



EVALUACIÓN DE PROGRAMAS Y DE CENTROS EDUCATIVOS. DIEZ AÑOS DE INVESTIGACIÓN.

Consuelo Vélaz de Medrano Ureta
Ángeles Blanco Blanco
Amalia Segalerva Cazorla
M.ª Esther del Moral Pérez

Ministerio de Educación y Ciencia

CIDE

**EVALUACIÓN DE PROGRAMAS Y
DE CENTROS EDUCATIVOS,
DIEZ AÑOS DE INVESTIGACIÓN**

**CONSUELO VÉLAZ DE MEDRANO URETA
ÁNGELES BLANCO BLANCO
AMALIA SEGALERVA CAZORLA
M.^a ESTHER DEL MORAL PÉREZ**

Número: 115

Colección: INVESTIGACIÓN

1. Evaluación.—2. Investigación Educativa.—3. Programa de Formación.—
4. Centro de Enseñanza.



© MINISTERIO DE EDUCACIÓN Y CIENCIA
Secretaría de Estado de Educación
Dirección General de Renovación Pedagógica
Centro de Investigación y Documentación Educativa
EDITA: CENTRO DE PUBLICACIONES—Secretaría General Técnica
Tirada: 1.200 ejs.
NIPO: 176-95-151-3
I.S.B.N.: 84-369-2741-9
Depósito legal: M-37.606-1995
Imprime: Solana e Hijos Artes Gráficas, S.A.
C/ San Alfonso, 26, La Fortuna (Madrid)

ÍNDICE

| | |
|--------------------|----|
| PRESENTACIÓN | 9 |
| INTRODUCCIÓN | 13 |

CAPÍTULO I

| | |
|---|----|
| CLARIFICACIÓN CONCEPTUAL EN TORNO A LA EVALUACIÓN EDUCATIVA | 17 |
| 1.1. La evaluación desde una perspectiva histórica | 17 |
| 1.2. Delimitación de conceptos: medida, evaluación, investigación | 23 |

CAPÍTULO II

| | |
|---|----|
| LA EVALUACIÓN DE PROGRAMAS Y CENTROS EN ESPAÑA EN EL PERÍODO 1982-1992 | 29 |
| 2.1. La evaluación y los períodos de reforma educativa .. | 29 |
| 2.2. Foros nacionales de discusión en torno a la evaluación de programas e instituciones educativas | 31 |

CAPÍTULO III

| | |
|--|----|
| LAS EVALUACIONES DE PROGRAMAS Y CENTROS EN EL CONJUNTO DE LA INVESTIGACIÓN FINANCIADA POR EL C.I.D.E | 35 |
| 3.1. Áreas prioritarias de investigación | 35 |
| 3.2. La evaluación de programas y centros en el contexto de la investigación educativa del Ministerio de Educación y Ciencia | 38 |

CAPÍTULO IV

| | |
|--|-----|
| REVISIÓN DE LAS EVALUACIONES DE PROGRAMAS Y CENTROS EN EL PERÍODO 1982-1992 | 41 |
| 4.1. Criterios utilizados en la revisión y análisis de las investigaciones | 41 |
| 4.2. Evaluación de programas | 49 |
| 4.2.1. Evaluación de reformas y de programas experimentales | 53 |
| 4.2.2. Evaluación de programas curriculares de aplicación restringida | 79 |
| 4.2.2.1. Evaluación de programas curriculares para distintas áreas y niveles educativos | 79 |
| 4.2.2.2. Evaluación de programas instructivos específicos | 87 |
| a) Evaluación de programas desarrollados en Educación Primaria y Secundaria: | 87 |
| a.1. Evaluación de programas instructivos centrados en los procesos didácticos .. | 87 |
| a.2. Evaluación de programas instructivos centrados en el uso de medios y recursos didácticos | 106 |
| a.3. Evaluación de programas instructivos contruidos sobre modelos teóricos definidos | 115 |
| b) Evaluación de programas desarrollados en Educación Superior | 127 |
| 4.2.3. Evaluación de Programas de Formación Inicial y Permanente del Profesorado | 133 |
| 4.2.3.1. Programas de Formación Inicial .. | 133 |

| | |
|--|-----|
| 4.2.3.2. Programas de Formación Perma- nente | 145 |
| 4.2.4. Evaluación de programas educativos para la prevención de drogodependencias en el ám- bito escolar | 164 |
| 4.3. Evaluación de centros | 174 |
| 4.4. Investigaciones centradas en la elaboración de mode- los para la evaluación de programas o centros educa- tivos | 202 |
| 4.4.1. Modelos de evaluación de programas de for- mación del profesorado | 202 |
| 4.4.2. Modelos de evaluación de centros | 214 |

CAPÍTULO V

| | |
|--|-----|
| CONSIDERACIONES FINALES: VISIÓN PANORÁMICA DE LAS INVESTIGACIONES | 229 |
| BIBLIOGRAFÍA | 247 |
| LISTADO DE INVESTIGACIONES REVISADAS | 251 |
| RELACIÓN ALFABÉTICA DE AUTORES | 257 |

PRESENTACIÓN

Hace dos años, el C.I.D.E. comenzaba a publicar una serie de trabajos con el propósito de pasar revista a la producción investigadora de la última década en algunos ámbitos educativos relevantes. Pretendía así cumplir con dos de sus misiones centrales, como son las de realizar el balance y la valoración de la investigación educativa desarrollada en España y difundir sus principales resultados. A través de los cinco volúmenes hasta ahora publicados ha presentado el estado de desarrollo de la investigación en otros tantos campos, poniendo a disposición de los investigadores, especialistas y estudiosos un conjunto de reflexiones y análisis, unos catálogos temáticos amplios y unas valiosas síntesis de resultados.

