

Revista de Investigación Educativa, 2000, Vol. 18, n.º 2, págs. 341-356

LA VALIDEZ EN LA EVALUACIÓN DE PROGRAMAS

J. M. Jornet, J. M. Suárez y A. Pérez Carbonell¹

INTRODUCCIÓN

El tema de la Validez, en cualquiera de los ámbitos que afectan a las CC. De la Educación, es un tema recurrente y, por lo tanto, tratado con exhaustividad y acierto por la gran mayoría de autores que, en algún momento, lo han atendido. Por ello, más que acercarnos a él desde una perspectiva estrictamente de fundamentación metodológica, en la que poco se puede aportar, estimamos que quizá es más adecuado revisar este tema desde las implicaciones que se derivan de las experiencias evaluativas en las que hemos estado inmersos, refiriendo eso sí los componentes teórico-metodológicos con los que nos sentimos más identificados. Por ello, este trabajo lo estructuramos en dos grandes apartados: a) Del concepto de Validez, y b) Estrategias de consecución de la Validez.

I. DEL CONCEPTO DE VALIDEZ

El Concepto de Validez es dependiente del proceso a que se aplique, así como de las características de la posición epistemológica desde la que se aborde. Por ello, estimamos necesario antes de referirnos al Concepto de Validez, definir nuestra posición respecto a la Evaluación de Programas, como proceso, y en cuanto a su orientación epistemológica y metodológica.

La *Evaluación de Programas*, como Investigación Aplicada, viene especificada fundamentalmente por su objeto, su finalidad o finalidades así como por el contexto. Como proceso de investigación, éste debe ser riguroso, estar bien controlado y realizarse de forma sistemática. Entendemos, pues, que la Evaluación de Programas (Alvira, 1996; Pérez Carbonell, 1998 y 1999; Rogers y Owen, 1999; Pérez Juste, 2000):

¹ Dpto. MIDE de la Universitat de València. Fac. Filosofía y CC. De la Educación. Avda. Blasco Ibáñez, 30. Tél. 963864430. e-mail: jornet@uv.es, rodrigus@uv.es, pereza@uv.es

- Requiere del mismo grado de control, rigor y sistematización que se le exige a cualquier otro tipo de investigación que pretende alcanzar un conocimiento, máxime teniendo en cuenta que pretende identificar elementos para intervenir sobre lo evaluado (es decir, de una u otra forma, pretende servir de base para un proceso en que se implican decisiones).
- Constituye un tipo de investigación con características muy definidas:
 - *Tiene una utilidad inmediata* que exige la aportación de un alto grado de confianza en los resultados que de ella se ofrecen así como en los procedimientos utilizados para alcanzar dichos resultados.
 - *Esta impregnada de cuestiones de valor*, como ¿conviene modificar algo?, ¿estamos realizando un trabajo efectivo?, ¿se está alcanzado lo pretendido?...
 - *Se encuentra condicionada por factores sociales y políticos* que la involucran en situaciones excesivamente complejas.
- Por ello, precisa de una *Planificación* en la que se anticipe el comportamiento de los componentes/elementos que intervienen en su propio proceso.

No obstante, bajo el término *Evaluación de Programas* coexisten diversos tipos de aproximaciones, que resultan en ocasiones muy divergentes en cuanto a los planteamientos epistemológicos y metodológicos en que se sustentan, o bien, simplemente, en cuanto a sus campos de actuación. En este caso, la reflexión que aportamos se centra respecto a la Evaluación de Programas como campo metodológico, es decir, asumiéndola como sinónimo de la Investigación Evaluativa.

Si realizamos un *breve recorrido histórico* sobre las corrientes que han estado presentes en el panorama de la Investigación Educativa, y que por lo tanto influyen en la delimitación del concepto de Validez en la Evaluación de Programas, hay que señalar que ha sido el positivismo y postpositivismo con su propio método científico el que se presenta con una mayor tradición científica dentro de las Ciencias de la Educación. Sin embargo, en esta tradición la aplicación de estas aproximaciones no ha sido uniforme y se puede observar un continuo que va desde el cumplimiento absoluto de todos los aspectos que marcan y caracterizan un planteamiento científico-naturalista a nivel epistemológico, ontológico y metodológico, hasta el cumplimiento más laxo de los mismos (De Miguel, 1988; Suárez, 1989; Dendaluze, 1995, 1997; Aliaga, 2000). Esto quiere decir, que si bien en la metodología Positivista y Postpositivista hay una serie de requisitos que deben estar presentes para llegar a estimar el impacto de un posible agente causal (variable Programa) en algún efecto potencial (variable Resultado al implementar el Programa), como son el control y la aleatorización, lo están pero con graduaciones que van desde un mayor a un menor cumplimiento. Hablamos, trasladándolo a la terminología que utilizamos, del continuo que va desde lo Experimental a lo No Experimental pasando por lo Cuasi-experimental.

En el momento en el que el positivismo (años 60 hasta años 80) recibe mayores críticas es cuando empieza a tomar fuerza, en cuanto a metodología de investigación en Educación, el *Constructivismo*. Como es sabido, este enfoque está integrado por una gran variedad de fuentes y que confluyen en torno al concepto de *Verstehen* (comprensión de los significados de las acciones humanas, Jacobs, 1987). Su finalidad es la de

llegar a comprender la intencionalidad en la actuación de un agente y no si ese agente es la causa o el efecto. Desde esta perspectiva, la explicación causal y la predicción del positivismo, es sustituido por la Comprensión —de la intencionalidad de las acciones y de las percepciones de los sujetos que pueden contemplarse desde la Investigación fenomenológica, la etnográfica, la participativa, la Investigación-Acción, la fenomenológica...—.

