan la aplicación o utilización de los resultados y que también deben ser considerados por los evaluadores con el fin de eliminar las barreras que dificultan que los procesos tengan consecuencias prácticas. Entre estos factores debemos plantearnos todos los problemas que surgen entre la entrega de los informes y la puesta en práctica de las recomendaciones oportunas; es decir, la difusión y diseminación de los informes, la interpretación de resultados, el conocimiento de los implicados en la toma de decisiones, los argumentos que avalan cada una de las alternativas sugeridas, etc.. El análisis de estos problemas nos permitirá buscar soluciones orientadas a eliminar o reducir sus efectos e incrementar así la eficacia de nuestros trabajos (Braskamp y Brown 1980, Ewell 1989, Torres, Preskill y Piontek 1996).

10. *Realizar estudios de metaevaluación sobre los trabajos efectuados*

Finalmente todos debemos considerar que la mejor forma de avanzar y profundizar sobre el marco teórico de nuestra disciplina y nuestra tarea es a través de la propia evaluación del trabajo que realizamos. De ahí que debemos someter a control todos nuestros trabajos con el fin de que, al detectar sus puntos fuertes y débiles, podamos tener una visión realista sobre la calidad del mismo y las limitaciones que presenta para su utilización posterior. Para ello debemos habituarnos a efectuar estudios de metaevaluación utilizando guías elaboradas con este fin y remitir nuestros informes a evaluación externa para recabar otras opiniones independientes. Es un camino que necesariamente debemos recorrer si queremos mejorar nuestro trabajo, mejorar las bases teóricas de nuestra disciplina, y ganar más credibilidad ante nuestras audiencias (Joint Committee 1988, Shadish, Newman, Scheirer y Wye 1995).

Finalmente, quisiera añadir que estas consideraciones que sugerimos son algunas de las reflexiones a las que hemos llegado desde nuestra experiencia personal y que, en la medida que están en línea con el pensamiento de otros autores, esperamos puedan ser útiles para quienes están preocupados por estas cuestiones. Los evaluadores presionados por la práctica hemos dejado poco espacio para la reflexión teórica y quizás ha llegado el momento en que nos tengamos que plantear la necesidad de fundamentar nuestra tarea ya que en ello no sólo se haya comprometida la identidad de esta disciplina sino también la misma credibilidad de nuestro trabajo.

BIBLIOGRAFÍA

- Alkin, M. y Ellett, J. (1990). «Development of Evaluation Models». En Walberg, H. y Haertel, G. (eds.). *The international encyclopedia of educational evaluation*. Oxford, Pergamon Press.
- Anguera, T. (1992). «Diseños diacrónicos en programas de intervención». En *Rev. Bordón*, 43 (4), 421-429.
- Anguera, T. (1995). «Diseños». En Fernández Ballesteros, R. (ed.). *Evaluación de programas*. Madrid, Síntesis.
- Apodaca, P. (1999). «Evaluación de los resultados y del impacto». *Revista de Investigación Educativa*, 17(2), 363-377.
- Bickman, L. (ed.) (1987). *Using Program Theory in Evaluation. New Directions for Program Evaluation* n° 33. San Francisco, Jossey-Bass.
- Bickman, L. (ed.) (1990). *Advances in Program Theory. New Directions for Program Evaluation* n° 47. San Francisco, Jossey-Bass.
- Blaskar, R. (1975). *A realist theory of science*. Brighton, Hasvester.
- Boruch, R. (1997). *Randomized controlled experiments for planning and evaluation: A practical guide*. Thousand Oaks, CA: Sage.
- Braskamp, L. y Brown, R. (1980). *Utilization of Evaluation Information. New Directions for program Evaluation* n° 5. S. Francisco, Jossey-Bass.
- Bryk, A. (ed.) (1983). *Stakeholder-based Evaluation. New Directions for program Evaluation* n° 17, S. Francisco, Jossey-Bass.
- Chelemsky, E. (1997). «The Coming transformations in Evaluation». En Chelemsky, E. y Sadish, W. (eds.). *Evaluation for the 21 Century*. London, CA: Sage.
- Chen, H.T. (1990). *Theory-driven evaluations*. Newbury, CA: Sage.
- Cook, T. (1997). «Lessons learned in Evaluation Over the past 25 Years». En Chelemsky, E. y Sadish, W. (eds.). *Evaluation for the 21 Century*. London, CA: Sage.
- Cook, T. y Reichardt, CH. (1986). *Métodos cualitativos y cuantitativos en la investigación evaluativa*. Madrid, Morata.
- Cronbach, L. (1982). *Designing evaluations of educational and social programs*. S, Francisco, CA: Jossey-Bass.
- Datta, L. (1997). «Multimethod Evaluations: Using case studies together with other methods». En Chelemsky, E. y Sadish, W. (eds.). *Evaluation for the 21 Century*. London, CA: Sage.
- De Miguel, M. (1989). «Modelos de evaluación sobre organizaciones educativas». En *Revista de Investigación Educativa*, 13, 21-56.
- De Miguel, M. (1990). «La investigación en la acción. ¿Una alternativa metodológica para las Ciencias Sociales? En Vázquez, J. y Otros. *Metodología de la Investigación científica*. Santiago de Compostela, ICE.
- De Miguel, M. (1991). «Utilización de indicadores en la evaluación de la docencia». En De Miguel, M.; Mora, J. y Rodríguez, S. (eds.). *La evaluación de las Instituciones Universitarias*. Madrid, Consejo de Universidades.
- De Miguel, M. (1995). «Participación ciudadana en la evaluación de las políticas públicas». *Aula Abierta*, 66, 211-218.