Para los que nos hemos ocupado profesionalmente de la evaluación de los sistemas educativos, es una buena noticia la aparición de este quinto volumen, dedicado precisamente a la evaluación de programas y de centros educativos. Es una buena noticia, como decía, aunque no sea una sorpresa, dada la importancia que la evaluación está cobrando tanto en el ámbito nacional como en el internacional.

En efecto, la evaluación va ocupando un lugar progresivamente más importante en el seno de los sistemas educativos actuales. No es cuestión de repetir aquí, en breves palabras, lo que ha sido ampliamente argumentado y publicado al respecto por diversos autores. Pero sí querría insistir en que nos encontramos ante un cambio profundo de tendencia en lo que se refiere a los modos de gestión y conducción de los sistemas educativos, más que ante un fenómeno de simple moda. Los mecanismos tradicionales de administración y control no acaban de adaptarse a las nuevas demandas que reciben las escuelas y al principio de autonomía que parece inspirar cada vez en mayor medida la organización educativa. La contrapartida de esa nueva situación es un mayor énfasis en los mecanismos de evaluación y de rendición de cuentas, entendidos de modo diverso en función de las

coordinadas específicas de cada sistema educativo, pero en proceso de expansión en un número creciente de países.

Así pues, no creo exagerado afirmar que la evaluación de los sistemas educativos está de actualidad y parece responder al sentido de los tiempos. Pero esta etiqueta común encubre realidades muy diversas entre sí. En algunos lugares, por evaluación del sistema educativo se entiende la puesta en marcha de pruebas nacionales de rendimiento de los alumnos, generalmente asociada a cambios curriculares de mayor o menor envergadura. En otros países, implica el desarrollo de planes sistemáticos de evaluación de centros. En ocasiones, se pone el énfasis en la construcción de sistemas de indicadores de la educación, siguiendo una tendencia de plena actualidad internacional. En ciertos casos, la evaluación intenta adoptar una perspectiva más amplia que las anteriores, abarcando el conjunto del sistema educativo. Y ello es así porque el término no tiene un referente único.

Al hablar de evaluación educativa, debemos distinguir entre ámbitos tales como los de evaluación de los aprendizajes, de los programas educativos, de los centros escolares o universitarios, de la función docente y de las políticas educativas. Cada uno de ellos tiene su objeto específico de evaluación (aun cuando a veces puedan existir ciertos solapamientos), sus propios enfoques, métodos y técnicas (aunque posean indudables semejanzas entre sí) y sus propios criterios de validez. El término de evaluación de los sistemas educativos tiende a englobar todos ellos, enfatizando el carácter integrador que debe tener un mecanismo coherente de evaluación.

Esta publicación que tiene el lector entre sus manos está concebida con cierta modestia, principal razón de su consistencia. Las autoras han sabido captar la distinción que antes mencionaba acerca de los diversos ámbitos que abarca el término. Y han apreciado, correctamente, que el grado de desarrollo de unos y otros es desigual. En vez de pretender cubrir todo el espectro de trabajos producidos en el amplio terreno de la evaluación educativa, han sabido acotarlo, centrando su atención en una parcela de los mismos. De ahí que hable de modestia.

Pero también me refería a su consistencia, y ello por tres razones. La primera, porque han sabido identificar una parcela que ha ido cobrando consideración propia dentro del campo de la evaluación. En efecto, la evaluación de programas está ya plenamente asentada, desde que hace unos treinta años comenzó una fuerte corriente en este sentido en los Estados Unidos, rápidamente expandida por otras

áreas geográficas. Y otro tanto puede decirse de la evaluación de centros, si bien en este caso se trata de una tendencia más reciente. La segunda razón consiste en que han adoptado una perspectiva amplia para acercarse a su objeto de estudio. Frente a otros autores que han utilizado definiciones más estrechas de los programas educativos, han preferido adoptar una caracterización más amplia y cubrir así una parcela del sistema educativo suficientemente significativa. Ello ha influido en la diversidad de estudios abarcados, que permiten hacerse una idea comprensiva del desarrollo experimentado en este ámbito. La tercera razón estriba en que, definiendo así el objeto de estudio, han puesto el énfasis en las líneas de investigación evaluativa que más han crecido en España en la última década. Siendo éste un campo de corta tradición entre nosotros, no abundan los trabajos de cierta envergadura. Pero hay que reconocer que la evaluación de programas y, en menor medida, la de centros han atraído la atención de un número significativo de investigadores, con lo que empezamos a poseer una cierta tradición en ese ámbito concreto.

