
PONENCIA II

PERSPECTIVAS METODOLÓGICAS ACTUALES DE LA EVALUACIÓN DE PROGRAMAS EN EL ÁMBITO EDUCATIVO

por

F. Javier Tejedor, Ana García-Valcárcel y M^a José Rodríguez Conde
Universidad de Salamanca

1. CONCEPTUALIZACIÓN DE LA EVALUACIÓN DE PROGRAMAS

La «evaluación de programas» surge en el ámbito educativo en el entorno de los años setenta con el objetivo fundamental de valorar la eficacia de los programas de acción social que se llevaban a cabo, ya que no había constancia alguna de la consecución de los objetivos previstos. La evaluación de programas de actuación pública o de intereses sociales supone, sobre todo a medida que el programa es ambicioso, analizar realidades complejas, dinámicas y en las que se desarrollan procesos de cambio de sistemas interdependientes (Colás, 1993, p. 10).

La evaluación de programas, dicen Cook y Shadish, es un modelo racional de solución de problemas de carácter psicológico, social, educativo,..., que requiere que: (1) los problemas se definan primero con claridad y después... (2) se genere una amplia variedad de soluciones potenciales... (3) algunas de las cuales se implementan y... (4) finalmente se evalúan. (5) El conocimiento de las soluciones positivas se presentan a los responsables políticos quienes... (6) lo utilizarán para tomar las decisiones políticas que correspondan (Cook y Shadish, 1986, p. 197).

La evaluación de un programa se planifica para tomar decisiones, por tanto los datos que se busquen deben facilitar dicha toma de decisiones. Debe ser un proceso cíclico y continuo, cuya puesta en marcha responda a una programación sistemática, para valorar si un programa consiguió o no alcanzar los objetivos para cuya consecución fue ideado. Este proceso consta de cuatro actividades básicas (Ato y otros, 1989):

- identificar los objetivos del programa
- transformar los objetivos en variables observables (indicadores)

- recoger datos empíricos relativos a las variables identificadas
- comprobar los objetivos logrados con los datos empíricos obtenidos, valorando el éxito o fracaso en su consecución y el balance económico del binomio coste/eficacia.

La finalidad de la evaluación es, pues, emitir un juicio valorativo sobre el programa y determinar en qué medida logra los objetivos marcados. Trata de maximizar la eficiencia y la eficacia de las acciones dirigidas a modificar segmentos de la realidad.

Hay que destacar también que la evaluación de programas pretende explicar los hallazgos encontrados (relación entre variables, incidencia de unas variables en otras,...) y no únicamente describir los resultados del proceso de medición.

La evaluación de programas ha ido perfilándose con una metodología concreta en los últimos quince años. Esta metodología queda delimitada por: una terminología propia, un conjunto de herramientas conceptuales y analíticas específicas, y unos procesos y fases también específicos. El conjunto define un campo floreciente en el que realmente lo específico es la mezcla y no las partes constitutivas.

Para Alvira (1991), algunas circunstancias que han contribuido a conformar este campo de actuación serían:

- a) Mayor interrelación entre las actividades de evaluación y programación. Si en una perspectiva clásica la evaluación se realiza después de las actividades de planificación, programación y desarrollo del programa, en la actualidad ambas actividades caminan de forma paralela. Es decir, frente al esquema de actuación tradicional se plantea una estrategia con mayor interrelación entre las partes, sin que ello suponga una pérdida de especificidad de cada uno de los componentes (cuadro 1):
- b) La evaluación de programas, desde la perspectiva metodológica que es la que aquí tratamos, es investigación evaluativa; es decir, es investigación (con todos los considerandos epistemológicos, metodológicos y técnicos que ello entraña), pero es más que investigación, ya que se entiende como un proceso cuya finalidad es aplicar procedimientos científicos para acumular evidencia válida y fiable sobre la manera y grado en que un conjunto de actividades producen resultados o efectos concretos. Intenta determinar el valor de la acción educativa, buscar soluciones a problemas concretos o mejorar las estrategias de actuación.
- c) En la evaluación de programas se produce una mayor implicación de las partes interesadas, lo que no necesariamente implica mayor aproximación a estrategias cualitativas (observación participante, investigación acción,...); lo que sí implica necesariamente es un mayor esfuerzo por atender las demandas de quienes patrocinan el trabajo y de los potenciales beneficiarios del mismo.
- d) En la evaluación de programas pueden integrarse estrategias de investigación muy diversas, tanto cuantitativas como cualitativas. Es un campo de trabajo que posibilita la utilización simultánea y sucesiva de diferentes metodologías para el tratamiento de los temas objeto de estudio e intervención. Será la

CUADRO 1
PERSPECTIVA ACTUAL EN EVALUACIÓN DE PROGRAMAS

Estrategia tradicional	Estrategia actual
1) Identificación de problemas y/o necesidades	Evaluación de necesidades
2) Programación y priorización de objetivos	Evaluación conceptual
3) Diseño del programa de intervención elegido	Evaluación del modelo
4) Ejecución del programa	Evaluac. de implementación Evaluación de la cobertura Evaluación del proceso
5) Evaluación	Evaluación de resultados

naturaleza específica del programa y la pericia del evaluador quienes a la postre determinarán la utilización de una o varias opciones metodológicas.

La evaluación de programas en el ámbito español, como señalan Colás y Rebollo (1993, p. 9), se caracteriza por las escasas realizaciones sistematizadas y el predominio de determinados enfoques.

A nivel internacional, la situación actual se define por la utilización de múltiples modelos que se aplican a distintos contenidos (necesidades, diagnóstico, productos, impactos,...), la multidisciplinariedad (análisis etnográfico y estadístico), multidimensionalidad en objetivos, en perspectivas y en niveles de valoración (validez, credibilidad, costes, rentabilidad, aplicabilidad). En síntesis, múltiples metodologías, múltiples impactos, múltiples formatos de informe,...Es una multiplicidad de lo múltiple (Colás y Rebollo, 1993, citando a Conner).

Una consecuencia de este panorama es sin duda la dificultad de conceptualizar la «evaluación de programas», encontrándonos con múltiples definiciones en las que cada autor destaca los elementos que más le interesan. En la búsqueda de patrones de similitud nos encontramos dos tipos básicos de apreciaciones: los que hacen hincapié en la dimensión técnica y los que ponen el énfasis en los considerandos ideológicos, aunque ni siquiera estas caracterizaciones extremas pueden considerarse tan distintas ya que realmente no son opuestas ni contradictorias, resultando en el fondo complementarias por estar destacando cada una de ellas momentos distintos del proceso: selección del contenido del programa, toma de decisiones, implicación

de los participantes, en un caso, y recogida y análisis de información orientada a la toma de decisiones, en otro.

Los considerandos básicos que caracterizan la evaluación de programas, en cualquiera de sus modalidades, serían, en opinión de Colás (1993) y Gairín (1991) que nosotros reinterpretemos, los siguientes:

- La evaluación de programas no debe entenderse como un proceso uniforme con un modelo único de evaluación o con una única metodología.
- La evaluación de programas se ajustará inicialmente a la idea que se tenga de evaluación. La concepción más amplia es aquella que entiende que evaluación es el conocimiento empírico del mérito o valor de las actividades y objetivos de un programa.
- Irá referida a un programa o conjunto sistemático de actuaciones que se ponen en marcha para alcanzar unos objetivos.
- Junto a la función educativa, el programa puede estar cumpliendo otras funciones de carácter ideológico y político.
- Planificación y evaluación deben pensarse simultáneamente.
- La evaluación de necesidades es una tarea previa a la planificación de un programa.
- Los modelos de evaluación deben ser al mismo tiempo modelo de análisis y de toma de decisiones.
- La planificación-evaluación del programa debe ser detallada si queremos facilitar las tareas posteriores de control.
- El evaluador debe estar familiarizado con el objeto de la evaluación (contacto y negociación previa).
- La participación de los interesados es deseable ya que aporta validez y es garantía de utilidad y aplicabilidad. Aunque en algunos diseños no se incluye de forma explícita o se duda de su utilidad, creemos que el problema actualmente no se plantea en términos de si admitir o no la participación de los interesados (gestores y destinatarios) en el proceso sino el delimitar la forma y el grado en que debe producirse (en algunos modelos, la participación es una fase del proceso).
- Los criterios e indicadores son básicos en la evaluación así como el establecimiento de estándares o grados aceptables de éxito.
- La evaluación se entiende como proceso, confiriéndole un carácter dinámico. El concepto de evaluación concreto que se adopte marcará la elección de posibles modelos teóricos y enfoques metodológicos a seguir, así como otros aspectos a considerar (tema, actividades, rol del evaluador, audiencia, informes,...).
- Deben aplicarse procedimientos científicos, que incluyan estrategias de diseño, recogida y análisis de información rigurosos y sistematizados, lo que no excluye en modo alguno la diversidad de enfoques metodológicos (técnicas de recogida y análisis de datos y estrategias de toma de decisiones). Tampoco el carácter científico del proceso evaluador debe hacernos concebir dicho

proceso como una actividad exclusivamente teórica, mera aplicación de los procedimientos científicos.

- Actualmente el énfasis de la tarea evaluativa se pone en la emisión del juicio de valor, a partir de una adecuada recogida y análisis de la información pertinente.
- Deben superarse las opciones excluyentes de metodologías específicas; la evaluación de programas permite utilizar de forma conjunta y coordinada metodologías diferentes.
- No todos los diseños son adecuados. Deben responder a criterios de validez y utilidad.
- Debe tener como proyección la toma de decisiones, lo que implica otorgarle una utilidad a la evaluación.
- Se remarca el carácter contextualizado, en el sentido de que ha de estar condicionada por las circunstancias y ha de ajustarse a las condiciones reales en las que se aplica el programa objeto de evaluación.
- La tipología de la evaluación depende del momento o fase del programa que quiera analizarse: de necesidades, de proceso, de producto,...
- Hay que prestar especial atención a los efectos no previstos, ya que pueden deberse tanto a la potencialidad del programa como a su incorrecta aplicación.
- El éxito de un programa no se mide únicamente por la consecución de objetivos ya que hay programas que no los formulan o que se plantean desde otras perspectivas menos funcionales.

2. BREVE PERSPECTIVA HISTÓRICA

Para algunos autores el comienzo de la Evaluación de Programas, tal y como hoy se entiende, se sitúa en la década de los años cincuenta en el entorno de las actividades de evaluación de objetivos educativos llevadas a cabo por Tyler.