Superado el debate paradigmático (Cook y Reichardt, 1986; Cook y Shadish, 1994; Reichardt y Rallis, 1994), en los últimos años los paradigmas *Cualitativos*, y en especial el *Interpretativo* y para el *Cambio*, influyen de forma decisiva en todos los acercamientos que pueden enmarcarse en la Investigación Educativa, y, con mayor énfasis, en sus ámbitos más aplicados (Colás, 1995; Chelimsky y Shadish, 1997). Las posiciones actuales más representativas abogan por la *Complementariedad Metodológica* (Krathwohl, 1987; Mateo, 1990; Tejedor, García-Valcárcel y Rodríguez, 1994; Dendaluze, 1995; Creswell, 1994, 1998; Bericat, 1998; Buendía, 1998). Lo cierto es que en el ámbito de la Evaluación de Programas, es difícil dar respuesta ajustada a los problemas que aborda, si no nos situamos en una posición flexible. Desde ella nos aproximaremos a la revisión del Concepto de Validez.

Desde este punto de vista, en la Evaluación de Programas se pueden identificar diversas *facetas del concepto de Validez* de especial énfasis. Para acercarnos a ellas partimos de una definición genérica de la Validez en la Evaluación de Programas:

El grado en que un proceso de Evaluación de Programas atiende de forma adecuada los componentes metodológicos implicados con el fin de:

- *representar de forma realista las características del objeto que evalúa,*
- *facilitar la consecución de los fines establecidos para el Plan de Evaluación,*
- *respetar los intereses de todos los implicados o afectados por la evaluación,*
- *representar el contexto en el desarrollo del Plan de Evaluación, en la modulación de la interpretación de resultados y en la adecuación de las consecuencias de la evaluación, y*
- *coadyuvar a la mejora o innovación de lo evaluado.*

En definitiva, la *Validez es un concepto multidimensional*, que ha sido referido en diversos momentos y situaciones haciendo énfasis en diferentes acepciones. Aplicado —y aplicable— a cualquier estrategia de acercamiento a la realidad (desde la investigación básica, los instrumentos de recogida de información o la investigación aplicada y el diagnóstico), es el criterio sustantivo desde el que valorar la calidad de cualquiera de estos procesos. Desde la posición que vamos desarrollando, la Validez alude entonces al *Criterio de Calidad de todo el proceso, y como tal es relativo* y por lo tanto gradual. No se puede hablar de la Validez como una característica que se dé o que esté ausente, sino que, en todo caso, intrínsecamente es cualitativo y globalizador, y está necesariamente vinculado a las características del objeto de evaluación, la finalidad y uso de los resultados, etc. (Swanborn, 1996; Perales 2000). Por ello, vamos a intentar aproximarnos a él, analizando diversas acepciones que le serían aplicables. Como medio de aproxima-

ción, y con el fin de poder establecer un orden en el análisis y exposición del concepto utilizamos una estrategia basada en el Análisis de Facetas de Guttman. Así, distinguimos en el concepto *dos tipos de Facetas*:

- las que responden a los componentes clásicos metodológicos de interpretación de la Validez, como referencias conceptuales desde las que abordar la reformulación del concepto, y a las que denominamos *Facetas subyacentes metodológicas de la Validez*,
- las que se pueden identificar en los componentes específicos de un proceso de Evaluación de Programas, que denominamos *Facetas adaptativas de la Validez*.

De acuerdo con lo expuesto anteriormente, pasamos a comentar las siguientes *Facetas de la Validez*:

a) *Facetas subyacentes metodológicas de la Validez*, nos referimos a aquellas que se pueden entender como componentes implicados en cualquier proceso de investigación, y que se especificarán en cuanto a su importancia o relevancia, atendiendo a las características concretas de dicho proceso (definidas por el objeto a investigar, la finalidad, etc...). Entre ellas, se pueden identificar los siguientes tópicos (Rutman, 1984):

- *Facetas genéricas*: son aquéllas que se pueden identificar con mayor o menor énfasis, dependiendo del tipo de situación (Programa, Plan de Evaluación, etc...).
 1. *Validez Interna*: se trata de demostrar que la asociación constatada entre el Programa implementado y el efecto medido se debe a una relación causal. Desde esta posición, la entendemos como *conjunto de criterios de calidad que se orientan por los principios de consistencia y credibilidad de los diferentes elementos que componen el proceso de evaluación*.
 2. *Validez Externa*: se trata de poder generalizar a constructos de mayor nivel teórico o amplitud partiendo de tratamientos y medidas específicas. Entendida como conjunto de criterios de calidad. Integra el *conjunto de criterios de calidad que se orientan por el principio de generalidad*: conseguir una adecuada pluralidad en los diferentes elementos que componen el proceso de evaluación *consistente* con la definición asumida en cada caso.
- *Facetas específicas*: son aquéllas que se pueden identificar con mayor o menor énfasis, dependiendo del tipo de situación (Programa, Plan de Evaluación, etc...).
 3. *Validez de Constructo*: se trata de realizar inferencias acerca de las construcciones teóricas que han estado orientando el planteamiento de la evaluación (por ejemplo, teorías organizacionales, en el caso de evaluaciones institucionales; o teorías instruccionales, en casos de programas de instrucción...). Respetando la ausencia de teorías formalmente construídas y adecuándose al nivel de desarrollo teórico existente en cada caso. En este sentido, hay que señalar que en el ámbito educativo, los procesos de evaluación de programas