- De Miguel, M. (1999). «La evaluación de programas: entre el conocimiento y el compromiso». *Rev. de Investigación Educativa*, 17 (2), 345-48.
- De Miguel, M. (1999b). «Problemas teóricos y metodológicos en la evaluación externa de programas multiculturales». En Martín, T. y Serrate, L. (Coord.). (1999). *Evaluación y Ámbitos Emergentes en Animación Sociocultural*. Madrid, Sanz y Torres.
- De Miguel, M. (2000). «Modelos y diseños en la evaluación de programas». En Sobrado, L. (ed.). *Orientación Profesional: Diagnóstico e Inserción Sociolaboral*. Barcelona, Editorial Estel.
- De Miguel, M. (2000b). «La evaluación externa de un programa de Educación Social» En Pérez Serrano, G. (Coord.). *Modelos de Investigación Cualitativa*. Madrid, Narcea.
- Dennis, M. y Boruch, R. (1989). «Randomized experiments for planning and testing projects in developing countries». *Evaluation Review*, 13 (3), 292-309.
- Drummond, M., Stoddart, G. y Torrance, G. (1991). *Métodos de evaluación económica de los programas de atención a la salud*. Madrid, Diaz de Santos.
- Ewell, T. (ed.). (1989). *Enhancing Information Use in Decision Making. New direcciones for institutional research*, nº 64. S. Francisco, Jossey-Bass.
- Fetterman, D., y Pitman, M. (1986). *Educational Evaluation*. London, Sage.
- Fetterman, D. (1994). «Empowerment evaluation». *Evaluation Practice*, 15(1), 1-15.
- Fetterman, D., Kaftarian, S. y Wasdersman, A. (eds.)(1966). *Empowerment evaluation*. London, CA: Sage.
- Fitz-Gibbon, C. y Morris, L. (1987). *How to design a program evaluation*. Newbury Park, CA: Sage.
- Folc'hlay, B. y Plottu, E. (1998). «Democratic evaluation». *Evaluation*, 4 (3), 261-277.
- Guba, E. y Lincoln, Y. (1981). *Effective evaluation*. S. Francisco, Jossey-Bass.
- Guba, E. y Lincoln, Y. (1989). *Fourth generation evaluation*. Newbury Park, CA.: Sage.
- Herman, J., Morris, L. y Fitz-Gibbon, C. (1987). *Evaluator's handbook*. Newbury Park, CA: Sage.
- Hesse, M. (1974). *The structure of scientific inference*. London, MacMillan.
- House, E. (1980). *Evaluating with validity*. Newbury Park, CA.: Sage.
- Howe, K. (1992). «Gedtting over the quantitative-qualitative debate». *American Journal of Education*, 100(2), 236-255.
- Joint Committee On Standars, (1988). *Normas de Evaluación*. México, Trillas.
- Keeves, J. (ed.). (1988). *Educational Research, Methodology and Measurement: A International Handbook*. Oxford, Pergamon Press.
- Lewin, H.M. (1983). *Cost-effectiveness* Newbury Park, CA: Sage.
- Lincoln, Y. y Guba, E. (1985). *Naturalistic inquiry*. Newbury Park, CA.: Sage.
- Lipsey, M. y Wilson, D. (1996). *Toolkit for Practical Meta-Analysis*. Cambridge, MA: Human Services Research Institute.
- MacDonald, B. (1977). «A Political classification of evaluation studies». En Hamilton, D. y al. (eds.). *Beyond the Numbers Game*. London, MacMillan.
- Madaus, G., Scriven, M. y Stufflebeam, D. (eds.) (1983). *Evaluations models: Viewpoints on educational and human services evaluation*. Boston, Mass.: Kluwer-Nijhoff.
- Martin, L. (1993). *Total quality management in human service organizations*. Newbury Park, CA: Sage.