En los planteamientos de Tyler la evaluación debe cumplir el papel de verificar la consecución de los objetivos operativos propuestos previamente a la realización de la actividad (bajo el supuesto de que esta realización posibilita la consecución de dichos objetivos). Cuanto más claramente estén explicitados los objetivos más fácil resultará la verificación o no de su consecución. La diferencia entre las medidas pre y post programa será un indicador preciso de sus efectos. Los esfuerzos metodológicos por garantizar la exactitud de esas medidas (y por tanto del juicio emitido) vinieron dados por Campbell y Stanley con la publicación en 1973 de su conocida obra sobre los diseños de investigación. Las pautas metodológicas contenidas en esa obra orientan todavía hoy la actividad de investigación evaluativa, aunque en el entorno de los años ochenta, con el cuestionamiento de la viabilidad de los estudios experimentales en el ámbito de las ciencias sociales, se abren nuevas perspectivas que suponen la incorporación de nuevas estrategias metodológicas, tanto cuantitativas (diseños correlacionales, series, técnicas de clasificación,...), como cualitativas.

Presentamos, de forma esquemática, un breve panorama de la evolución histórica de las prácticas evaluativas.

1) Antecedentes: En torno a los años 20 (evaluación de 1ª generación); corriente positivista y empiricista; se centra en la medición de características individuales (testing).

2) Consolidación: Años 30-50 (evaluación de 2ª generación); referencia a los objetivos (Tyler); se pretende conseguir la descripción de la medida.

3) Auge: Años 55-70 (evaluación de 3ª generación); evolución de la etapa de la inocencia al realismo; se pretende medir, describir e interpretar; predomina la evaluación como investigación empírica positiva, que utiliza procedimientos estandarizados para la recogida y análisis de datos orientados a la toma de decisiones.

4) Actualidad: A partir de los años 70; investigación evaluativa (diversidad de paradigmas); coincidiendo con la aparición de las prácticas cualitativas, la evaluación de programas incorpora modelos que se centran en los participantes; los procedimientos de investigación incorporados provienen del campo de la investigación cualitativa (naturalista, etnográfica,...); a partir de 1980 (evaluación de 4ª generación) se habla del modelo crítico de evaluación, sin una personalidad muy reconocida ya que la mayoría de los autores lo consideran una modalidad de la perspectiva cualitativa; actualmente, se presenta la evaluación como negociación (Guba y Lincoln).

Los ochenta, en opinión de Alvira (1991), se caracterizan por:

- Una expansión de la evaluación de programas a los países occidentales a la par que se restringe el dinero para éstas en Estados Unidos.
- La aparición de una pluralidad de perspectivas, exageradamente denominadas modelos, y de nuevos métodos.
- Un énfasis creciente en lograr que las evaluaciones sean útiles.
- Ampliación del tipo de preguntas a las que debe dar respuestas una evaluación, rebasando con creces la evaluación de resultados que había sido predominante anteriormente.

Hoy se vive en la metodología de la evaluación una situación plural, compleja y que tiende al realismo en mayor grado que anteriormente; las nuevas perspectivas no sustituyen a las anteriores sino que las complementan. Los cambios habidos en estos últimos años en la conceptualización de la evaluación de programas quedan recogidos en el cambio recogido en las dos posturas siguientes: la postura más tradicional que entendía la evaluación como el empleo de métodos científicos para medir el desarrollo y los resultados de programas que luego se utilizarán para la toma de decisiones, y la postura más actual que entiende la evaluación como el conocimiento empírico del mérito o valor de las actividades y objetivos de un programa.

3. LA INVESTIGACIÓN EVALUATIVA Y LOS PROCESOS DE INTERVENCIÓN

La investigación evaluativa es una manera de aumentar la racionalidad de las decisiones en relación con la puesta en marcha de programas sociales, evitando una excesiva dependencia de criterios políticos. Los objetivos de la investigación evaluativa suelen ser de dos tipos: elaborar juicios sobre la eficacia de los programas y ayudar en las decisiones a los responsables.

La investigación evaluativa es un modo de investigación que implica un proceso riguroso, controlado y sistemático de recogida y análisis de información fiable y válida para tomar decisiones sobre un programa educativo. Vemos que aparecen importantes puntos de contacto respecto a lo que entendemos por «investigación»: proceso sistemático, riguroso, controlado; utilización de técnicas de medición y análisis de datos (la investigación se distingue de la evaluación por los objetivos y no por los métodos).

Sin embargo, creemos que conviene diferenciar ambas tareas. Las características peculiares de la investigación evaluativa que señala De la Orden (1985 b) son:

- a) Los juicios de valor acompañan al desarrollo de todo el programa: desde la selección del problema a estudio hasta en la aplicación de metodologías concretas.
- b) Resulta difícil (a veces inadecuado) formular hipótesis precisas.
- c) La replicación es prácticamente imposible dadas las características casi irrepetibles de la ejecución de un programa de intervención.
- d) La recogida de datos está condicionada por la viabilidad del proceso.
- e) Multitud de variables relevantes sólo son susceptibles de un control superficial limitándose por tanto la utilización de diseños de carácter más experimental.
- f) Las decisiones importantes en relación con la paralización, sustitución o repetición del programa no son competencia del evaluador.
- g) El informe debe adaptarse a las exigencias de quien lo financia.

En el cuadro 2 se recogen, para algunos de los tópicos principales, las diferencias comentadas (y algunas otras) entre investigación (en términos generales) e investigación evaluativa.

Una característica importante de la investigación evaluativa es que se lleva a cabo en un proceso de intervención, en la misma acción, que es lo realmente importante. La investigación, señalaba Weis (1982), es una prioridad secundaria, lo que debe llevarnos a pensar que la evaluación/investigación tiene que adaptarse al contexto del programa y perturbar lo menos posible su desarrollo.

Esta aproximación de la investigación a la práctica educativa, ¿la hace ser más útil? Puesto que la Evaluación de Programas está siendo planteada, desde una perspectiva metodológica, en términos de investigación evaluativa, al plantearnos su utilidad nos estamos realmente preguntando por la utilidad de los resultados de la investigación. La respuesta, en todo caso, es compleja. En numerosas ocasiones se

CUADRO 2
INVESTIGACIONES VS INVESTIGACIÓN EVALUATIVA

<u>Tópicos</u>	<u>Investigación</u>	<u>Investigación evaluativa</u>
Propósito	incrementar el conocimiento sacar conclusiones	modificar la realidad tomar decisiones
generalización de resultados	a otros ámbitos (validez externa)	limitado a grupo de referencia
juicios de valor	conocimiento objetivo (la verdad)	asignar criterios de valor
papel investigador	autónomo	dependiente del gestor
temas	interés personal y/o científico	interés del gestor
metodología	básicamente cuantitativa	diversas
diseño	preestablecido	preestablecido pero flexible
hipótesis	explícitas	no explícitas
replicación	exigible	no exigible
datos	a partir del problema	a partir de la viabilidad
control y manipulación	deseable (exigible en la básica)	deseable, pero muy difícil
aleatorización	básica	no procede
criterios de validez	interna, externa y de constructo	utilidad credibilidad
interpretación de datos	de acuerdo a reglas definidas	valoración por los responsables del programa
informe	adaptado a comunidad científica	adaptado a usuario y gestores

han puesto de manifiesto las limitaciones que concurren en la investigación educativa para proyectar los resultados a la práctica cotidiana.

El proceso de evaluación de programas debe concebirse como una auténtica estrategia de investigación sobre los procesos educativos, en cuyos resultados deberían basarse las pautas sugeridas para orientar los procesos de intervención. Investigación e intervención se apoyan mutuamente. La intervención sacará partido de la investigación en la medida en que ésta se extienda a la solución de problemas prácticos. La investigación activa se asimila, enriqueciéndola, a la noción de investigación aplicada.

El método, como forma de la actividad humana, puede adoptar dos modalidades básicas, según que dicha actividad sea de carácter humano interno (métodos de pensamiento) o de carácter externo (métodos de acción). Los métodos de acción pueden a su vez concebirse como: a) técnicas o métodos de investigación, orientados al conocimiento de la realidad observada y b) técnicas o métodos de actuación dirigidas a transformar la realidad.

Toda situación educativa es susceptible de este doble tratamiento. Por una parte, existe la necesidad de «conocerla» y de «explicarla». Por otra, es preciso «comprenderla» y «mejorarla». Se requiere que la investigación alcance un nivel de integración suficiente, en este caso, entre el pensamiento y la decisión.

Entre el terreno del pensamiento teórico, donde sólo interesa el saber y en el que sólo el saber tiene valor propio, y el de la acción informada, donde lo útil y lo realizable van en primer lugar, hay una vasta zona en las ciencias del hombre donde las preocupaciones teóricas y utilitarias se mezclan en distinto grado: la preocupación por el saber se une de alguna manera a la preocupación por el actuar. Este campo de estudio, que algunos denominan «investigación orientada», está a la base de toda investigación evaluativa.

La investigación orientada/evaluativa nace como respuesta a necesidades sociales y se elabora en función de tales necesidades; se realiza con objeto de llegar a la solución de los problemas que exige una acción informada científicamente. Habrá de tener necesariamente un carácter interdisciplinar y viene determinada por los objetivos, los resultados y las motivaciones.

En estos momentos, lo que más nos podría interesar sería analizar los cauces que sigue la investigación orientada para contribuir a la utilización práctica de los resultados obtenidos, es decir, la contribución a la toma de decisiones, y a la inspiración de la acción educativa.

Pienso que la evaluación de programas, como técnica de investigación sobre los procesos de intervención socioeducativos, es una estrategia que contempla en su realización un potencial riquísimo para contribuir a la mejora de la calidad de la acción educativa pues tiene todos los requisitos exigibles a la investigación educativa útil.

No obstante, habría que preguntarse si siguen hoy siendo válidas las pesimistas reflexiones que Weis (1982) exponía hace unos años al valorar como no demasiado fructífera la breve historia de la investigación evaluativa; es más, decía, algunos

autores la consideran decepcionante. Pocos ejemplos de aportaciones importantes a las políticas y/o los programas pueden citarse. Ello se debe, en parte, a las notables resistencias de las organizaciones a ser evaluadas, a que sus programas sean evaluados, a ofrecer la información que se precisa en la forma en que se necesita. Otra parte de la culpa estriba en la manera en que a veces está estructurada, dotada de personal y ejecutada la misma evaluación. En ocasiones hay incoherencias entre los fines propuestos en la evaluación y el tipo de estudios o programas evaluados. Gran parte de la decepción que venimos comentando pueden deberse al carácter poco realista de las expectativas que se plantean, a la necesidad de continuidad en los estudios, a los esfuerzos por presentar adecuadamente el programa a los usuarios,...

4. MODELOS DE EVALUACIÓN. TIPOS Y PAUTAS PARA SU SELECCIÓN

En estos últimos años ha proliferado, quizá en exceso, la formalización de modelos de evaluación, lo que ha supuesto un cierto confusiónismo para el usuario. Algunos autores opinan (Alvira, 1991) que en la mayoría de los casos no se trata de modelos en sentido estricto sino de diferentes aproximaciones a unas pocas variantes básicas de la metodología de la evaluación.