se enfrentan a situaciones muy diversas en las que se pueden integrar, a su vez, variables e indicadores de muy diversa construcción teórica. En este caso, se trata de adecuar el proceso metodológico atendiendo al grado de desarrollo existente para cada caso, tratando cada situación en virtud del desarrollo teórico existente. Efectivamente, son muchas las situaciones en las que no existe ni ya una teoría consolidada que pueda orientar en el proceso, sino la debilidad de la mayoría de las existentes, o la ausencia de teorías para gran cantidad de situaciones. No obstante, la Evaluación de Programas, debe ofrecer respuestas en cualquier situación, si bien no debe forzar el desarrollo de sus procesos en virtud de requerimientos metodológicos que se han podido desarrollar en el marco de usos guiados por teorías. Así, el planteamiento metodológico deberá adaptarse a las ausencias o debilidades teóricas, radicando en este respeto a la realidad de la representación, precisamente, la validez de constructo del proceso evaluativo. Finalmente, hay que destacar el potencial de la Evaluación para aportar conocimiento —y, por tanto, para facilitar información para la construcción teórica— (De Miguel, 1999).

4. *Validez de Conclusión*: se trata de realizar inferencias válidas acerca de si un Programa y su posible efecto están relacionados, aún sin saber si esa relación es o no causal. Las Conclusiones, en este caso, deben identificarse en relación con los fines declarados en la Evaluación. Así, deben cumplir necesariamente los principios de consistencia y generalidad, atendiendo a la *relevancia y utilidad* de la información derivada de la evaluación.

No obstante, como señalamos anteriormente, estas facetas constituyen referencias globales en cuanto al planteamiento metodológico para cualquier tipo de Evaluación de Programas. La importancia diferencial de cada una de ellas, dependerá, como es obvio, del tipo de Programa a Evaluar, así como de los demás factores implicados, fundamentalmente la finalidad de la evaluación y su contexto, y condiciones de desarrollo. En este sentido, únicamente señalar algún ejemplo:

- Durante el Diseño de un Programa de Formación se realiza una evaluación para comprobar su eficacia, eficiencia y funcionalidad²; el objetivo es identificar los elementos a mejorar y, en definitiva, optimizar el Programa. En este supuesto, pueden concurrir todas las facetas descritas.
- Por otra parte, podemos acercarnos a otro supuesto: la evaluación del grado en que un Programa de Formación Ocupacional, orientado a la inserción laboral de las personas que asisten a él, logra sus objetivos. En este caso, la finalidad del Programa de Formación (inserción laboral), así como las condiciones de control en que se desarrolla, sitúan el énfasis en otros aspectos de la Validez, como la calidad en la representación del logro de los objetivos que se perseguían en su puesta en marcha, o la utilidad de la información que de él se deriva para aquellos que deben tomar decisiones acerca del Programa. El impacto del Programa,

2 En el sentido definido por De la Orden (1990).

por tanto, sitúa el interés de la Evaluación en la esfera de sus consecuencias, más que en la de sus procesos. En esta línea, la evaluación de este tipo de Programas puede orientarse primordialmente hacia el impacto del programa, por lo que la Validez de la Evaluación, también se dirigirá hacia este aspecto (Mohr, 95; Apodaka, 99). Lo más importante en este caso será que la Evaluación pueda ofrecer información adecuada acerca de la inserción profesional de los asistentes al Programa. La amplitud del concepto que reflejen los indicadores de inserción, así como los elementos explicativos de la misma, constituirán elementos clave para la validación del plan de evaluación. Así, por ejemplo, no se podrá considerar igualmente válido un Programa que se limite a controlar cuántas personas han encontrado trabajo después de realizar el Programa de Formación (en diversos momentos: después de un mes, tres meses o un año...), que otro que además considere otros elementos, como la adecuación del trabajo conseguido en función de la formación recibida, la covariación posible de la inserción con otras variables (personales —edad, sexo, curriculum...—, sectoriales y/o sociales —tejido empresarial de la localidad en que se encuentra trabajo, elementos cíclicos de la economía, ...—, etc...), o el análisis de los beneficios personales obtenidos a partir del Programa (con independencia de la mera inserción).

Estas matizaciones nos llevan a considerar otras facetas de la Validez de la Evaluación de Programas, que cobran especial énfasis en virtud de las características de la Evaluación de Programas como ámbito aplicado de investigación. A ellas nos referimos en el apartado siguiente.

b) *Facetas adaptativas de la Validez*, porque se implican en tanto en cuanto el proceso de Evaluación de Programas debe atender a todos los componentes necesarios para conseguir sus fines (mejora, innovación, cambio...). Se pueden identificar en los siguientes grandes núcleos procesuales:

- *En relación con la Representación del Objeto*. Se trata del grado en que, en la Evaluación, se seleccionan indicadores y criterios que se ajusten a aquello que se desea evaluar. La identificación de instrumentos, fuentes de información, tipo de información que se va a facilitar, utilización de la misma, etc..., forman parte de la misma.

En este sentido, hay que señalar que en la Evaluación de Programas, en múltiples ocasiones, se producen disfunciones entre los planteamientos que animan los procesos de evaluación y su traducción u operativización en los indicadores que los sustentan. Así, revisemos únicamente algún ejemplo: los procesos de mejora de la calidad aluden a este concepto como la referencia que orienta la evaluación, así como la finalidad de la misma (mejora de la calidad). No obstante, las definiciones acerca de la Calidad de cualquier programa educativo suelen ser muy genéricas. Recuerdan en alguna medida el rol que juegan las definiciones de Rasgos Psicológicos como orientadoras de los tests psico-

métricos³. De esta forma bajo una misma definición de Calidad de un Programa se pueden identificar evaluaciones muy diversas, orientadas hacia el Programa, o hacia su efecto, o únicamente sobre algunos elementos procesuales⁴ (Dendaluze, 1991).