El hecho de haber sabido recoger los trabajos producidos en estos años (hasta alcanzar un número de cuarenta y nueve) y haberlos analizado cuidadosamente constituye una importante aportación de las autoras. Aunque el procedimiento de trabajo que han adoptado pueda en ocasiones privilegiar la descripción frente al análisis, no cabe duda de que en una obra que pretende ser herramienta de trabajo para los investigadores y los estudiosos esta decisión constituye más bien una virtud. El libro nos enseña muchas cosas, nos ayuda a sistematizar este campo de investigación pero, aún más, nos desperta interrogantes, nos acerca a los planteamientos de los evaluadores y nos ayuda a reflexionar sobre nuestra propia práctica. No se le debe pedir más de lo que es lícito pedirle.

Ahora que la evaluación del sistema educativo comienza su andadura firme en nuestro país, resulta un momento muy oportuno para realizar una primera síntesis del camino recorrido. Por provisional que dicha síntesis sea, el C.I.D.E. ha demostrado sentido de la oportunidad al difundirla. Y no cabe duda de su carácter provisional, que las propias autoras reconocen, por haber incluido únicamente las investigaciones financiadas desde el propio C.I.D.E. y por haber dejado al margen (por motivos que saben justificar) todo el campo de la evaluación de los aprendizajes. La dificultad objetiva para abarcar otras fuentes de promoción, incentivación y financiación de los estudios

evaluativos no debe hacer desistir de esfuerzos como el aquí realizado. Será responsabilidad de otras instituciones y Administraciones realizar un ejercicio similar, catalogando, reseñando y analizando los trabajos desarrollados por iniciativa suya.

En un momento como éste, en que la evaluación del sistema educativo atrae un número creciente de miradas, es fundamental ejercitar la transparencia de sus prácticas y la crítica de sus fundamentos, procesos y logros. Esta publicación se orienta en esta doble dirección, realizando una contribución tan modesta en su intención como sustantiva en su resultado. Y de ahí que, nuevamente, podamos felicitarnos de que ya esté a nuestra disposición.

ALEJANDRO TIANA FERRER

Director del Instituto Nacional de Calidad y Evaluación

Madrid, mayo de 1995

INTRODUCCIÓN

Este trabajo se propone ofrecer una síntesis comentada de las investigaciones que, sobre el tema de la evaluación de programas y centros educativos, han sido financiadas en el período de 1982 a 1992 por el Centro de Investigación y Documentación Educativa (C.I.D.E.), que durante una época (1989-94) se denominó Centro de Investigación, Documentación y Evaluación.

Esta obra no pretende ser una revisión exhaustiva de todos los estudios evaluativos sobre programas y centros realizados en España durante este período. Por un lado, y como ya se ha dicho, la presente revisión alcanza exclusivamente a la investigación financiada por el C.I.D.E. pero, además, trata de englobar sólo aquellas investigaciones que constituyen auténticas evaluaciones. En este sentido, cabría mencionar la dificultad que, en muchos casos, ha supuesto la decisión de incluir o excluir algunas investigaciones en esta obra. En la práctica, los límites de lo que es investigación o evaluación educativa son a menudo poco nítidos, por lo que se determinaron desde el principio unos criterios de demarcación de acuerdo a la literatura actual sobre el tema. Así, con el propósito de fundamentar las decisiones tomadas, y de exponer los criterios que las han sustentado, se contemplan en esta obra el significado y evolución del concepto de evaluación educativa, así como una tipología de la misma en función de distintos criterios. Con todo y con ello, en algunos casos, el propósito expresado por parte de un equipo de investigación de evaluar un programa o un centro se ha considerado como uno de los argumentos importantes para incluir el estudio en esta obra, con independencia de que el resultado se ajuste *sensu estricto* a lo que se entiende por evaluación.

Con las investigaciones que aquí se presentan, se trata de dar una idea general de qué aspectos han suscitado el interés de los investigadores en estos diez años, qué tipos de centros o de programas se han evaluado, qué enfoques y metodologías se han empleado y, en la me-

dida de lo posible, qué empleo se ha hecho de los resultados obtenidos. Asimismo, el lector encontrará una visión panorámica de las investigaciones incluidas en la que es posible apreciar una caracterización de los equipos de investigación, de los períodos de más intensidad de trabajo, de los niveles educativos más evaluados, etc.

Sin duda habrá quien se pregunte la razón de que en un compendio sobre evaluación no se hayan incluido investigaciones centradas en la evaluación del aprendizaje de los alumnos, limitándose a la evaluación de programas y centros. La razón es sencilla, aunque no lo fue su adopción.

Al inicio de esta obra su inclusión estaba fuera de duda, sin embargo, la revisión exhaustiva de las investigaciones realizadas en el período considerado, llevó a la conclusión de que apenas existían trabajos cuyo objeto de estudio fuese la evaluación de los aprendizajes. De hecho, se observó que casi siempre que se había emprendido la evaluación del rendimiento de los alumnos se había hecho con el fin de evaluar la pertinencia de un programa, material o método educativo. Así, al no encontrar apenas investigaciones en las que evaluar los aprendizajes constituyera un objeto de estudio en sí mismo, se ha optado por la clasificación que finalmente se presenta.