Recordando que la evaluación de programas requiere la emisión de juicios de valor (y por tanto el establecimiento previo de criterios para la realización de esos juicios) podría entenderse que las diferentes perspectivas difieren simplemente en los criterios de valor utilizados, que vendrán determinados por los propios objetivos de la evaluación.

Algunos de estos modelos básicos son los conocidos modelos de Tyler, Stake, Schuman, Stufflebeam, Scriven, ... Dado el carácter de este escrito no vamos a entrar en su descripción; en la bibliografía se ofrecen referencias para que el lector interesado pueda profundizar en el estudio de cada uno de ellos.

Los elementos básicos que componen un modelo de evaluación son, en opinión de Colás (1992), los siguientes: finalidad científica y/o política, contenido, toma de decisiones, ámbito o unidad de evaluación, rol del evaluador, enfoque y proceso metodológico. Como vemos, en la evaluación de programas el modelo es una síntesis entre ideología y metodología: se fundamenta ideológicamente y se concreta después es una propuesta metodológica específica, aceptándose actualmente cinco modalidades principales: experimentalista, ecléctica, descriptiva, coste-beneficio y crítica (Talmage, 1982). Y siempre aceptando mejor a quienes defienden la complementariedad entre los distintos métodos que a quienes la rechazan.

Creo que la opinión mayoritaria actual sobre esta temática es clara: actitud de búsqueda de la complementariedad de los diferentes enfoques, evitando la oposición, y sobre todo la descalificación, a posturas diferentes a las personales. Esta posición ecléctica, vinculada a mayores dosis de pragmatismo, entiende que la relación entre paradigmas y métodos es de tipo instrumental, siendo la metodología una simple colección de técnicas y procedimientos (Cook y Reichardt, 1986). Cada día parece más evidente que un determinado problema, el análisis de una determi-

nada realidad, puede ser realizado desde distintas perspectivas; lo realmente importante será la coherencia entre los objetivos del estudio, el tipo de datos y los análisis que se lleven a cabo. Con estudios bien hechos —sean del tipo que sean— resultará más fácil detectar similitudes interpretativas a través de los procedimientos adecuados (triangulación, matriz multirrasgo-multimétodo, confrontación de resultados,...).

Se han utilizado distintos criterios parciales (contenidos y efectos de la evaluación, enfoque metodológico,...) para presentar diferentes clasificaciones de los modelos de evaluación. En algún caso, se mezclan en esos criterios aspectos teóricos y metodológicos lo que proporciona esquemas más coherentes. Seleccionamos para presentar aquí las propuestas de clasificación de Popham y Rebollo.

Popham (1980) nos sugiere la siguiente clasificación de los modelos:

a) *Modelos de concreción de metas:*

— *Modelo de Tyler*, caracterizado por:

- establecer las metas y objetivos
- ordenar los objetivos en amplias clasificaciones
- definir los objetivos en términos operativos
- establecer en qué condiciones puede demostrarse la consecución de objetivos
- explicar los propósitos de la evaluación a los destinatarios
- recopilar datos
- comparar los datos con los objetivos del programa

b) *Modelos de enjuiciamiento que acentúan criterios intrínsecos:*

- el evaluador tiene un papel muy importante
- es un planteamiento técnico-profesional
- carece de conexión con el contexto

c) *Modelos de enjuiciamiento que acentúan criterios extrínsecos:*

— *Modelo de Scriven:*

- distinción evaluación formativa-sumativa
- se demanda especialmente el interés de los objetivos

— *Modelos de Stake* (evaluación respondiente)

d) *Modelos de facilitación de decisiones:*

— *Modelo de Cronbach*

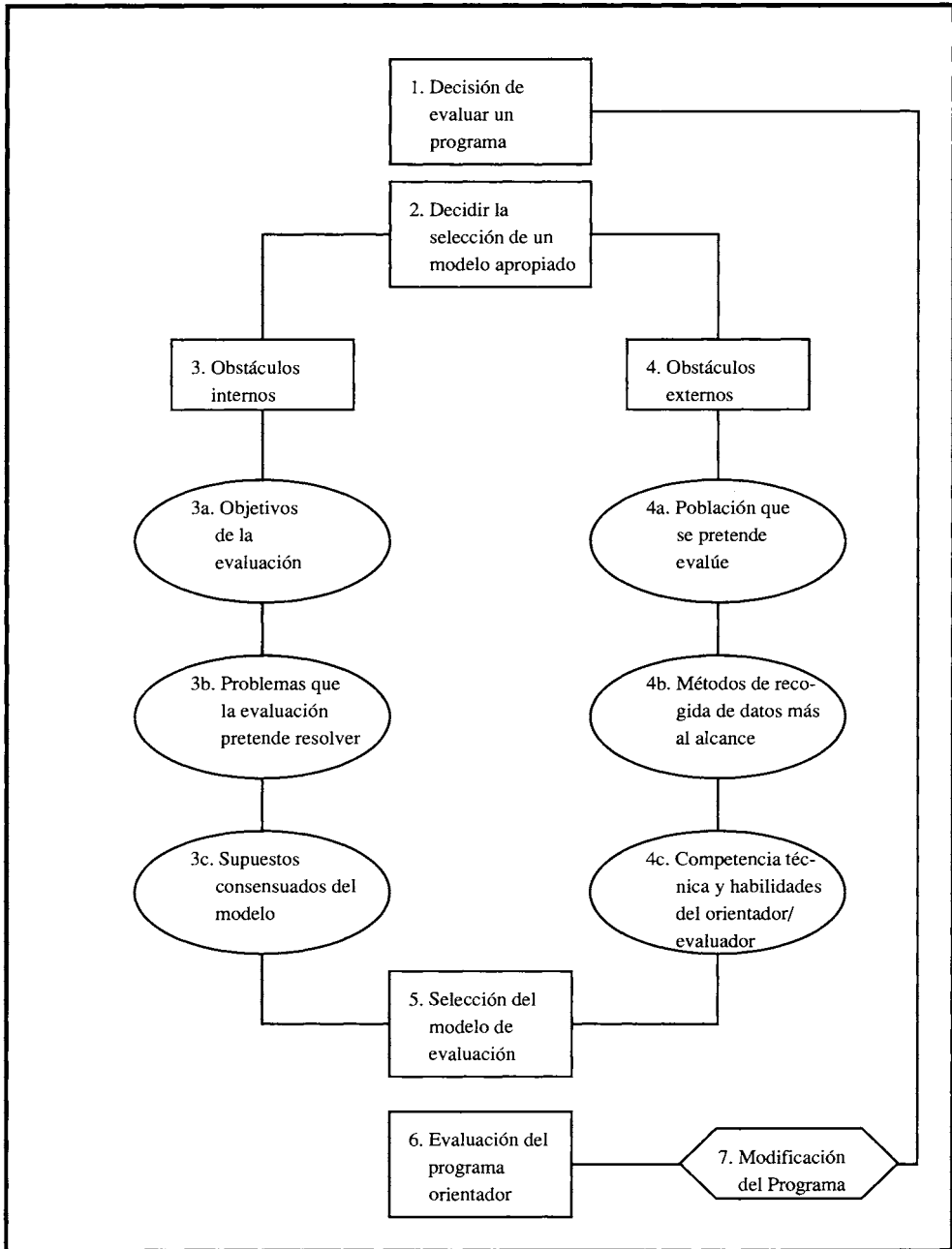
— *Modelo CIPP de Stufflebeam*

— *Modelo CSE de Alkin (variante del CIPP)*

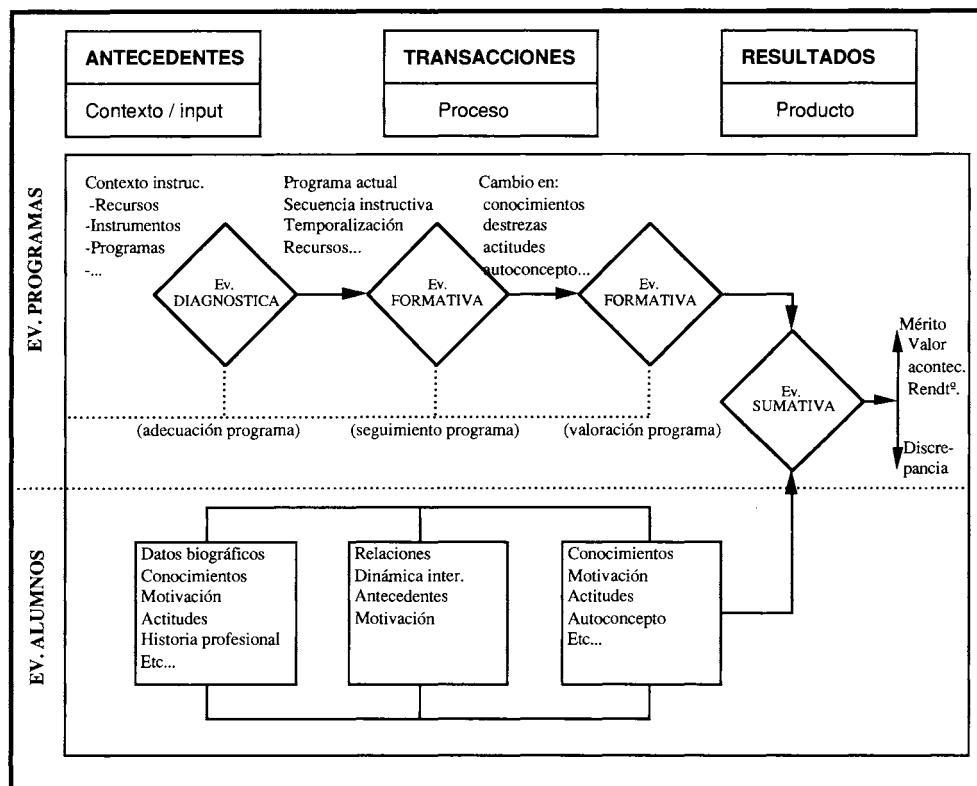
e) *Evaluación como negociación:*

— *Modelo respondiente construtivista de Guba y Lincoln*

CUADRO 3
METAMODELO DE EVALUACIÓN DE PROGRAMAS DE DANIELS (1981)



CUADRO 4
PROPUESTA DE MODELO DE EVALUACIÓN DE PROGRAMAS



La propuesta más reciente de Rebollo (1993) plantea la clasificación de los modelos en tres grandes bloques:

a) *Modelos objetivistas*, aquellos que se conciben desde una perspectiva técnica, científica. Algunos de los modelos que pueden incluirse aquí serían:

- Evaluación basada en objetivos (Tyler)
- Planificación evaluativa (Cronbach)
- C.I.P.P. (Stufflebeam)
- Evaluación sin referencia a objetivos (Scriven)

b) *Modelos subjetivistas*, que plantean la evaluación como estrategia de comprensión y valoración de los procesos y resultados de un programa. Se diferencia de la postura anterior, sobre todo, en la concepción de la realidad y en la forma de entender el conocimiento. Especialmente interesa evaluar los procesos. Podríamos incluir los modelos siguientes:

- Evaluación respondiente (Stake)
- Evaluación iluminativa (Parlett y Hamilton)
- Evaluación democrática (McDonald)

c) *Modelos críticos*, que entienden la evaluación de programas como un proceso de recogida de información que fomenta la reflexión y que persigue la transformación de los destinatarios del programa. La evaluación se centra en el análisis de las circunstancias personales, sociales y políticas de los destinatarios. El contenido básico será tanto el proceso como las actividades de su desarrollo. No hay modelos explícitamente vinculados a esta modalidad aunque si hay experiencias interesantes que siguen las pautas señaladas (Martínez y otros, 1990; Pozo y otros, 1990).