En la Evaluación de Programas, la *representación del objeto, necesariamente, nace del compromiso entre todos los elementos de definición de la evaluación*. No sólo se trata de definir qué se evalúa, sino que hay que entender que ello es necesariamente interdependiente con todos los restantes componentes de la evaluación: para qué, cómo, a quién y de qué modo se pide información, qué información se va a aportar, qué usos se harán con ella, etc... Este compromiso global de todos los componentes en la definición del objeto a evaluar persigue la idea de *Coherencia* como base para aspirar a la Validez en cualquier proceso de evaluación. Por ello, es un aspecto central en la Planificación y diseño de la Evaluación.

Implica por lo tanto, todas las facetas que se relacionan con el planteamiento metodológico e instrumental: establecimiento del diseño, determinación de los procesos de la evaluación, selección de indicadores e instrumentos, tipos de análisis de la información, usos de la misma, etc...

- *Adaptación al proceso evaluado*. Se trata de llevar a cabo los diferentes elementos del proceso de la Evaluación, respetando el proceso natural de desarrollo del Programa Educativo. En esta línea, hay que señalar los elementos que coadyuvan a la consecución de la Validez, se pueden sintetizar en el *Respeto* a los procesos que se evalúan, no interfiriendo en los mismos. La introducción de expectativas por el hecho de la realización de la evaluación, los condicionantes que pueda suponer su realización para el proceso o para los diferentes actores implicados en el mismo, entre otros aspectos, pueden introducir elementos que interfieran en el normal desenvolvimiento de lo que se está evaluando.
- *Integración Contextual*. Otro elemento sustancial de la Validez es la consideración de las diferentes variables que pueden afectar al Programa evaluado y, por ello, deben ser atendidas como elementos de contextualización. Estos elementos servirán de base para la extensión de la Comprensión, y en su caso, Explicación de los resultados. Por ello, son fundamentales como componentes de modulación de la interpretación y comprensión de la situación evaluada.
- *Participación e implicación de audiencias*. Desde la posición epistemológica y metodológica en que nos situamos, es necesario considerar que la consecución de los fines de la evaluación sólo son posibles si se atiende adecuadamente la Partici-

3 Recuérdese, por ejemplo, la siguiente definición: el test ... está destinado a valorar la capacidad para conceptualizar y aplicar el razonamiento sistemático a nuevos problemas y apreciar las funciones centrales de la inteligencia: abstracción y comprensión de relaciones. Según los autores del mismo es el objetivo del Test D-48. La excesiva distancia entre la definición del objeto a evaluar (Rasgo: Inteligencia), y la operativización de su medida es un claro indicador de falta de Validez (en este caso, hablaríamos de carencias en cuanto a su Validez de Constructo y de Contenido, como instrumento de medida).

4 En este mismo número monográfico hacemos, junto a la profesora M^a J. Perales, una revisión de los Planes de Evaluación de la Formación Ocupacional y Continua. En él se incluye una referencia concreta a la consideración del concepto de Calidad en estas Evaluaciones. Puede ilustrar los extremos que en este punto comentamos.

pación de los diferentes actores implicados en el Programa (De Miguel, 1995). La Validez de la Evaluación, desde esta perspectiva, dependerá del grado en que se haya ajustado la participación de todos ellos, de acuerdo con su nivel de implicación y expectativas. Estos aspectos se conjugan en diferentes momentos del proceso, desde la participación en la definición del objeto a evaluar y sus finalidades, pasando por el tipo de información que se requiere y cómo se recoge, hasta la información que se hace disponible y cómo se difunda. Lo cierto es que impregna todo el proceso de evaluación y, por ello, debe ser adecuadamente atendida en toda su Planificación y Desarrollo.

- *Compromiso con el Cambio: Credibilidad y Utilidad.* Finalmente, esta faceta es probablemente la más determinante en cuanto a las consecuencias de la Evaluación. La Evaluación de Programas, desde la posición de que partimos, sólo es posible considerarla válida si aporta los elementos de mejora e innovación que se pretendía identificar y que justificaban la puesta en marcha del proceso. Estas facetas, ya descritas por otros autores (Tejedor, 2000; Mateo, 1990), son a su vez consecuencia de haber atendido adecuadamente todas las anteriormente descritas, y base para la valoración global del proceso. En definitiva, procesos evaluativos metodológicamente bien contruados y desarrollados, no son asumidos como válidos, por la falta de credibilidad que dimana de alguno de los componentes anteriores que no se han atendido bien, o por la falta de utilidad de la Evaluación.

En cualquier caso, esta faceta no depende de los momentos finales del proceso evaluativo, sino que empieza a construirse en el momento en que se comienza con la evaluación, adaptando de forma coherente todo el proceso y atendiendo todas las facetas señaladas.