La detección de esta ausencia es ya un hallazgo importante de la revisión emprendida, evidenciándose la necesidad de fomentar en mayor medida la investigación centrada en el complejo tema de cómo evaluar los aprendizajes. En este sentido, y aunque el período comprendido por la presente revisión de investigaciones sobre evaluación se cierra antes de poder recoger los estudios financiados en las convocatorias posteriores, hay que decir que la convocatoria de 1993 del Concurso Nacional de Investigación Educativa del C.I.D.E., incluyó como uno de los temas de estudio “La evaluación de los aprendizajes”, con tres subapartados más específicos:

- Procedimientos, estrategias e instrumentos para la evaluación de capacidades de los alumnos en relación con la evaluación de los contenidos en diversas áreas y etapas educativas.
- Relaciones entre procesos de evaluación, calificación y decisiones de promoción de los alumnos.
- La información proporcionada a los diversos componentes de la comunidad escolar sobre la evaluación de los aprendizajes.

La intención de abordar una revisión de la evaluación de programas y centros educativos es múltiple. Por un lado, responde a la necesidad que existe siempre de aportar visiones generales, sistemáticas y ordenadas de los trabajos de evaluación financiados por un organismo que, como el C.I.D.E., está encargado del fomento de la investigación educativa. Por otro, el propio ámbito de la evaluación de programas y centros es en estos momentos una de las prioridades del Ministerio de Educación y Ciencia, por considerarlo una medida esencial para la mejora de la calidad de la enseñanza.

Pero, sin duda, ha habido otros motivos de peso para realizar una publicación de estas características. La reflexión sobre lo investigado en el pasado nos permite no sólo recoger y divulgar sus aportaciones, sino hacer prospectiva, facilitando la previsión de necesidades de investigación y, por tanto, una mejor planificación de los temas objeto de estudios futuros. Por último, es también una medida de control social de los rendimientos de la investigación en función del gasto realizado. Así, esta obra forma parte de una serie de publicaciones del C.I.D.E. que, con tales propósitos, han visto la luz en los dos últimos años. Los títulos que hasta el momento componen la serie son: *Diez años de investigación sobre profesorado, 1983-1993*; *Diez años de investigación e innovación en enseñanza de las Ciencias, 1983-1993*; *Diez años de investigación en didáctica de las Matemáticas, 1983-1993*; *Historia de la Educación en la España contemporánea. Diez años de investigación: 1983-1993*.

La evaluación de programas y centros, siendo el descriptor principal de esta revisión, abarca un campo muy extenso que requiere algunas precisiones, máxime cuando, como se ha dicho, en algunos casos no es fácil apreciar en la práctica las diferencias entre trabajos de investigación y trabajos de evaluación educativa o, por poner otro ejemplo, no siempre está claro qué se entiende por "programa" en educación. Ello puede haber motivado, en alguna ocasión, inclusiones o exclusiones discutibles. De hecho, prácticamente toda la investigación educativa aborda en alguna medida el tema de la calidad, la eficacia, la pertinencia, etc. de los programas e instituciones educativas. Por eso debe considerarse que, en las páginas siguientes, únicamente se recogen aquellos estudios que han llevado a cabo la evaluación de programas o centros educativos, siendo estos, eso sí, de muy diversas características.

La obra aparece estructurada en cinco capítulos. El primero es un capítulo introductorio que trata de clarificar el concepto y tipos de

evaluación educativa desde un punto de vista evolutivo. En el segundo capítulo se realiza una breve exposición de la situación de la evaluación de programas y centros en nuestro país para, en el capítulo tercero, proceder a encuadrar estos estudios de evaluación dentro del conjunto de investigaciones financiadas por el C.I.D.E. en el período tratado. Esto permite obtener una aproximación a la importancia relativa de este ámbito de estudio.

El cuarto capítulo, más extenso que los anteriores, se ocupa ya de la revisión del conjunto de evaluaciones de programas y centros financiadas en el período 1982-1992.

Este capítulo se inicia con una visión panorámica de las evaluaciones presentadas en función de diferentes variables (equipo de investigación, procedencia, nivel educativo, temporalización, etc.). A continuación, se presentan los criterios utilizados en la revisión y análisis de las evaluaciones, seguidos de tres grandes apartados: evaluación de programas, evaluación de centros e investigaciones centradas en la elaboración de modelos para la evaluación de programas y centros. Dentro de cada uno de ellos se ha procedido, a partir de los criterios definidos previamente, a la descripción y análisis detallado de las distintas investigaciones que conforman cada grupo. Este trabajo de análisis-síntesis ha estado presidido por el afán de preservar lo más fielmente posible el espíritu y la letra de los trabajos de los investigadores, objetivo sin duda imprescindible que se espera haber cubierto satisfactoriamente en todos los casos.

Para finalizar, se ofrecen algunas conclusiones o consideraciones generales acerca de los ámbitos de estudio abordados, los objetivos, la metodología, etc.