Aunque la selección de un modelo adecuado de evaluación debe realizarse en función de las características específicas del programa a desarrollar, consideramos pertinente ofrecer las pautas de carácter general que pueden ayudarnos a realizar la selección, presentadas por Daniels y otros (1981), y que configuran un metamodelo (cuadro 3).

Una propuesta concreta de modelo de evaluación, aplicada en este caso a la formación ocupacional, es la realizada por Tejada (1991) y en la que se excluye la evaluación previa a la planificación y se integran elementos teóricos de diversos modelos (Scriven, Stake y Stufflebeam). Dejamos al lector que realice su propia interpretación (cuadro 4).

5. FORMULACIÓN DE OBJETIVOS E INDICADORES

Tal y como viene planteándose la evaluación de programas, en cualquiera de sus modalidades o enfoques, la formulación de objetivos es una tarea fundamental que en buena medida va a condicionar el desarrollo del programa y las posibilidades reales de una evaluación correcta y fructífera. Por ello conviene sistematizar al máximo dicha tarea, lo que puede lograrse siguiendo cualquiera de las taxonomías o modelos de formulación de objetivos conocidos. En cualquier caso, la adecuación de estos objetivos a programas de intervención en el ámbito psicopedagógico, no exclusivamente de carácter instructivo, es una importante tarea a realizar ya que la mayoría de las taxonomías conocidas están pensadas específicamente para los procesos de enseñanza-aprendizaje.

Quizá por esta razón sigue proponiéndose (Colás, 1993) el modelo de Hammond como adecuado para la formulación de objetivos. El modelo, que aparece comentado en la referencia dada, incorpora cuatro dimensiones: programas, objetivos, acciones y agentes.

Realizada la selección de objetivos a partir de algunas pautas como las anteriormente reseñadas y antes de llegar a una propuesta definitiva, tendremos que asegurarnos de que:

- son relevantes, desde la perspectiva educativa, social, política,...., en la que se orienta el programa

- son útiles, para los sujetos a los que va destinado
- son susceptibles de ser conseguidos con la realización del programa, lo que implica que se han formulado tomando en cuenta el contexto y los recursos
- son susceptibles de ser evaluados.

Tarea nada fácil será convertir los objetivos seleccionados en indicadores. En evaluación de programas los indicadores se sitúan en la confluencia entre lo teórico y lo empírico y mientras que la definición conceptual establece el significado en términos abstractos, los indicadores se refieren a sus características observables, medibles y empíricamente detectables (como eslabón intermedio de esta cadena actuarán los objetivos operativos).

6. LOS DISEÑOS EN LA EVALUACIÓN DE PROGRAMAS

Un diseño es un plan que fija qué medidas y cuándo serán tomadas en el curso de un proceso de evaluación. Plantearse la especificación de un diseño supone haber realizado reflexiones previas en torno a cuestiones relacionadas con el concepto general de evaluación y con las finalidades de las diferentes opciones metodológicas; posteriormente delimitaremos, como parte de la planificación prevista, los fundamentos de la misma, los objetivos, los elementos personales implicados en el programa, los ámbitos de intervención, los instrumentos de recogida de datos, las técnicas de análisis y los momentos de la evaluación; todo ello, de acuerdo al modelo que se haya elegido para realizar la actividad evaluativa y teniendo en cuenta que en todos estos elementos existen diferencias según la perspectiva o enfoque (paradigma o modelo) por el que se haya optado.

En cualquier caso, independientemente del modelo que se haya elegido, cualquier mínima planificación tiene que atender las siguientes tareas:

- establecer las actividades que se van a evaluar
- fijar los criterios de evaluación
- elegir las estrategias para la obtención de información
- analizar la información
- tomar decisiones a partir de la información analizada

En su concreción, la variedad de diseños es amplísima ya que no hay limitación alguna en la elección, si bien dadas las características peculiares de la investigación evaluativa existen diseños más adecuados para planificar la intervención mediante programas. La propuesta de Fitz-Gibbon y Morris (1987, pp. 53-60) hace referencia a siete modalidades de diseño:

- 1) Diseño pretest post-test con grupo de control
- 2) Diseño post-test (únicamente) con grupo de control
- 3) Diseño pretest post-test con grupo de control no equivalente
- 4) Diseño de series temporales con un único grupo (experimental solamente)
- 5) Diseño de series temporales con grupo de control no equivalente
- 6) Diseño antes-después

7) Diseños complejos (experimentales y cuasiexperimentales, resueltos mediante análisis de varianza).

Creemos que esta propuesta olvida la referencia a determinadas estrategias descriptivas y correlacionales, razón por la cual sugerimos la siguiente:

A) Diseños para la investigación cuantitativa

- 1) Diseños descriptivos-comparativos (estudios tipo survey)
- 2) Diseños longitudinales
 - secuenciales (diacrónicos)
 - transversales (sincrónicos)
- 3) Diseños correlacionales:
 - correlación y regresión (discontinuidad en la regresión)
 - técnicas de clasificación
- 4) Diseños causales
- 5) Diseños experimentales:
 - a) preexperimentales
 - pretest con un solo grupo
 - pretest post-test con un solo grupo
 - b) cuasiexperimentales
 - pretest posttest con grupo control no equivalente
 - series con pretest y posttest con un solo grupo
 - series con grupo control
 - c) experimentales de campo
 - pretest con grupo control
 - pretest posttest con grupo control
 - n=1 (línea base)
 - intrasujetos
 - complejos

B) Diseños para la investigación cualitativa

- 1) Estudios de casos
- 2) Estudios del grupo
- 2) Estudios observacionales
- 3) Estudios de comunidades (etnográficos)
- 4) Estudios de situaciones o contextos
- 5) Estudios de investigación-acción

Nos parece importante insistir en esta distinción ya que tradicionalmente el diseño de evaluación se ha asociado a la práctica de evaluaciones sumativas, de carácter cuantitativo. Mientras que la evaluación formativa no suele planificarse de acuerdo a pautas de diseños cerrados, lo que no quiere decir que este tipo de

evaluación no requiera reflexiones en torno a la organización del trabajo a desarrollar.

Hay campos de aplicación de la acción educativa (educación compensatoria, educación especial) en los que la planificación de diseños supone muchas más dificultades que en otros (educación de adultos, proceso instructivo,...). En cualquiera de los campos, es fundamental admitir la diversidad de aplicaciones en función de la modalidad de evaluación: cabe pensar en un diseño para la evaluación de necesidades, en otro para la evaluación de contextos y en otro para la evaluación de resultados. Igualmente será importante el campo específico de trabajo, con ciertos tipos de diseños más adecuados para programas de educación especial, por ejemplo, y diseños más propios para programas de procesos instructivos.

7. FASES DEL PROCESO EVALUATIVO

Presentamos las fases correspondientes al diseño de programas de corte tradicional y de corte participativo.

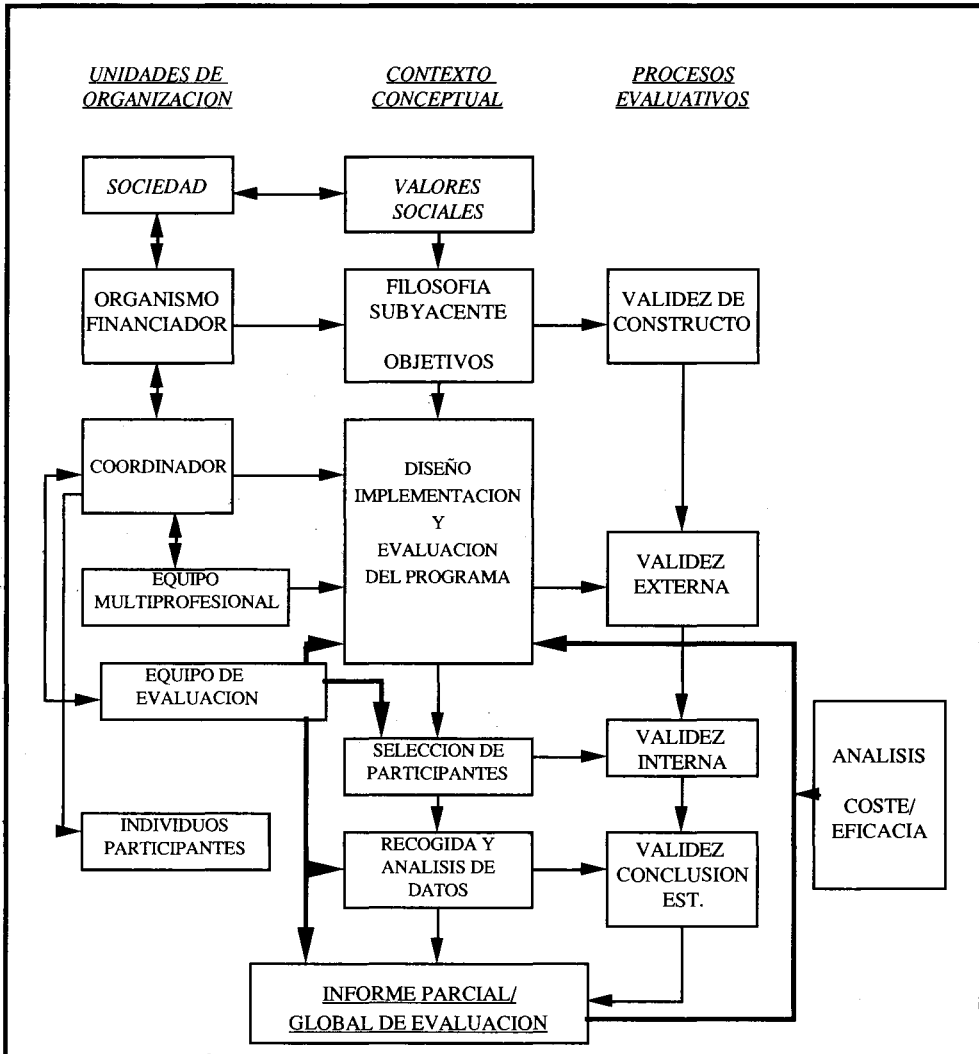
A) *Esquema tradicional*

El esquema de planificación de un programa incluiría las siguientes fases:

- 1) Análisis o evaluación de necesidades
- 2) Programación:
 - desarrollo de objetivos y procedimientos de evaluación
 - selección de estrategias o actividades para lograr los objetivos
- 3) Implementación:
 - desarrollo de estrategias
 - construcción de instrumentos, que implica:
 - delimitación de variables e indicadores
 - selección y listado de criterios
 - redacción de ítems
 - aplicación de instrumentos
 - seguimiento del proceso
- 4) Toma de decisiones (evaluación propiamente dicha):
 - sintetizar, analizar y valorar la información
 - proceso dinámico que genera información
 - retroalimentación del programa
 - carácter formativo vs sumativo

Un esquema de corte tradicional similar al presentado es el ofrecido por Alvira (1991, p. 22). Algo más exhaustivo y orientador de las actividades a llevar a cabo en la puesta en marcha de un programa es el modelo de Wortman (1983), que presentamos a continuación (cuadro 5).