2. ESTRATEGIAS DE CONSECUCIÓN DE LA VALIDEZ EN EVALUACIÓN DE PROGRAMAS

En los procesos de investigación y evaluación enmarcables dentro de posiciones postpositivistas, la Validez —tanto Interna como Externa— se vincula a diversos factores, favorables a su consecución, o bien que atentan contra ella. Desde la posición que revisamos, aunque constituye una reinterpretación conceptual, obviamente también los hay. Sin embargo, dadas las características de esta presentación, no vamos a realizar un análisis exhaustivo de los mismos, sino que vamos a acercarnos a su delimitación tomando como base las diversas etapas de desarrollo en que podría enmarcarse un proceso de Evaluación de Programas⁵. Para ello, integramos las acciones que se implican

5 Identificar *Etapas* alude a la referencia de una secuencia lógica que nos permita centrar el análisis. Entendemos que en estos Procesos debe adoptarse una actitud dinámica que los convierta en un esquema de trabajo flexible. Ello no quiere decir que queden fuera de este planteamiento situaciones como los diseños emergentes. Es obvio que, en ellos, no sería posible identificar estas grandes categorías de funcionamiento, tal cual lo hacemos aquí. Sin embargo, la elección del esquema de reflexión se basa en la pretensión de abordar el tema desde una perspectiva general, aunque produzca problemas en cuanto a su aplicación a tipos de aproximaciones más específicas. El fundamento del enfoque, no obstante, es el mismo.

en los procesos de evaluación en dos grandes conjuntos: a) la Planificación de la Evaluación, y b) la Gestión de la Información. Pasamos a comentarlas a continuación.

- Planificación de la Evaluación: Construcción del plan de trabajo, adaptación y desarrollo del mismo.

Por *Planificación* entendemos el conjunto de actividades que van anticipando y haciendo posible todos los elementos a tener en cuenta en el proceso de evaluación (Jornet, Suárez y Belloch, 1998). Incluirá desde los elementos metodológicos clásicos (decisiones en cuanto al diseño de la evaluación, elementos de operativización, etc...), junto a todos los componentes psicosociales implicados en el desarrollo de una evaluación (identificación de audiencias, promoción de la aceptación de la evaluación, establecimiento de consensos en cuanto a todos los componentes del proceso, etc...).

En cualquier caso, debe tenerse en cuenta que hablar de Planificación no implica necesariamente referirse a procesos en que es posible anticipar todos los elementos a tener en cuenta. Por ello, si bien se puede entender como una acción que comienza con anterioridad a la puesta en marcha de cualquier evaluación, debe abordarse como un componente de flexibilización, que debe presidir todo el desarrollo del proceso, facilitando las modificaciones y adaptaciones del mismo según sea necesario.

Por ello, los elementos que pueden coadyuvar en esta actividad a la consecución de la Validez son muy diversos. Entre ellos los más relevantes pueden ser:

- *Participantes, roles y responsabilidades.*

Probablemente uno de los elementos clave de cualquier proceso de evaluación es la integración de los diversos implicados de acuerdo con su grado de responsabilidad y rol. Obviamente, desde la definición de lo que va a ser objeto de evaluación, pasando por la finalidad de la misma y uso de la información, etc...la Evaluación de Programas debe estar enmarcada en un proceso claro de participación, facilitando el desarrollo de evaluaciones en los que todos los elementos estén consensuados. Es claro que en este tipo de procesos nos encontramos con situaciones muy diversas, desde evaluaciones orientadas a la rendición de cuentas en las que se sustenta el proceso sobre evaluadores externos hasta otras netamente orientadas a la mejora e innovación basadas en la autoevaluación. En cualquier caso, en todas ellas, habrá que analizar los grados de participación de todos los colectivos implicados y favorecer su participación, promoviendo estrategias de evaluación basadas en la *aceptación y credibilidad*.

En este marco, son importantes las estrategias basadas en la creación de comités que integren a representantes de los diversos colectivos implicados, determinar el rol de los evaluadores en los mismos. Estas estrategias pueden constituir la base para la consecución del consenso en los diferentes momentos por los que transita un proceso de evaluación.

- *Definición del Objeto a Evaluar y Finalidad de la Evaluación*

El objetivo inicial es delimitar el objeto y finalidades de la Evaluación en términos de *validez/credibilidad*, tanto *interna* —la consistencia del planteamiento del problema respecto a la situación específica que se pretende abordar y a las otras respuestas que se han dado en situaciones similares—, como *externa* —credibilidad para los agentes y audiencias implicadas en los mismos—.

En este sentido, las estrategias básicas de acercamiento a la *Validez*, requieren, junto a un planteamiento metodológico riguroso y ajustado, de la *participación de todos los implicados*. Se trata de valorar y adecuar la orientación de la evaluación, identificar los obstáculos que puedan producirse para llevar a cabo el proceso (desde su realización a sus consecuencias), promover la aceptación de la evaluación y anticipar los límites de utilización de la Evaluación. Para ello, la *negociación* constituye la base más adecuada para desarrollar este proceso.

Los componentes básicos en este caso son:

- Determinar qué es lo que se va a evaluar a nivel operativo, y
- Definir los usos finales de la evaluación.

En este caso, un elemento que se va introduciendo paulatinamente es el de *Contrato de Evaluación*. Una fuente frecuente de invalidez de los procesos de Evaluación de Programas reside en los cambios que se van introduciendo a lo largo del proceso de evaluación, que pueden llevar a conllevar utilidades no previstas de la misma. Asimismo, en muchos casos hemos observado como la institucionalización de un proceso en lugar de constituir un elemento de mejora del mismo, conlleva perjuicios en los usos de los sistemas de evaluación; la presencia de cambios sucesivos en los gestores de la institución usuaria de la evaluación, facilita este tipo de cambios⁶.