Queda señalar, por último, una información que el lector ha de tener en cuenta respecto a alguno de los trabajos que aquí se incluyen. Como es sabido, las convocatorias de Ayudas o los Concursos de investigación educativa del C.I.D.E. son de carácter anual y bienal, respectivamente. Esto, unido a que algún equipo ha podido prorrogar el plazo de finalización de su investigación, hace que en este período haya proyectos de investigación que, en la fecha de esta publicación, no han culminado todavía. Pero no por eso se ha querido dejar de consignar su existencia y características, mencionando en cada caso la situación del proyecto.

CAPÍTULO I

CLARIFICACIÓN CONCEPTUAL EN TORNO A LA EVALUACIÓN EDUCATIVA

1.1. La evaluación desde una perspectiva histórica

El ámbito de la evaluación constituye un elemento central y recurrente de la teoría y la práctica educativas. Proceder a una caracterización formal y exhaustiva del área de estudio excede con mucho la intención y espacio de este trabajo y, probablemente, tampoco sea éste el lugar para proceder a la formulación de una definición concluyente del término, máxime cuando la literatura sobre el tema proporciona un abundante número de ellas¹.

En consecuencia, se ha optado por señalar algunos de los contextos más significativos en los que se emplea el concepto de evaluación para, en un segundo momento, tratar de exponer las características básicas que definen y acotan el término y que, por lo mismo, son comunes a las múltiples definiciones de evaluación disponibles.

Definir el concepto de evaluación pasa en cierto modo por realizar una caracterización, en perspectiva histórica, de las formas en que se han venido conceptualizando y desarrollando las prácticas evaluativas, lo que ha de hacerse necesariamente al hilo de los dos tipos de evaluación más evolucionados: la evaluación del rendimiento o de las capacidades de los sujetos, y la evaluación de programas.

Los orígenes de lo que ha llegado a ser la evaluación educativa cabe situarlos en la preocupación por la medición científica de los ras-

¹ Pueden consultarse, a título de ejemplo, las veinte definiciones recogidas por Lázaro (1990), entre las que se incluyen formulaciones realizadas por autores españoles y extranjeros.

gos y conductas humanas iniciada a finales del siglo XIX y principios del XX. Se define en ese momento una evaluación centrada en el establecimiento de diferencias individuales, en un intento por determinar la posición relativa del sujeto dentro de un grupo normativo. La aplicación de un test de rendimiento era sinónimo de evaluación educativa, dejando en el olvido la representatividad y congruencia con los objetivos educativos. La evaluación así entendida, tenía poca relación con los programas y con el funcionamiento de las escuelas. Los tests informaban sobre los alumnos, pero nada sobre los programas o procesos de enseñanza (Guba y Lincoln, 1982).

Será en los años 30 cuando un estudio realizado por Tyler y Smith abra un nuevo período en la teoría y práctica evaluativas, al superar el concepto de evaluación como mero proceso externo de calificación de habilidades aprendidas por los sujetos sometidos a un proceso instructivo.

El modelo clásico propuesto por Tyler, vino a definir la evaluación como el proceso que tiene por objeto determinar en qué medida se han logrado unos objetivos previamente establecidos. Las discrepancias entre la ejecución y los objetivos orientarían las modificaciones a realizar, y el ciclo de la evaluación se iniciaría y completaría de nuevo. Toda la teoría y práctica evaluativas y aún educativas posteriores, se verán fuertemente influenciadas por este enfoque.

Como señala Carballo (1990) la gran influencia del concepto de Tyler en las posteriores concepciones de la evaluación se debió a que era un concepto lógico, aceptable desde un punto de vista científico, y fácil de aplicar para los evaluadores de programas. Se desarrollaron, pues, modelos que asumían y desarrollaban los presupuestos citados.

No obstante, el modelo de Tyler fue y es objeto de algunas críticas achacándosele, entre otras limitaciones, la falta de un auténtico componente evaluativo (al facilitar más la medida y la evaluación de los objetivos, que la obtención de juicios explícitos de valor que permitan apreciar el éxito global de un programa), la falta de criterios para juzgar la importancia de las discrepancias observadas, y su dependencia de una filosofía que promueve un estudio de la evaluación excesivamente lineal e inflexible. Al mismo tiempo, se ha destacado el hecho de que cuando un programa no alcanza sus objetivos, no significa necesariamente que el programa no tenga calidad o sea ineficaz, habiendo otros factores de proceso y/o de contexto que pueden explicar esa situación.

En Estados Unidos, los años 60 y 70 se caracterizaron por la búsqueda de modelos de evaluación que permitieran superar el modelo desarrollado y publicado por Tyler en 1942, siendo éste objeto de numerosas revisiones críticas y superadoras en períodos posteriores (Bartolomé 1990).

Carballo (1990) ofrece una descripción sintética de la evolución del concepto de evaluación en este período, atendiendo al desarrollo de los modelos de evaluación de programas, que es el que se adopta en estas páginas para su caracterización.