CUADRO 5
MODELO DE WORTMAN



B) *Esquema participativo*

Las dos peculiaridades principales que desde un principio incorpora la investigación evaluativa respecto a otros esquemas más tradicionales de investigación (básica, aplicada,...), son, como ya hemos señalado, la participación de los interesados en algunas de las fases del proceso de evaluación y la utilización de los resultados para

la mejora del propio programa (feed-back para la toma de decisiones). Sin embargo, no siempre la investigación evaluativa ha respondido a estas expectativas, por lo que desaparecen los matices diferenciales de esta estrategia de investigación en relación con otras de corte más clásico.

Potenciar explícitamente esos matices como criterios de identificación más específicos de la investigación evaluativa es lo que pretenden conseguir algunas de las últimas propuestas realizadas (el «modelo de evaluación responsable» de Stake y el «modelo de evaluación dirigida a resultados» de Patton).

La investigación evaluativa tiene que preocuparse de proporcionar a los interesados la información que necesitan, estudiando los problemas que ellos consideran relevantes. El hincapié en estos aspectos no altera el flujo (proceso/fases) de una evaluación, pero sí cambia el énfasis y contenido de cada fase, de modo que la identificación de los problemas (objetivos) a través de contactos con los interesados (audiencia) y la consideración de los distintos métodos para establecer el consenso entre las partes, toma preminencia sobre otras fases de la investigación.

A partir de estos considerandos, y tomando como referencia el modelo de Patton, en el trabajo ya comentado de Alvira (1991, p. 31), nos presenta el esquema de actuación para el desarrollo de una investigación evaluativa que incorpora explícitamente los presupuestos participativos y que por su interés presentamos aquí al lector sin comentarios añadidos.

- 1) Identificación de partes interesadas
- 2) Organización con ellas de un grupo de trabajo
- 3) Delimitación de preguntas y temas de la evaluación
- 4) Anticipación de las respuestas y valoración de las mismas a fin de determinar si procede seguir con la evaluación
- 5) Decisiones sobre diseño, métodos y medición
- 6) Recogida de información
- 7) Análisis de datos
- 8) Implicación en el análisis del grupo de trabajo
- 9) Decisión sobre cómo utilizar y aplicar los hallazgos
- 10) Decisión sobre difusión de los hallazgos

8. TIPOS DE EVALUACIÓN EN FUNCIÓN DE LAS ETAPAS DEL PROGRAMA

8.1. Etapas en la Evaluación de Programas

Pueden establecerse básicamente tres etapas: evaluación de la formulación del programa; evaluación de la aplicación del programa y evaluación de los resultados del programa.

Estas diferentes etapas de la evaluación pueden presentarse como diferentes tipos de evaluaciones. Angulo (1988), como ya expusimos en un trabajo anterior (Tejedor, 1990, p. 20), relaciona las distintas inquietudes planteadas en el desarrollo

CUADRO 6
ESTRUCTURA METODOLÓGICA

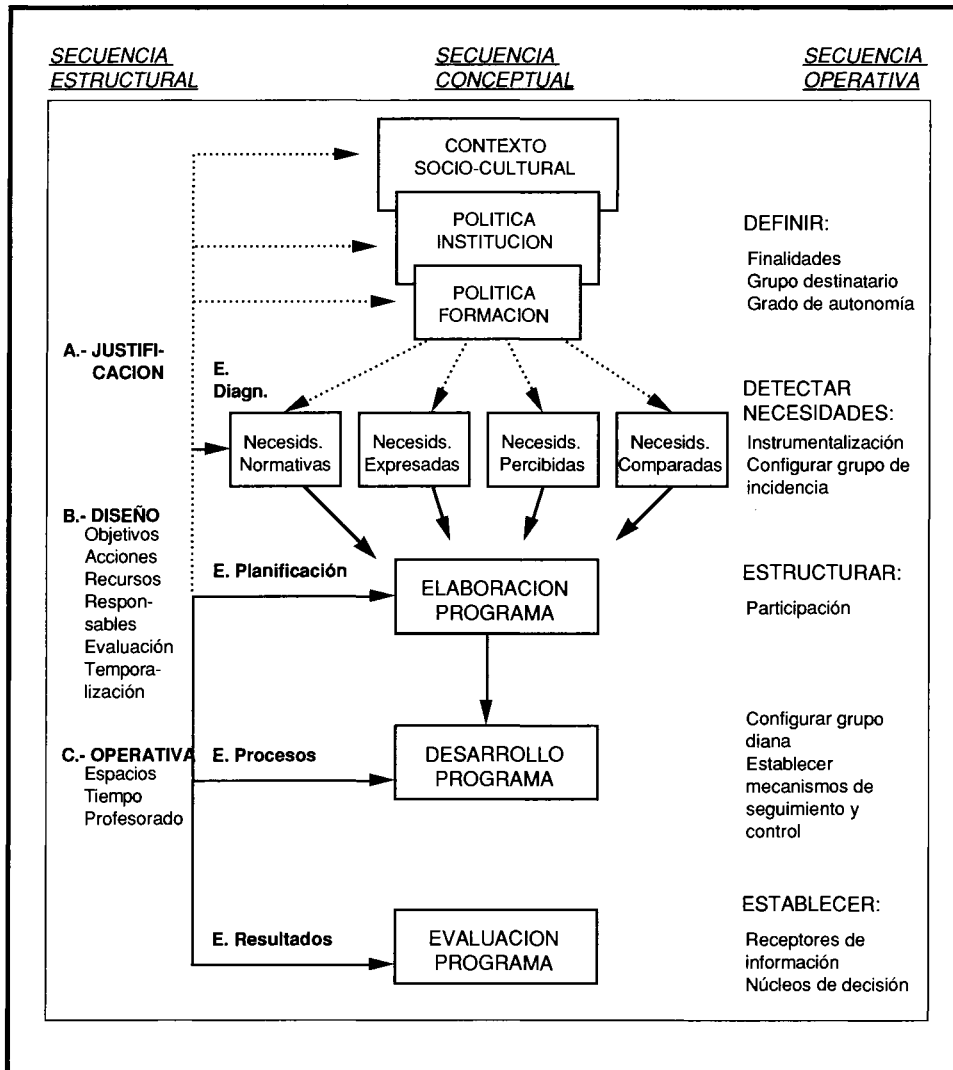
Estructura metodológica de la evaluación en el desarrollo comunitario			
Tipo de evaluación	Referente	Tarea	Ambito de aplicación
Del contexto necesidades	territorio y sociedad: comunidad, grupos, sujetos,...	estudio previo de la realidad:problemas, necesidades, recursos, conflictos,...	diagnóstico, planificación
Del diseño	expertos y profesionales en programar y planificar	análisis de criterios programáticos, justificación coherencia y pertinencia del diseño	diseño y planificación del programa
Del proceso	profesionales participantes en programa	análisis desarrollo del programa: estrategias, niveles de ejecución, participación,...	ejecución de programa
Del producto	equipo evaluador	valoración de los resultados en relación con objetivos:eficacia, eficiencia,...	conclusiones toma de decisiones

del programa con los diferentes tipos de evaluaciones. Podemos pensar también que, como recogemos en el cuadro 6, cada etapa del programa lleva asociada un tipo de evaluación (Caride, 1990), con especificación de ámbitos, tareas y referentes de evaluación.

Recientemente Alkin (1991) identificaba cinco tipos de evaluación vinculados a las distintas fases en la evaluación de programas:

- 1) Evaluación de sistemas (fase de preplanificación o de evaluación de necesidades)
- 2) Evaluación de planificación del programa (evaluación del diseño)
- 3) Evaluación de la puesta en marcha del programa (evaluación de proceso)

CUADRO 7
SECUENCIA DE ACTIVIDADES EN EVALUACIÓN DE PROGRAMAS



- 4) Evaluación de perfeccionamiento del programa (evaluación de resultados)
- 5) Evaluación de certificación del programa (evaluación de eficacia y evaluación de impactos).

Una panorámica óptima de la secuencia de actividades a desarrollar para la elaboración de programas es la que nos presenta Gairín (1991), que nosotros recogemos en el cuadro 7.

8.2. Contenidos o referentes de la evaluación

Desde una perspectiva metodológica, la puesta en marcha de un programa supone una actuación que debe atender ordenadamente a las distintas fases del proceso. La supervisión del desarrollo de cada una de esas fases en términos evaluativos nos lleva al establecimiento de diferentes tipos de evaluación que pueden llegar a tener interés en sí mismos, incluso con su metodología propia (necesidades, objetivos, proceso, resultados,...); pero no podemos olvidar que todos ellos forman parte de un esquema único de investigación que nos informará sobre la puesta en marcha del programa y la calidad de sus logros.

Tenemos entonces, de acuerdo con la fase del proceso a la que se atiende, la posibilidad de establecer diversos tipos de evaluación:

1) *Evaluación de necesidades*

Debe ser la primera actividad a llevar a cabo en la puesta en marcha de un programa de intervención. Aunque, como ya señalamos en escritos anteriores (Tejedor, 1990), algunos autores plantean la evaluación de necesidades con el mismo énfasis que la evaluación de programas, considero que es más pertinente concebirla como una fase de dicho proceso, limitando entonces sus objetivos, lo que no implica desvirtuar su esquema operativo que entiendo está suficientemente conformado.

Es decir, podemos pensar, desde un punto de vista metodológico, que la evaluación de necesidades, siendo una fase del proceso de investigación para la puesta en marcha de un programa de intervención, tiene entidad propia y es susceptible de ser analizada con un esquema completo de investigación evaluativa (Bank y Morris 1979; Witkin, 1984).