- Martin, L. y Kettner, P. (1996). *Measuring the performance of human service program*. London, Sage/SHSG.
- McLaughlin, G. y McLaughlin, J. (1989). «Barriers to Information Use: The Organizational Context». En Ewell, T. (ed.). *Enhancing Information Use in Decision Making. New directions for institutional research*, nº 64. S. Francisco, Jossey-Bass.
- McLaughlin, M y Phillips, C. (eds.) (1991). *Evaluation and Education: At quarter Century*. Chicago, NSSE.
- Nas, T. (1996). *Cost-benefit analysis. Theory and application*. Newbury Park, CA: Sage.
- Nevo, D. (1990). «Role of the evaluator». En Walberg, H. y Haertel, G. (eds.): *The international encyclopedia of educational evaluation*. Oxford, Pergamon Press.
- Patton, M. (1986). *Utilization-focused evaluation*. (1ª ed., 1978). Newbury Park, CA: Sage.
- Patton, M. (1987). *How to use qualitative methods in evaluation*. Newbury Park, CA.: Sage.
- Patton, M. (1994). «Developmental Evaluation». *Evaluation Practice*, 15(3), 311-319.
- Pawson, R. y Tilley, N. (1997). *Realistic Evaluation*. Newbury Park, CA.: Sage.
- Pineault, R. y Daveluy, C. (1989). *La planificación sanitaria: conceptos, métodos y estrategias*. Barcelona, Masson.
- Rebien, C. (1997). «Development assistance evaluation and the foundations of program evaluation». *Evaluation Review*, 21 (4), 438-460.
- Reichart, C. y Rallis, S. (eds.) (1994). *The qualitative-quantitative debate: new perspectives. New directions for Evaluation*, nº 61. S. Francisco, Jossey Bass.
- Rossi, P. y Freeman, H. (1989). *Evaluación. Un enfoque sistémico para programas sociales*. México, Trillas.
- Rossi, P., Freeman, H. y Lipsey, M. (1999). *Evaluation. A systematic approach*. Thousand Oaks, Sage.
- Ruttman, L. (1984). *Evaluation research methods: a base guide*. Beverly Hills, Sage.
- Sayer, A. (1984). *Method in social science*. London, Hutchinson.
- Schalock, R. (1995). *Outcomes-based evaluation*. Newbury Park, CA.: Sage.
- Scriven, M. (1991). *Evaluation Thesaurus*. Newbury Park, CA.: Sage.
- Scriven, M. (1997). «Truth and Objectivity in Evaluation». En Chelimsky, E. y Sadish, W. (eds.). *Evaluation for the 21 Century*. London, CA: Sage
- Shadish, W., Cook, T. y Leviton, L. (1991). *Foundations of program evaluation*. Newbury Park, CA.: Sage.
- Shadish, W., Newman, D., Scheirer, M. y Wye, C. (eds). (1995). *Guiding Principles for Evaluators. New Directions for Evaluation*, nº 66. S. Francisco, Jossey-Bass.
- Sheerens, J. (1987). «Beyond decision-oriented evaluation». *International Journal of Educational Research*, 11 (1) 83-93.
- Smith, M. (1998). «Empowerment evaluation: Theoretical and methodological considerations» *Evaluation and Programming Planning*, 21, 255-261.
- Sonnichsen, R. (1994). «Evaluators as Change Agents». En Wholey, J., Hatry, H. y Newcomer, K. (1994). *Handbook of Practical Program Evaluation*. S. Francisco, Jossey-Bass Publishers.
- Stake, R. y Mabry, L. (1997). *Advances in program evaluation*. Newbury Park, CA.: Sage.
- Stecher, B. y Davis, W. (1991). *How to focus an evaluation*. Newbury Park, CA.: Sage.

- Stufflebeam, D. y al. (1971). *Educational Evaluation and decision-making in Education*. Illinois, Peacock-Itasca.
- Stufflebeam, D. y Shinkfield, A. (1987). *Evaluación sistemática*. Barcelona, Paidós/MEC.
- Stufflebeam, D. (1994). «Empowerment evaluation, Objectivist evaluation and Evaluation Standars». *Evaluation Practice*, 15, 321-338.
- Torres, R., Preskil, H. y Piontek, M. (1996). *Evaluation Strategies for Communicating and Reporting*. Thousand Oaks, CA: Sage.
- Vedung, E. (1997). *Evaluación de políticas públicas y programas*. Madrid, Instituto Nacional de Servicios Sociales.
- Walker, J. y Evers, C. (1988). «The epistemological unity of educational research». En Keeves, J. (ed.). *Educational Research, Methodology and Measurement: A International Handbook*. Oxford, Pergamon Press.
- Weis, C. (1975). *Investigación evaluativa*. México, Trillas.
- Weis, C. (1997). «How can Theory-based Evaluation make Greater Headway?». *Evaluation Review*, 21(4), 501-524.
- Wholey, J., Hatry, H. y Newcomer, K. (1994). *Handbook of Practical Program Evaluation*. S. Francisco, Jossey-Bass Publishers.
- Worthen, B. y Sanders, J. (1987). *Educational evaluation*. London, Longman.
- Worthen, B., Sanders, J. y Fitzpatrick, J. (1997). *Program evaluation alternative approaches and practical guidelines*. Newbury Park, CA.: Sage.