Efectivamente, las dudas en torno a la pertinencia de determinar la calidad de un programa exclusivamente en función de la consecución de sus objetivos, motivó el desarrollo de los *modelos de evaluación basados en la experimentación*, los cuales asumían, genéricamente, que para demostrar la eficacia de un programa educativo la evaluación tenía que establecer una relación causal entre el rendimiento y el propio programa. En el marco de este enfoque se defendían dos diferentes tipos de diseños evaluativos: los diseños experimentales puros, o los diseños dotados de mayor flexibilidad. Esta polémica acabó orientándose hacia la defensa de una evaluación comparativa, capaz de determinar la bondad relativa de un programa educativo en relación con los resultados obtenidos por otro programa, es decir, a la estimación normativa de la eficacia desde un planteamiento causal.

Las dificultades intrínsecas a este enfoque comparativo plantearon, por vez primera, la necesidad de valorar y juzgar la conveniencia y utilidad de los propios objetivos de los programas. Cuando distintos autores comienzan a preguntarse por el valor educativo que se le puede atribuir a un programa que se muestra eficaz, pero que tiene unos objetivos educativamente irrelevantes, se inicia el desarrollo de los llamados *modelos orientados al juicio*, centrados en el valor de la competencia profesional del evaluador para juzgar el programa educativo en cuestión.

Frente a la caracterización estrictamente “objetiva” y “científica” hasta entonces dominante en los enfoques de la evaluación centrados en los objetivos y en los experimentos, se empieza a abrir paso la consideración de los criterios y de los juicios valorativos como rasgos fundamentales de la actividad evaluadora. Así, Scriven (1967) defendió la acción de valorar, como un elemento esencial de toda evaluación, aplicándolo a los componentes de los programas, a sus resultados y a las decisiones a las que tienen que hacer frente sus responsables, bases sobre las cuales desarrolló su conocido modelo.

En esta línea pueden situarse dos modelos de evaluación concretos. El *modelo de la figura* de Stake, centrado en dos tareas: la descripción del programa evaluado y la emisión de una valoración global sobre el mismo; y el *modelo de consulta a un experto* de Eisner, en el que el evaluador desempeña el papel de crítico.

Al reconocimiento de las aportaciones de este enfoque se suman, no obstante, algunas críticas referidas al carácter y naturaleza de las valoraciones que realiza un único evaluador basándose en su sola competencia, y al contexto y consecuencias de dichas valoraciones. Para superar estas limitaciones, se desarrollarían los denominados *modelos políticos o de debate*, englobando aquellas evaluaciones en las que diferentes evaluadores, o grupos de evaluadores, ofrecen sus distintos puntos de vista sobre los diferentes aspectos de un programa. Modelos representativos de esta perspectiva son los *modelos de perspectiva contrapuesta* o los de *orientación judicial*.

Estas consideraciones llevaron a adoptar distintas posturas relativas al papel del evaluador de programas, así como nuevos criterios de valoración. Entre estos últimos destaca la consideración del coste de un programa como objeto de evaluación o, *modelo de evaluación de costos*. Dentro de esta perspectiva las propuestas más interesantes son la evaluación de las relaciones costo-beneficio, costo-eficacia, costo-utilidad y costo-viabilidad, estudios que permiten valorar si las inversiones efectuadas en un programa o intervención educativa son adecuadas y rentables.

Desde este momento se hará evidente la necesidad de valorar los programas no sólo atendiendo a la consecución de sus objetivos, sino también a las consecuencias no previstas que se pueden derivar de su aplicación. Fundamentalmente será Scriven (1967) el autor que más significativamente contribuya al establecimiento y desarrollo del *modelo de evaluación libre de objetivos*.

Por otro lado, el reconocimiento de la importancia de la evaluación en cualquier proceso de toma de decisiones (administrativas, políticas o pedagógicas) ha constituido un referente común a todas las propuestas realizadas.

Las posiciones adoptadas en relación a este último aspecto permiten diferenciar distintos modelos centrados en las necesidades de las personas directamente implicadas en la toma de decisiones, en los que la coordinación del evaluador con los responsables de tomar dichas decisiones se considera un requisito de la eficacia de la evaluación.

Destacan en este sentido el *modelo de la planificación educativa* de Cronbach, el *modelo C.I.P.P. de evaluación global* de Stufflebeam, el *modelo centrado en el cliente* de Scriven, o el *modelo de evaluación formativa* de Brinderhoff, entre otros.

Con el paso del tiempo, se ha venido reivindicando un papel más comprometido para el evaluador, así como una mayor importancia a las valoraciones y necesidades de los distintos sujetos y grupos implicados en un programa o institución educativa. Se espera que el evaluador considere e interprete las diferentes posturas, sopesando de manera intuitiva y equilibrada dicha pluralidad. En definitiva, desde el enfoque de la evaluación como iluminación, o evaluación pluralista, los destinatarios de un programa educativo son quienes mejor pueden informar sobre la calidad del mismo. Ejemplos significativos de esta perspectiva son el *modelo de evaluación respondente* de Stake, la *evaluación iluminativa* de Parlett y Hamilton, la *evaluación transaccional* de Rippey y la *evaluación democrática* de McDonald.