2) *Evaluación del diseño del programa de intervención*

Supone la preocupación por la secuencia de los pasos a seguir, por la claridad de formulación de los objetivos planteados, por la calidad de los instrumentos o técnicas de obtención de información, por la suficiencia de la muestra fijada, por la adecuada localización de las unidades muestrales, por la procedencia de las actividades sugeridas, por la adecuación de las técnicas de análisis,... (Mayer, 1985; Bartolomé, 1990).

3) *Evaluación de la evaluabilidad*

La experiencia determina que pocos de los programas llevados a la práctica son evaluables al tener problemas de especificación de objetivos, de implementación, de adecuación de las actividades a la consecución de objetivos, de carencia de medios para su desarrollo,... Por ello estamos de acuerdo con quienes sugieren la conveniencia de llevar a cabo una revisión previa de las posibilidades evaluativas del

CUADRO 8
EVALUABILIDAD DE MODELOS

El decisor político asigna recursos para la evaluación de un programa
Determinación de la evaluabilidad (actividades)
<p>Determinación de la evaluabilidad (productos)</p> <ol style="list-style-type: none"> 1) Objetivos y expectativas 2) Medidas del desempeño 3) Factibilidad de medir el desempeño 4) Usos probables de la información 5) Opciones de evaluación
<p>Uso de la información de evaluación</p> <ol style="list-style-type: none"> 1) Mantener o cambiar los objetivos 2) Mantener o cambiar las actividades 3) Mantener o cambiar las estrategias de desarrollo del programa 4) Seleccionar una o más de las opciones de evaluación

programa, lo que supone fundamentar adecuadamente las pautas metodológicas a seguir, dado que, al igual que la evaluación de necesidades, es un campo de referencia con tratamiento específico ya contrastado (Rutman, 1980; Shipman, 1989; Anguera, 1990; Berk y Rossi, 1990).

Wholey (1979) presenta un esquema para determinar la evaluabilidad de un programa (cuadro 8).

4) Evaluación de la implementación

Este tipo de evaluación pretende analizar el funcionamiento del programa y determinar si existen diferencias entre el funcionamiento real y el previsto inicialmente. Si hubiera diferencias notorias entre ambos esquemas de funcionamiento

trataríamos de hallar las causas y, si fuera posible, subsanar las deficiencias (Leithwood y Montgomery, 1980; Finnegan y otros, 1988).

5) *Evaluación de la cobertura*

Trataremos de analizar si el programa cubre todos los objetivos previstos y si llega a todos los sujetos para los que se había pensado, determinando, si procede, los sesgos de cobertura existentes y las causas que los han motivado.

6) *Evaluación del proceso (seguimiento)*

En algunos casos, sobre todo si el programa es de larga duración, conviene realizar un análisis de su funcionamiento a fin de verificar si su desarrollo sigue las pautas previstas. Implica la revisión de todos los elementos constitutivos del programa (personales, funcionales y materiales), en su dimensión interactiva: calidad de las actividades, distribución del tiempo y del trabajo, utilización de recursos,...

Pretende describir y valorar las actividades que se llevan a cabo durante el desarrollo del programa intentando solucionar los fallos que pudieran encontrarse.

La evaluación adecuada del proceso, en tanto en cuanto que implica que el programa se ha puesto en marcha, supone la consideración de todos sus elementos por lo que se sugiere un esfuerzo previo de reconsideración respecto a las fases anteriores de ejecución del programa (necesidades, objetivos, diseño,...).

7) *Evaluación de resultados*

Hasta hace relativamente poco tiempo era el único tipo de evaluación que se llevaba a cabo tras la finalización de las actividades que implicaba la ejecución del programa.

En términos generales, se trata de analizar el grado en que el programa consigue los resultados previstos (también pueden evaluarse, siempre que se hayan previamente identificado, resultados no previstos inicialmente). Se pretende determinar hasta qué punto pueden vincularse unívocamente los logros obtenidos con los efectos del programa, lo que nos sitúa de lleno en la problemática de la validez interna (control de variables) de todo trabajo de investigación, cuestión que, como es sabido, tiene especial interés en los diseños de intervención dada la dificultad para establecer nexos de causalidad y covariación entre las variables consideradas en el programa.

La evaluación de resultados admite la modalidad de «evaluación de verificación», que pretende efectuar comprobaciones a base de patrones claramente identificados, tanto en base a datos cuantitativos como cualitativos, y la modalidad de «evaluación comprensiva», que pretende ampliar o profundizar la comprensión de situaciones y que utiliza básicamente datos cualitativos.

La temática de la evaluación de resultados es amplia y compleja, apareciendo tratada en la bibliografía con diferentes terminologías: evaluación de productos, de la efectividad, de la eficacia, de la eficiencia, ... Las propuestas metodológicas (conceptualización del producto, selección de productos relevantes, diseños, instrumentos, medida y análisis de datos) para el análisis de resultados son igualmente muy variadas (Judd y Kenny, 1981; De la Orden, 1985 a, 1990).

En algunos casos, la evaluación de resultados se plantea en términos de «evaluación del impacto», aunque la mayoría de los estudiosos del tema reserven esa terminología para referirse a la evaluación de resultados en la población del entorno de los usuarios del programa; es decir, se trataría tanto de analizar las consecuencias no previstas al planificar el programa como de evaluar su incidencia en una población más amplia que la directamente receptiva del programa.

La estimación del impacto nos permite identificar y prevenir las consecuencias de un programa y desemboca en la valoración de los efectos, esperados o no. Las técnicas más utilizadas en la estimación de impactos son, en opinión de Anguera (1989):

- las matrices de revisión causa-efecto, para la búsqueda de relaciones entre las variables (para cada una de las actividades básicas de intervención se revisan los factores ambientales y/o sociales que puedan verse afectados)
- matrices cruzadas, para el establecimiento de relaciones primarias, secundarias,...
- modelos de simulación
- análisis de supervivencia
- estimación de demandas futuras
- modelos logit de respuestas multinómicas

8) *Evaluación económica*

Trata de analizar el coste del programa, tanto en términos absolutos como en términos relativos (respecto a resultados conseguidos).

El análisis de los costes surge en el ámbito de los estudios de las ciencias económicas y se plantea en relación directa con los beneficios. Como en la intervención social no tiene mucho sentido hablar de «beneficios», el estudio de los costes se plantea en términos de análisis «coste/efectividad» o «coste/utilidad», lo que vincula este tipo de evaluación a la evaluación de resultados, debiendo entenderse como una consecuencia de ella y, por tanto, realizarse con posterioridad (Levin, 1985).

9. LOS CRITERIOS Y NORMAS DE VALOR

Sin duda, uno de los problemas básicos que rodean a toda actividad evaluativa es el problema del establecimiento de los criterios, que nos van a permitir la emisión de juicios de valor respecto a la consecución o no de los objetivos planteados a

partir de las actividades llevadas a cabo. Es un proceso en el que pueden señalarse tres fases:

- 1) Especificación de los criterios que nos señalen lo que tiene que hacer o cómo tiene que funcionar un programa para ser etiquetado como bueno.
- 2) Para cada criterio hay que especificar estándares de funcionamiento que detallen niveles o grados de mérito.
- 3) La evaluación consistirá simplemente en la recogida de información respecto a esos criterios para estimar si se han alcanzado o no los estándares prefijados.

No siempre es fácil establecer criterios y estándares pero su determinación es un requisito previo para una adecuada evaluación. Desde luego, la formulación clara de objetivos, y más si éstos están debidamente operacionalizados, facilitará la tarea del establecimiento de los criterios (en algún caso, los objetivos así establecidos se identifican con los propios criterios).

Brunet (1987) señala los cuatro aspectos principales que hay que controlar para lograr que la medición del criterio sea lo más válida posible:

- la insuficiencia del criterio (comprobar si el criterio que se quiere medir incluye todos los objetivos deseados del programa)
- la contaminación del criterio (se refiere al control de ciertos elementos externos del programa que pueden afectar a los resultados evaluados)
- los sesgos causados por las características del grupo, que se producen cuando los participantes en el programa poseen ya las características medidas por el criterio
- la pluridimensionalidad del criterio, lo que exige evaluar el criterio con múltiples medidas.

Van de Ven y Ferry (1980) proponen una guía de seis puntos para llevar a cabo la tarea de fijación de los criterios de evaluación:

- 1) Deben escogerse criterios que tengan gradaciones para que puedan ser discriminativos.
- 2) Como es imposible medir todos los criterios que podemos considerar como relevantes de las distintas variables, conviene investigar aquellos que parecen explicar o están relacionados con el mayor número de variables (la parsimonia es por tanto un aspecto a tener en cuenta).
- 3) Se han de buscar aquellos criterios que son menos costosos de medir, tanto económicamente como en el tiempo.
- 4) Hay que determinar si cada criterio se asocia con una variable dependiente, independiente o moduladora.
- 5) Hay que clasificar los criterios por niveles de análisis.
- 6) Tanto los usuarios como los evaluadores han de realizar una exhaustiva revisión de los criterios obtenidos hasta alcanzar el consenso sobre los más adecuados.

Aunque cada programa deberá formular sus propios criterios, podemos hablar de criterios de valor de referencia general: utilidad, viabilidad, formalidad, exactitud

CUADRO 9
CRITERIOS E INDICADORES DE EVALUACIÓN

<u>CRITERIOS</u>	<u>INDICADORES</u>
<u>Pertinencia:</u> Adecuación de un programa con la política de formación y el contexto de aplicación.	<i>Nivel de coherencia: relación entre los objetivos asignados y los objetivos propuestos.</i>
<u>Actualización:</u> Adecuación de los objetivos del programa y las necesidades reales (sociales e individuales).	<i>Relación entre los objetivos propuestos y las necesidades detectadas (objetivos de formación-demanda/oferta laboral).</i>
<u>Objetividad:</u> Adecuación de las leyes y principios científicos.	<i>Relación entre los objetivos asignados y contenidos (selección y secuenciación de contenidos).</i>
<u>Aplicabilidad:</u> Posibilidad de puesta en práctica de los objetivos propuestos.	<i>Relación entre el programa y la inserción laboral.</i>
<u>Suficiencia:</u> Grado con que un programa satisface las necesidades detectadas.	<i>Nivel de exhaustividad: relación entre los objetivos asignados y las necesidades detectadas.</i>
<u>Eficacia:</u> Nivel de logro de los objetivos asignados.	<i>Relación entre los objetivos asignados y los objetivos alcanzados.</i>
<u>Eficiencia:</u> Grado de implicación de recursos humanos, materiales y funcionales.	<i>Relación entre los objetivos logrados y los recursos implicados (ratio profesor/alumno, hora/profesor/alumno, etc.)</i>
<u>Comprensividad:</u> Grado de optimización alcanzado.	<i>Relación entre el nivel de entrada y nivel de salida de un programa.</i>
<u>Relevancia:</u> Grado de importancia del programa para cubrir necesidades individuales y sociales.	<i>Relación entre los objetivos propuestos y necesidades sociales e individuales (objetivos de formación-necesidades, expectativas, intereses, cambio laboral y polivalencia laboral, etc.)</i>

(precisión y validez), contextualización y diversidad en la valoración de las consecuencias/efectos del programa, juzgadas desde el punto de vista de los técnicos, de los gestores y de los usuarios.