- *Construir el plan de trabajo*

Se trata de traducir en términos de actuación el problema que se ha planteado. Esta traducción que guía todo el plan de trabajo debe reunir condiciones de *fidelidad y generalidad*. Así, por una parte, es preciso mantener una conexión permanente con las ideas clave o características fundamentales de la definición del

6 Obviamente, además de un problema de validez para las Evaluaciones de Programas, constituye un problema ético para los evaluadores. Lo cierto es que en este contexto, es claro que si un sistema se desarrolla para una institución, atendiendo a un uso determinado, si bien el sistema es de la Institución, los implicados en su desarrollo deberían estar amparados en algún elemento contractual que, cuanto menos, asegurara los usos previstos, de forma que en caso de que la institución deseara introducir variaciones, lo debiera hacer a través de alguna adaptación en la que la responsabilidad quedara circunscrita a quienes lo promuevan y utilicen. Este tipo de situaciones, que conlleva la progresiva desvirtuación de procesos que se diseñaron para propósitos diferentes, por desgracia es habitual y va más allá de la implicación de la validez. Este es un elemento que se ha desarrollado en aquellos contextos en los que la Evaluación presenta un mayor nivel de consolidación social y profesional, particularmente en los EE.UU. de Norteamérica.

objeto y finalidad de la evaluación, realizando una traducción *fidedigna* de las mismas —de forma que, por ejemplo, estén todas las dimensiones que aparecen en el problema y que se seleccione una información relevante para las mismas—. Y por otra parte, la *generalidad* de la traducción implica atender a la validez externa del plan de trabajo. Es decir, supone un planteamiento de las claves del mismo desde la perspectiva no sólo de la situación concreta que se aborda, sino que es conveniente atender a las soluciones observadas en situaciones de la clase a la que pertenezcan.

En la Evaluación de Programas es frecuente que se produzca una tensión entre ambos principios, *fidelidad y generalidad*, más agudizada que en otros ámbitos de trabajo. En este caso, el proceso se mueve por una demanda concreta de un «cliente» sobre una situación específica, así como otras audiencias. Ello conlleva que en ocasiones se mantenga la fidelidad en detrimento de la generalidad. No obstante, la clave de la validez en este caso es la fidelidad, si bien la generalidad puede constituir un componente que, de poderse dar de forma concomitante, facilitaría la explicación e interpretación de resultados al disponer de elementos de comparabilidad y contextualización externos a la propia evaluación.

En este caso, la fidelidad de la traducción del problema debe asegurarse en función de los siguientes componentes:

- Determinación de la estructura del diseño: determinación de los roles de las variables e indicadores sobre las que se sustentan, momentos de recogida de información, evaluabilidad de las informaciones, etc...
 - Selección de indicadores adecuados,
 - Identificación de la mejor fuente de información, en cada caso, para cada indicador.
- *Recogida de Información.*

Se trata de realizar la recogida de información atendiendo a las condiciones en que pueda asegurarse la máxima *calidad y representatividad* de la misma, en virtud de los objetivos y finalidades de la evaluación, no produciendo para ello disrupciones en la dinámica habitual del Programa. Así, en esta etapa existen diversas condiciones a tener en cuenta:

- Conseguir una relación adecuada entre la *fuentes* que nos va a dar la información y el *procedimiento* que vamos a utilizar para recogerla.
- Seleccionar los instrumentos y técnicas de recogida de información de acuerdo a criterios de bondad de los mismos (fiabilidad y validez).
- Recoger la información en el momento más adecuado considerando tanto la no intromisión en el desarrollo del Programa, como la calidad derivada del mismo (por ejemplo, evaluando el impacto de un programa de inserción, momentos de recogida de información posteriores a la realización del Programa: a corto, medio y largo plazo...).

— Utilizar, en lo posible, estrategias que permitan el contraste y validación de la información, a través de fuentes, procedimientos y/o instrumentos múltiples (por ejemplo, triangulación, validación convergente...).

- *Análisis de Datos, síntesis de la información y su preparación para la Toma de Decisiones.*

Las dimensiones básicas a atender como criterios de calidad del proceso serían garantizar la *consistencia* de la síntesis de información con el planteamiento del problema y la *credibilidad* de la misma para las audiencias implicadas. En esta etapa, los componentes básicos son los siguientes:

- La *coherencia interna* de la evaluación, se basa en el desarrollo de un proceso de análisis de datos ajustado al planteamiento del problema y uso posterior de la información derivada de la evaluación, atendiendo a los diversos roles de las variables implicadas, con el fin de ampliar al máximo los elementos de comprensión y explicación de resultados.
- La *coherencia externa*, alude a la *credibilidad* de los resultados para los diferentes actores implicados. Esto supone valorar la consistencia externa para estas audiencias de la síntesis de información que se ha llevado a término. En realidad este proceso se sitúa entre esta Etapa de actuación y la siguiente, dado que una parte del mismo se lleva a cabo dentro del Proceso de Difusión de la información.

- *Gestión de la información.* Constituye el segundo núcleo de acciones desde el que abordar el problema de los criterios de calidad del proceso de evaluación. En él diferenciamos tres grandes momentos: durante el proceso de evaluación, en su finalización y después de haberse realizado la evaluación.

- *Durante el proceso de evaluación.*

Una característica básica que es conveniente tener en cuenta a lo largo de cualquier proceso de Evaluación de Programas es la flexibilidad para adaptar el mismo en virtud de las informaciones que vayan recabándose y analizándose. Es decir, la planificación se retroalimenta a través del ajuste sucesivo de todos los componentes del Plan de Evaluación durante su implementación. Esta adaptabilidad del Plan está al servicio de la fidelidad y la generalidad del mismo, así como coadyuva en el logro de la credibilidad y utilidad de la evaluación.

- *En la finalización del proceso de evaluación: Difusión de Información.*

En la Etapa de la Difusión de la información, el objetivo central es ofrecer información —al cliente y a las demás audiencias implicadas—, sobre cuales son los resultados y los criterios básicos en que puede sustentarse la interpretación de los mismos y la toma de decisiones. En este caso, las dimensiones vuelven a ser la *coherencia* en la difusión de la información y la *credibilidad* de los medios utilizados para los agentes implicados.