En esta aproximación histórica a la evaluación se ha optado por una de las aproximaciones disponibles, ya que existen abundantes trabajos que responden a otros tantos criterios de clasificación de los modelos de evaluación. En cualquier caso, se cree haber puesto de manifiesto que el período tratado constituye una etapa de desarrollo espectacular de la evaluación educativa —sobre todo en el contexto anglosajón de la evaluación de programas—, que surgió como consecuencia del descontento con los resultados de la enseñanza pública y la provisión de ingentes cantidades de fondos públicos destinados a subvencionar nuevos programas educativos, lo que demandaba una adecuada evaluación de su calidad. Este movimiento de responsabilidad y transparencia, que se desarrolló en los Estados Unidos con el nombre genérico de “accountability” a principios de los años 70, consistió en una demanda y una voluntad de rendir cuentas sobre la efectividad o mérito de los programas educativos, que ha tenido su continuidad hasta nuestros días.

Dicho período ha cristalizado en la profesionalización de la evaluación, en los primeros intentos de superar la dicotomía entre las metodologías cualitativa y cuantitativa en la investigación evaluativa (Reichardt y Cook, 1979), y en el desarrollo de unas normas para la evaluación de materiales y programas educativos (Joint Committee on Standards for Educational Evaluation, 1981), que constituyen una conocida y sistemática propuesta de metaevaluación.

Los años 80, en opinión de Alvira (1991) se caracterizaron por la aparición de una pluralidad de perspectivas, exageradamente consideradas modelos, y de nuevos métodos de evaluación; por un esfuerzo creciente en lograr unas evaluaciones útiles, y por la ampliación de los objetivos y los ámbitos de la evaluación educativa, rebasando con creces la evaluación de resultados predominante en períodos anteriores.

El mismo autor señala también cómo las restricciones de los fondos destinados a la evaluación de programas en Estados Unidos, coincidió con el incremento de la evaluación educativa en otros países occidentales. En este sentido, Rutman (1984) afirma que mientras que en los años 80 en Estados Unidos el campo de la evaluación se iba reduciendo considerablemente y el interés por la evaluación de programas declinaba, en otros países parecía desarrollarse un gran interés por la evaluación de programas educativos².

Por su parte, Tejedor y otros (1993), desde un punto de vista más metodológico, describen la situación actual de la evaluación como plural y compleja, tendente a un mayor realismo en sus planteamientos, de modo que las nuevas perspectivas no sustituyen a las anteriores, sino que las complementan. A juicio de este autor, los cambios habidos en los últimos años pueden ser sintetizados en dos concepciones básicas. La tradicional, desde la que la evaluación se entiende como el empleo de métodos científicos para medir los logros educativos con el fin de iluminar la toma de decisiones, y la postura más actual, que entiende la evaluación como el conocimiento del mérito o valor de un programa, centro, etc., en función de unos criterios que permiten tomar decisiones ajustadas a las necesidades de los usuarios y/o responsables del programa.

En cualquier caso, la evaluación de programas y la evaluación de instituciones educativas constituyen hoy dos ámbitos de estudio diferenciados, con una terminología, un conjunto de herramientas conceptuales y analíticas, y unos procesos y fases, propios y específicos. Al mismo tiempo, se observa cómo el trasvase de aportaciones de un ámbito a otro es constante.

A juicio de Bartolomé (1990) en nuestro país ha sido notable en los últimos años el incremento de la financiación de los estudios evaluativos, destacando el hecho de que la reforma educativa promovida

² Puede citarse como ejemplo, la reunión de científicos y políticos de los países de la OTAN celebrada en La Haya en diciembre de 1990 bajo el título: "*International Developments in Program Evaluation Research*".

por la L.O.G.S.E. ofrece un gran respaldo e impulso a la evaluación, tanto a nivel estatal, como autonómico.

Los primeros momentos de la reforma educativa supusieron el diseño y desarrollo de un considerable número de programas experimentales de iniciativa institucional, de mayor o menor alcance, que fueron objeto de evaluación. Junto a ellos, la propia reforma del currículo (complejo conjunto de programas educativos) ha sido objeto de sucesivos y amplios estudios evaluativos, muchos de ellos realizados por el C.I.D.E.. El lector podrá contrastar personalmente esta situación a lo largo de las páginas de este libro.

También en el ámbito de la enseñanza universitaria se ha ido desarrollando un movimiento evaluativo que, aunque centrado fundamentalmente en la evaluación institucional, abarca aspectos clave para la mejora de la calidad de la enseñanza, entre los que se incluye la evaluación de programas.

En conclusión, la evaluación, y especialmente la evaluación de programas, dispone ya de una historia propia construida durante los últimos treinta años. En conjunto, y por lo que se refiere a nuestro país en particular, en los últimos años se ha asistido a una notable demanda de evaluación sistemática de la educación, demanda que ha sido asumida por el propio sistema educativo, incorporando procedimientos intencionales y planificados de intervención en este sentido (Tiana, 1993).

1.2. Delimitación de conceptos: medida, evaluación, investigación

A la complejidad conceptual del tema de la evaluación se suma el hecho de que, en castellano, este término engloba un conjunto de significados diversos lo que, unido a reduccionismos de distinto signo, ha oscurecido notablemente el significado de la evaluación y su relación con términos más o menos próximos. Por ello, quizá sea de utilidad establecer un paralelismo idiomático con los distintos términos utilizados en el idioma inglés, en el que la ambigüedad terminológica es, en este caso, más reducida.