Ferrández (1990) realiza una interesante propuesta en la que aparecen detallados criterios e indicadores, con un matiz operativo y referido a la puesta en marcha de un programa de formación ocupacional (cuadro 9).

Desde una perspectiva más global (considerando la evaluación del programa en su totalidad), también podemos hablar, de forma genérica, de criterios o normas para la realización de un proceso evaluativo.

Las normas que deben orientar las realizaciones de las prácticas evaluativas, definidas como principios comúnmente aceptados para determinar el valor o la calidad de una evaluación, han sido fijadas por colectivos de investigadores, pudiendo citarse como referencias claves los acuerdos del Joint Committee on Standards for Educational Evaluation de 1975 y de 1981 (Standards for Evaluation of Educational Programs, Projects and Materials, publicadas bajo la supervisión de Stufflebeam).

Su establecimiento pretende proporcionar, en opinión del propio Stufflebeam (1991):

- definiciones operativas para la evaluación de programas y de personal
- grados de acuerdo respecto a métodos y significados de la evaluación educativa
- estrategias para afrontar los problemas evaluativos
- guías prácticas para planificar las evaluaciones
- criterios aceptados para valorar los planes de las evaluaciones
- referentes conceptuales para fundamentar la evaluación
- pautas para llegar a la profesionalización de la tarea evaluadora
- diversos puntos de vista sobre los diferentes tipos de evaluación
- sugerencias para formar a los evaluadores.

Una línea similar siguen las conocidas 95 tesis para la reforma de la práctica evaluativa de Cronbach (1980).

10. TÉCNICAS DE RECOGIDA Y ANÁLISIS DE LA INFORMACIÓN

10.1. Técnicas de recogida de datos

La información que se precisa en cada tipo de estudio evaluativo es muy diversa y por tanto el único consejo válido al evaluador es seleccionar adecuadamente de entre el amplio espectro de técnicas las que sean más adecuadas en función de los objetivos del estudio y de las posibilidades reales del mismo, recordando que no hay técnicas específicas aunque sí existen algunas más adecuadas que otras para recoger información pertinente según la modalidad de evaluación que se está llevando a cabo.

Sin pretender ser exhaustivos, recogemos en el cuadro 10 los instrumentos adecuados en función del aspecto a evaluar.

CUADRO 10
INSTRUMENTOS DE MEDICIÓN

<p>1) Características de los sujetos</p> <p>inteligencia-aptitudes</p> <p>personalidad</p> <p>rendimiento</p> <p>actitudes, opiniones, intereses</p> <p>datos sociológicos</p>	<p>test estandarizados</p> <p>test y cuestionarios estandarizados</p> <p>pruebas estandarizadas y no estandarizadas</p> <p>escalas, cuestionarios, entrevistas, encuestas diarias</p> <p>entrevistas, cuestionarios</p>
<p>2) Características sociales del grupo</p> <p>relaciones del grupo</p> <p>contexto económico y social</p> <p>sistemas organizativos</p>	<p>test sociométrico</p> <p>observación</p> <p>incidente crítico</p> <p>escalas BARS</p> <p>indicadores, cuestionario</p> <p>documentos, encuesta</p> <p>entrevistas</p> <p>técnica delphi</p> <p>técnicas negociación</p>

No pensamos que la utilización de estas técnicas de obtención de información en el campo de la evaluación de programas tenga que responder a criterios diferentes a los establecidos para su utilización en otros campos de la investigación social. Sin duda, lo más importante a tener en cuenta desde un principio será adaptar el proceso de construcción de instrumentos (cuestionarios, entrevistas, escalas de observación,...) al tipo de evaluación que se quiera llevar a cabo (descriptiva, comprensiva); a los objetivos y condiciones específicas del trabajo a realizar; conocer a quién van

dirigidos (padres, alumnos, profesores, autoridades,...) y las posibilidades de aplicación (amplitud de la muestra, disponibilidades económicas,...); y disponer de los recursos necesarios para codificar y analizar adecuadamente la información recogida así como para difundir los informes pertinentes.

En escritos anteriores hemos comentado las peculiaridades de estas técnicas en su aplicación a la evaluación de necesidades (Tejedor, 1990). Pensamos que aquellas opiniones son extensibles al marco más general de la evaluación de programas por lo remitimos al lector interesado a dicho escrito.

10.2. Técnicas de análisis de datos

Insistiendo en algunas opiniones ya expuestas en escritos anteriores (Tejedor, 1990), creemos que con las técnicas a utilizar en el análisis de datos de la evaluación de programas ocurre otro tanto que lo señalado en relación a los instrumentos; no existen técnicas específicas pero conviene seleccionar las que mejor se ajusten a los presupuestos metodológicos implicados: tipo de evaluación seleccionada, peculiaridades del diseño, instrumentos utilizados en la recolección de datos, características de las variables y tipo de información esperada.

Las consideraciones a tener en cuenta para la elección de la técnica concreta de análisis incluirá referencias a:

- adecuación de la técnica al enfoque dado en la investigación
- posibilidad de evaluar los objetivos formulados a partir de los datos procedentes de los análisis sugeridos
- adecuación al tipo de datos obtenidos (por tanto, posibilidad de utilizar distintas técnicas si tenemos distintos tipos de datos)
- disponibilidad de la infraestructura para el tratamiento adecuado de la información (análisis de entrevistas, número de observadores, soporte informático,...).

Es un campo en el que no se puede, a priori, pensar en la improcedencia de aplicación de técnica alguna, ya que se mueve entre la aplicación del análisis de diarios y la de los modelos matemáticos rigurosos; que puede utilizar desde las técnicas de observación menos estructuradas a los diseños experimentales de campo; que puede incorporar datos de muy variada naturaleza,... Stufflebeam y otros (1984, pp. 88-89) presentan una relación interesante de las técnicas que pueden asociarse con este tipo de estudios.

Aún aceptando este contexto de diversidad, pienso que pueden señalarse algunas técnicas como más específicamente adaptables al análisis de datos en el estudio de la evaluación de programas:

- a) Todas aquellas que permiten comparar resultados o situaciones diversas, bien por la vía del establecimiento de diferencias significativas bien por la vía del reconocimiento de ganancias o pérdidas.
- b) Aquellas técnicas adaptables al análisis de contingencia, tanto bidimensiona-

les como multidimensionales. En este sentido consideramos del todo pertinente las aportaciones recientes en este campo que nos posibilitan desde el estudio de relaciones simétricas o asimétricas al estudio de modelos descriptivos (log-linear) o explicativos (logit); en todos los casos con posibilidad de estimar los parámetros del modelo, de analizar las interacciones entre las variables y de verificar el ajuste del modelo respecto a la hipótesis de trabajo planteada.

- c) Aquellas técnicas que están específicamente pensadas para el establecimiento de tipologías, tanto de variables como de sujetos (análisis clusters) y que actualmente pueden adaptarse a datos correspondientes a escalas de medida muy diversas (métricas o no métricas).
- d) Aquellas técnicas cuyo objetivo básico es el establecimiento de relaciones de covariación y causalidad. Este grupo de técnicas (regresión y análisis causal) merecen una mención específica ya que pueden permitirnos —el análisis causal— no sólo detectar las relaciones entre las variables sino establecer las causas que las originan. Para algunos investigadores, el análisis causal incluso sobrepasa el carácter de técnica específica para alcanzar el nivel de modelo, de estrategia. Interés particular puede tener el método de análisis causal denominado Fault Tree Analysis (FTA), tanto en su dimensión cualitativa como cuantitativa y que tanta utilización provechosa ha tenido en el estudio del establecimiento de causas que han orientado la toma de decisiones, la solución de problemas y la adopción de prioridades (Witkin, 1982).
- e) Las técnicas vinculadas al estudio de datos de carácter longitudinal, en sus dimensiones diacrónica y sincrónica (diseños secuenciales, time-lag, transversales, secuencial-transverso,...).
- f) Por último, podemos hacer referencia a aquellas técnicas que se vinculan de una forma específica con los procesos de toma de decisiones, es decir, con la elección de una determinada opción entre diversas alternativas. Cada opción dará lugar a un resultado distinto que puede ser medido en términos de utilidad, coste, beneficio o cualquier otra magnitud, dependiendo del problema que se esté considerando. Nos va a interesar reducir la incertidumbre que conlleva la elección de una opción, lo que se conseguirá a medida que seamos capaces de prevenir.

Desde un punto de vista metodológico, los métodos de previsión pueden agruparse en dos grandes bloques: cualitativos y cuantitativos. Los cualitativos se utilizan cuando el pasado no proporciona una información directa sobre el fenómeno considerado; las técnicas estadísticas desempeñan en ellos un papel secundario y lo más importante es contar con expertos dotados de intuición y sagacidad. Podríamos incluir entre estas técnicas el brainstorming, el delphi y el cross-impact.

En las posiciones de carácter cuantitativo se parte del supuesto de que se tiene registrada información sobre el pasado, generalmente en forma de series temporales (conjunto de observaciones acerca de una variable, observada a intervalos regulares

de tiempo). La misión del estadístico consiste en extraer toda la información posible contenida en los datos y, en base al patrón de conducta seguida en el pasado, realizar conjeturas sobre el futuro. En el análisis de series de datos de una única variable suele interesarnos conocer sus componentes (tendencias, factores cíclicos, movimientos estacionales y movimientos irregulares), de forma global o local (alisados exponenciales). Si el objetivo, más complejo, es identificar los modelos generados de las observaciones y, una vez verificado el modelo, predecir valores futuros, estaremos en el campo de los modelos ARIMA (utilización plena del análisis de series temporales).

Estas técnicas apenas si se han incorporado a la investigación educativa; quizá la consideración rigurosa del estudio de la evaluación de programas nos exija una mayor familiarización con ellas si no queremos depender excesivamente de otros investigadores que sí las han incorporado a este tipo de estudios con probada utilidad.

REFERENCIAS SOBRE PRÁCTICAS DE EVALUACIÓN DE PROGRAMAS LLEVADAS A CABO EN NUESTRO PAÍS

Recogemos, sin ánimo de ser exhaustivos, algunas referencias sobre trabajos de evaluación de programas de intervención llevados a cabo en España en los últimos años.