El criterio de *coherencia interna* supone atender a la conexión entre los Informes y lo previamente establecido de manera consensuada en las etapas anteriores del proceso. Así, será preciso aportar la información que se haya pactado al inicio del proceso con cada audiencia, atendiendo a su grado de responsabilidad e implicación en el Programa y en la Evaluación, sus intereses y posibles usos de la información. Además se deben atender condiciones de calidad en los mecanismos de difusión de la Información: utilización de un lenguaje ajustado a la audiencia a la que se dirige, estructuración de la información adecuada a las preguntas que se pretende responder, atención a la oportunidad temporal para los objetivos básicos propuestos (Tejedor, 1997; Elliott, Fischer, y Rennie, 1999). Estos mismos aspectos deberán considerarse desde la perspectiva de la *coherencia externa* para garantizar la credibilidad de los mecanismos de la difusión de la Información a las diferentes audiencias implicadas.

- *Después del proceso: valoración global del mismo y seguimiento de las consecuencias de la evaluación.*

Un elemento adicional, pero de gran importancia, lo constituyen las estrategias metaevaluativas (De la Orden y Martínez de Toda, 1999) y las de seguimiento de los procesos de implementación de las soluciones aportadas a través de la evaluación (Rodríguez Espinar, 1997, 1998). El análisis global del proceso y la comprobación de los compromisos mantenidos respecto de las consecuencias de la evaluación pueden constituir elementos clave de validación global del Plan de Evaluación.

A MODO DE CONCLUSIÓN

En primer lugar, es preciso recordar que la *Evaluación de Programas* es algo más que una aplicación de métodos a través de los que se enumeran resultados. Hay que ser conscientes de que actualmente los investigadores no son los únicos que deciden evaluar y qué métodos utilizar. Ellos intervienen en un sistema de decisiones complejo en el que también participan otros colectivos implicados en el problema objeto de evaluación, como legisladores, administradores, planificadores, profesores y alumnos, etc... (De la Orden y Mafokozi, 1999). Por ello, es necesaria una profunda reflexión en cuanto a las soluciones metodológicas y sus componentes, en un ámbito en el que la complementariedad no puede asumirse como una mera adición de aportaciones que provienen desde diversas opciones epistemológicas.

Asimismo, la atención prioritaria que debe prestarse a las características del objeto evaluado, las finalidades de la evaluación, sus usos, etc..., como elementos que deben conducir en la selección de la solución metodológica desde la que abordar el problema, debe acompañarse de una reflexión respecto al rol que desarrollamos los que, desde la Evaluación, participamos en estos procesos. Los valores, la ideología, la forma de estructurar cognitivamente la información —producto de nuestra formación metodológica y de nuestra forma de ser— inciden en los planteamientos que realiza-

mos (House, 1994). Por ello, es conveniente asumir como punto de partida un acercamiento abierto y tolerante, que no es patrimonio exclusivo de ninguna posición ideológica o epistemológica más en la práctica que en el discurso teórico en que se basa. La formación y profesionalización en el ámbito de la Evaluación de Programas puede constituir un instrumento desde el que acompañar este proceso de mejora de este tipo de procesos (Scriven, 1994; González Such, Pérez Carbonell y Perales, 2000). No obstante, el camino está por recorrer, si bien se identifican líneas de trabajo prometedoras que pueden abocar en soluciones ciertamente alternativas que superen las dificultades existentes.

BIBLIOGRAFÍA

- Aliaga Abad, F. (2000). *Bases epistemológicas y Proceso de Investigación psicoeducativa*. Valencia: C.S.V.
- Alvira, F. (2ª Ed.) (1996). *Metodología de la Evaluación de Programas*. C.I.S. Madrid: Cuadernos Metodológicos, 2.
- Apodaka, P. (1999). Evaluación de los resultados y del impacto. *Revista de Investigación Educativa*, 17, 2, págs. 363-377.
- Bericat, E. (1998). *La integración de los métodos cuantitativo y cualitativo en la investigación social. Significado y medida*. Barcelona: Ariel Sociología.
- Buendía, L. (1998). La Investigación sobre Evaluación Educativa. *Revista de Investigación Educativa*, 14, 2, págs. 5-24.
- Buendía, L. (2000). Ética de la Evaluación. Ponencia presentada en las I Jornadas de Medición y Evaluación Educativas. Valencia: en prensa.
- Chelmsky, E. y Shadish, W.R. (1997). *Evaluation for the 21st Century*. Sage Publications.
- Colás, M.P. (1995). Diseño y evaluación de programas. En J.L. Rodríguez Diéguez y O. Sáenz Barrio, *Tecnología educativa. Nuevas tecnologías aplicadas a la educación* (pp. 45-66). Alcoy: Marfil.
- Cook, T.D. y Reichardt, Ch.S. (1986). Hacia una superación del enfrentamiento entre los métodos cuantitativos y cualitativos. En: T.D. Cook y CH. S. Reichardt (1986). *Métodos Cualitativos y Cuantitativos en Investigación Evaluativa*. (3ª Ed.) Madrid: Morata.
- Cook, T.D. y Shadish, W.R. (1994) Social experiments: some developments over the past fifteen years. En *Annual Review of Psychology*, 45, págs. 545-580.
- Creswell, J.W. (1994). *Research Design. Qualitative & Quantitative Approaches*. Thousand Oaks. CA. Sage Publications.
- Creswell, J.W. (1998). *Qualitative inquiry and research design. Choosing among five traditions*. Thousand Oaks, CA: Sage.
- De la Orden, A. (1990). Evaluación de los Efectos de los Programas de Intervención. *Revista de Investigación Educativa*, 8, 16, págs. 61-76.
- De la Orden, A. y Mafokozi, J. (1999). La investigación educativa: naturaleza, funciones y ambigüedad de sus relaciones con la práctica y la política educativas. *Revista de Investigación Educativa*, 17, 1, págs. 7-29.
- De la Orden, A. y Martínez de Toda, M.J. (1991) Metaevaluación educativa. *Bordón*, 43, 4, págs. 7-29.