- ABARCA, P. (coord.) (1989): *La evaluación de Programas Educativos*. Escuela Española, Madrid. Esta publicación recoge los trabajos presentados a la reunión anual de AEDES, celebrada en Madrid, con aportaciones a la evaluación de programas en distintos ámbitos psicopedagógicos.
- AGUILERA, M^a J. y otros (1990): *Evaluación del programa de integración escolar*. CIDE, Madrid.
- ÁLVARO, M. y otros (1990): *Evaluación externa de la reforma Experimental de las E.E.M.M. (II)*: CIDE, Madrid.
- ALVIRA, F. (1989): Los centros ocupacionales y los centros de recuperación de minusválidos físicos del INSERSO. Un ejemplo de evaluación externa. *Anales de Psicología*, nº 5, 67-76.
- BARRIO, V. del (1989): Diseño de evaluación de un Programa de Educación Vial. *Anales de Psicología*, nº 5, 77-102.
- DUEÑAS, M. L. (1993): *Evaluación de un programa de apoyo a la integración escolar de deficientes mentales* (Madrid): Citado y comentado en Colás y Rebollo (1993).
- GARCÍA PASTOR, C. y GARCÍA JIMÉNEZ, E. (1993): *Evaluación de los efectos de la integración escolar en niños deficientes visuales* (Sevilla): Citado y comentado en Colás y Rebollo (1993).
- MARTÍNEZ, A. y otros (1993): *Evaluación de los programas de intervención en las instituciones de menores* (Valencia): Citado y comentado en Colás y Rebollo (1993).
- VARIOS (1989): *Evaluación de la Integración Escolar (2º Informe)*: MEC. Dirección General de Renovación Pedagógica. Madrid.
- VARIOS (1990): *Evaluación del Programa de Integración de Alumnos con Necesidades Educativas Especiales*. Tercer informe. MEC. Madrid.
- VARIOS (1992): *Evaluación de la calidad de los Servicios del Programa de Termalismo Social*. Ministerio de Asuntos Sociales, INSERSO, Madrid.

REFERENCIAS BIBLIOGRÁFICAS

- ABARCA, P. (coord.) (1989): *La evaluación de Programas Educativos*. Escuela Española, Madrid.
- ALKIN, M. C. (1991): Evaluation theory development. En M. W. McLaughlin y D. C. Phillips (eds.): *Evaluation and education*. At Quarter Century, National Society Of Education, Chicago.
- ÁLVAREZ, M. y otros (1988): «Evaluación de programas de Orientación: La evaluación del contexto y del diseño». Ponencia presentada al *IV Seminario de Modelos de Investigación Educativa*. Santiago de Compostela.
- ALVIRA, F. (1991): *Metodología de la evaluación de programas*. CIS, Cuadernos Metodológicos nº 2, Madrid.
- ANGUERA, M. T. (1989): Innovaciones en la metodología de la evaluación. *Anales de Psicología*, nº 5, pp. 14-42.
- ANGUERA, M. T. (1990): Programas de intervención. ¿Hasta qué punto es factible la evaluación? *Revista de Investigación Educativa*, nº 16, pp. 77-93.
- ANGULO, J. F. (1988): «Evaluación de programas sociales: De la eficacia a la democracia». *Revista de Educación*, 286, pp. 193-207.
- ATO, M. y otros (1989): Evaluación de programas: Aspectos básicos. *Anales de psicología*, nº 5, pp. 3-12.
- BANK, A. y MORRIS, L. L. (1979): *A needs assessments kit: Guidelines and Resources for Educators*. Center the Study of Evaluation. University of California, Los Ángeles.
- BARTOLOMÉ, M. (1990): Evaluación y optimización de los diseños de intervención. *Revista de Investigación Educativa*, nº 16, pp. 39-60.
- BLANCHET, A. (1989): «Entrevistar» en A. Trognon y otros: *Técnicas de investigación en ciencias sociales*. Narcea, Madrid.
- BERK, R. y ROSSI, P. M. (1990): *Thinking about program evaluation*. Sage, Beverly Hills.
- BRUNET, L. (1987): La investigación evaluativa en el medio organizacional. En Y. Bordelau: *Modelos de investigación en recursos humanos* (pp. 122-149): Trillas, México.
- CAMPBELL, D. T. y STANLEY, J. C. (1973): *Diseños experimentales y cuasiexperimentales en la investigación social*. Amorrortu Editores, Buenos Aires (versión española de 1978).
- CARIDE, J. A. (1989): «De la evaluación de necesidades a la evaluación de programas sociales en el desarrollo comunitario». Ponencia inédita.
- COHEN, E. y FRANCO, R. (1989): *Evaluación de proyectos sociales*. G.E.L., Buenos Aires.
- COLÁS, P. (1992): Evaluación de programas en el estudio de cuestiones educativas. *Cuestiones Pedagógicas*, nº 4-5, pp. 79-91.
- COLÁS, P. y REBOLLO, M. A. (1993): *Evaluación de programas. Una guía práctica*. Editorial Kronos, Sevilla.
- COOK, T. D. y REICHARDT, CH. (1986): *Métodos cualitativos y cuantitativos en la investigación educativa*. Morata, Madrid.
- COOK, T. D. y SADISH, W. R. (1986): Program Evaluation: The worldly science. *Annual Review of Psychology*, 37, 193-231.
- CRONBACH, L. J. (1980): *Toward reform of program evaluation*. Jossey-Bass, San Francisco.
- DANIELS, M. H. y otros (1981): A meta-model for evaluating counseling programs. *Personnel and Guidance Journal*, 59, 578-582.
- DE LA ORDEN, A. (1985 a): Hacia una conceptualización del producto educativo. *Revista de Investigación Educativa*, nº 16, 61-76.
- DE LA ORDEN, A. (1985 b): Investigación evaluativa. En A. de la Orden: *Investigación Educativa*. Diccionario de Ciencias de la Educación (pp. 133-137): Anaya, Madrid.

- DE LA ORDEN, A. (1990): Evaluación de los efectos de los programas de intervención. *Revista de Investigación Educativa*, nº 16, pp. 61-76.
- DUNN, R. y DUNN, K. (1977): *Administrator's guide to new programs for faculty mangement and evaluation*. Parker Publ., New York.
- FERRÁNDEZ, A. (1990): Identificación de la evaluación en Educación de Adultos: las alternativas. *V Jornadas de Educación de Adultos*. Zaragoza.
- FINNEGAN, J. R. y otros (1988): Measuring and tracking education program implementation: The Minnesota Heart Health Program. *Health Educational Quarterly*, 16 (1), 77-90.
- FITZ-GIBBON, C. T. y MORRIS, L. L. (1978): *How to design a Program Evaluation*. Jossey-Bass, San Francisco.
- FLANAGAN, J. C. (1954): «The critical incident technique». *Psychological Bulletin*, 54, 327-358.
- FORNI, F. (1988): *Formulación y evaluación de proyectos de acción social*. Humanitas, Buenos Aires.
- GAIRÍN, J. (1991): *La evaluación de programas y cursos*. Documento policopiado. Bellaterra, Barcelona.
- GARANTO, J. (1989): Modelos de evaluación de programas educativos. En *La evaluación de Programas Educativos*. Escuela Española, Madrid, pp. 43-78.
- JUDD, C. M. y KENNY, D. A. (1981): *Estimating the effects of social interventios*. Cambridge Univ. Press, New York.
- LEITHWOOD, J. y MONTGOMERY, D. J. (1980): Evaluating program implementation. *Evaluation Review*, 4 (2), 193-214.
- LEVIN, H. M. (1985): *Cost-effectiveness*. Sage, Beverly Hills.
- MARTINEZ, A. y otros (1990): Evaluación de los programas de intervención en las instituciones de menores (Valencia): *Revista de Investigación Educativa*, nº 16, pp. 325-331.
- MASSONNAT, J. (1989): «Observar» en A. Trognon y otros: *Técnicas de investigación en ciencias sociales*. Narcea, Madrid.
- MAYER, R. R. (1985): *Policy and program planning*. Prentice Hall, Englewood Cliffs, N.J.
- POPHAM, W. J. (1980): *Problemas y técnicas de evaluación educativa*. Anaya, Madrid.
- POZO, M. T. y otros (1990): Evaluación de un programa de intervención en educación ambiental. La investigación-acción como estrategia de conservación del medio natural. *Revista de Investigación Educativa*, nº 16, pp. 309-315.
- RUTMAN, L. (1980): *Planning useful evaluations. Evaluability assessment*. Sage, Berverly Hills.
- SANZ ORO, R. (1990): *Evaluación de programas de orientación educativa*. Pirámide, Madrid.
- SELDIN, P. (1985): *Changing practices in faculty evaluation*. Jossey-Bass, San Francisco.
- SHIPMAN, S. (1989): General criteria for evaluating social programs. *Evaluation Practice*, 10 (1), 20-26.
- STUFFLEBEAM, D. L. y otros (1984): *Conducting Educational Needs Assessment*. Kluwer-Nighott Publ., Boston.
- STUFFLEBEAM, D. y SKINKFIELD, A. (1987): *Evaluación sistemática: Guía teórica y práctica*. Paidos-Mec, Madrid.
- STUFFLEBEAM, D. (1988): Normas para evaluadores. En R. Pascual (eds.): *La gestión educativa ante la innovación y el cambio*. Narcea, Madrid.
- STUFFLEBEAM, D. (1991): Professional standards a ethics for evaluators. En M. W. McLaughlin y D. C. Phillips (eds.): *Evaluation and education*. At Quarter Century, National Society Of Education, Chicago.
- TALMAGE, H. (1982): Evaluation of programs. En *Encyclopedia of Educational Research*. McHollan, New York.
- TEJADA, J. (1991): *La evaluación en la formación ocupacional*. Documento policopiado. Bellaterra, Barcelona.
- TEJEDOR, F. J. (1990): Perspectivas metodológicas del diagnóstico y evaluación de necesidades. *Revista de Investigación Educativa*, nº 16, pp. 15-38.

- VAN DE VEN, A. H. y FERRY, D. L. (1980): *Mesaruring and assessing organizations*. Wiley, New York.
- WEIS, C. H. (1982): *Investigación evaluativa. Métodos para determinar la eficiencia de los programas de acción social*. Trillas, México.
- WHOLEY, J. S. (1979): *Evaluation: Promise and Perfomance*. The Urban Institute, Washington.
- WITKIN, B. R. (1982): «New York Association for Learning Disabilities Project: Formative Evaluation-Fault tree Compnent. Final report» en K. G. Butler y otros *Evaluation of Sulffolk and Nassau County Chapter Projects*. Syracuse University.
- WITKIN, B. N. (1984): *Assessing needs in Educational and Social Programs*. Jossey-Bass Publ., London.
- WORTMAN, P. (1983): Evaluation research: A methodological perspective. *Annual Review of Psychology*, 34, 223-240.