- De Miguel, M. (1988). Paradigmas de la investigación educativa española. En I. Dendaluze (Coord.), *Aspectos metodológicos de la investigación educativa* (págs. 60-77). Madrid: Narcea.
- De Miguel, M. (1995). Participación ciudadana en la evaluación de las políticas públicas. *Aula Abierta*, 66, págs. 211-218.
- De Miguel, M. (1999). La evaluación de programas: entre el conocimiento y el compromiso. *Revista de Investigación Educativa*, 17, 2, págs. 345-348.
- Dendaluze, I. (1991). *Dimensiones e investigabilidad de la Calidad de la Educación*. En P. Braza y otros. (Eds.). Actas del I Congreso Internacional sobre Calidad de la Educación Universitaria. Cádiz: Universidad de Cádiz.
- Dendaluze, I. (1995). Avances en los métodos de investigación educativa en la intervención psicopedagógica. *Revista de Investigación Educativa*, 26, págs. 9-32.
- Dendaluze, I. (1997). La validez teórica de las investigaciones empíricas en Ciencias Sociales. *Revista internacional de estudios vascos*, 42, 1, págs. 77-101.
- Dendaluze, I. (1998). Algunos retos metodológicos. *Revista de investigación educativa*, 16, 1, págs. 7-24.
- Elliott, R., Fischer, C.T., y Rennie, D.L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology*, 38, págs. 215-229.
- González Such, J., Pérez Carbonell, A. y Perales Montolío, J. (2000). Perfil profesional del evaluador. Ponencia presentada en las I Jornadas de Medición y Evaluación Educativas. Pendiente de publicación. Valencia.
- House, E.R. (1980). *Evaluating with validity*. Beverly Hills, CA: Sage.
- House, E.R. (1992). Tendencias en Evaluación. *Revista de Educación*, 299, págs. 43-55.
- House, E.R. (1994). *Evaluación, ética y poder*. Madrid: Morata.
- Jornet, J.M., Suárez, J.M., y Belloch, C. (1998). Metodología de Evaluación de Programas de Formación Profesional, Ocupacional y Continua. *Leonardo Project: Eurodialogue Evaluation*.
- Krathwohl, D.R., (1987). *Social and behavioral science research: a new framework for conceptualizing implementing and evaluating research studies*. San Francisco: Jossey Bass.
- Mateo, J. (1990). La toma de decisiones en el contexto de la evaluación educativa. *Revista de Investigación Educativa*, 16, 8, págs. 95-114.
- Mohr, B. (2ª Ed.) (1995). *Impact Analysis for Program Evaluation*. Newbury Park, Ca: Sage Publication.
- Morse, J.M. (1997). Considering theory derived from qualitative research. En J.M. Morse (Ed.), *Completing a qualitative project. Details and dialogue* (págs. 163-189). California: Sage.
- Perales, M.J. (2000). *Enfoques de Evaluación de la Formación Ocupacional y Continua. Estudio de Validación de un modelo*. Tesis Doctoral. Universitat de València. (en prensa, Microfichas).
- Pérez Carbonell, A. (1998). *Metodología de la Evaluación de Programas: Evaluación de programas de postgrado en la Universitat de Valencia*. Col·lecció: Tesis Doctorals en microfiches. Universitat de Valencia: SERVEI DE PUBLICACIONS ETD, S.A. (publicación: 1999)

- Pérez Juste, R. (2000). Evaluación de programas educativos. *Ponencia presentada en las I Jornadas de Medición y Evaluación Educativas*. Valencia, Marzo (en prensa).
- Reichardt, C. y Rallis, S. (Eds.) (1994). *The qualitative-quantitative debate: new perspectives*. S. Francisco: Jossey Bass.
- Rodríguez Espinar, S. (1997). La evaluación institucional universitaria. *Revista de Investigación Educativa*, 15, 2, págs. 179-214.
- Rodríguez Espinar, S. (1998). El proceso de evaluación institucional. *Revista de Educación*, 315, págs. 45-65.
- Rodríguez, G., Gil, J., y García, E. (1996). *Metodología de la investigación cualitativa*. Archidona (Málaga): Aljibe.
- Rogers, P. y Owen, J. (1999). *Program evaluation. Forms and approaches*. Sage.
- Rutman, L. (1984). *Evaluation Research Methods: A basic Guide*. Inc. Beverly Hills, California: Sage Publications.
- Scriven, M.S. (1994). Evaluation as a discipline. *Studies in Educational Evaluation*, 20, págs. 147-166.
- Suárez, J.M. (1989). *Proyecto Docente de Investigación*. Proyecto inédito, Universitat de València.
- Swanborn, P.G. (1996). A common base for quality control criteria in quantitative and qualitative research. *Quality and Quantity*, 30, págs. 19-35.
- Tejedor, F.J. (1997). Informes en la Evaluación Institucional. *Revista Española de Pedagogía*, 208, págs. 471-498.
- Tejedor, F.J., García-Valcarcel, A. y Rodríguez, M.J. (1994). Perspectivas Metodológicas Actuales de la Evaluación de Programas en el Ámbito Educativo. *Revista de Investigación Educativa*, 23, págs. 93-127.