El término *testing* está referido a la aplicación de ciertos instrumentos o pruebas objetivas utilizadas en los procesos de evaluación, o a ciertas modalidades psicométricas que en cualquier caso no agotan los instrumentos o estrategias a utilizar en la evaluación de alumnos.

Asimismo, Choppin (1985) establece una diferencia, ya clásica, entre el término inglés *assessment*, entendido como valoración, apreciación o juicio cuando se aplica sobre las características de las personas, y el término *evaluation*, preferiblemente utilizado en el caso de entidades abstractas como programas, currículos y variables organizativas. La distinción pues entre ambos términos cabría establecerse en función del objeto, siendo *evaluation* el término vinculado al análisis de la posible influencia de los distintos componentes del proceso de enseñanza en el aprendizaje, mientras que la acción a que se refiere el término *assessment* está en relación con los individuos o grupos receptores y responsables de la enseñanza (Scriven, 1991).

A tales apreciaciones, Angulo (1990) añade un segundo criterio de diferenciación: el sentido del proceso. El término *evaluation* hace referencia a la valoración del servicio educativo prestado, y *assessment* a la de los receptores y/o agentes de dicho servicio.

Desde el punto de vista teórico resulta hoy clara la diferenciación entre la evaluación y otros conceptos próximos como medida o calificación. El solapamiento de estos conceptos o su identificación, constituyen una simplificación de lo que implica la acción de evaluar.

En un sentido general, medir puede entenderse como la asignación de cantidades a determinadas magnitudes, o en términos ya clásicos de Stevens, como la asignación de un numeral a un objeto. Resulta, pues, un proceso libre de referencias valorativas que se limita a estudiar atributos que pueden ser observados empíricamente. Los procesos de medida, en su sentido más amplio y en este contexto, han de entenderse como un componente de la evaluación, con los que en ningún caso puede ser identificada, por relevante que sea el papel que el hecho de medir desempeñe en la misma. Como se ha mencionado, la corriente positivista dominante en el ámbito de la evaluación educativa en torno a los años 20, aparecía vinculada a la medición de las características individuales (Tejedor, 1993), visión reduccionista ampliamente superada en la actualidad. A partir de su propia etimología, y tal como se desprende de las definiciones más académicas, la valoración constituye la esencia del acto de evaluar.

En definitiva, al tratar de definir el concepto de evaluación educativa, diversos autores coinciden en hacerlo en función de los distintos rasgos que hasta aquí se le han venido atribuyendo. Se caracteriza, así, como un *proceso* en el que tienen lugar procedimientos de *recogida de información* así como su posterior *interpretación* en función del con-

traste con determinadas instancias de referencia o patrones de deseabilidad, de modo que se posibilite la emisión de un juicio de valor de la realidad evaluada, capaz de proporcionar determinadas pautas para orientar la acción o la toma de decisiones.

Estos elementos, de un modo u otro, y con diversos matices, pueden considerarse constitutivos de todo proceso evaluador, con independencia de los diferentes modelos o perspectivas existentes y del objeto de la evaluación. Sin embargo, cabe destacar que no ha sido fácil ni frecuente encontrar la presencia de todas estas acciones en las evaluaciones recogidas en esta publicación. A menudo no se explicitan los criterios de referencia empleados para valorar, o no se hace mención a cómo los resultados de la evaluación orientarán la acción o la toma de decisiones. De hecho, muy pocas veces se ha explicitado el concepto o modelo de evaluación que los investigadores han manejado en su trabajo.

En cualquier caso, conviene detenerse brevemente en cada uno de los elementos o características de la acción evaluadora.

El énfasis puesto en el carácter procesual de la evaluación pretende destacar que no se trata de una acción ajena o separada de la intervención educativa, sino que constituye una dimensión muy importante de la misma.

Por otra parte, es un hecho asumido que objeto de evaluación puede ser cualquier fenómeno, situación o actor educativo. Así, el ámbito y campo de aplicación de la evaluación educativa se extiende a alumnos, profesores, directivos, al currículo, las instituciones, los programas educativos, la Administración, etc.

En el curso de cualquier actividad evaluativa, se planifican y llevan a cabo procesos de recogida sistemática de información. Tal y como apunta Pérez Juste (1989) ninguna valoración posterior ni, en consecuencia, las decisiones que se toman sobre la base de los juicios de valor, van a ser mejores que la información en que se basen. Ello determina el carácter planificado y sistemático que se les supone a los procedimientos de recolección de datos. Se trata de obtener datos representativos, información valiosa. Esta posición no prejuzga ni especifica la clase de información que se ha de recoger, o el tipo de metodología a utilizar para ello (aplicación de pruebas normalizadas, de cuestionarios, realización de entrevistas, observación directa, etc., o empleo de una metodología cuantitativa, cualitativa o mixta para el análisis de los datos). En principio nada se excluye, y el empleo de unas u otras lo determinará el tipo o modelo de evaluación por el que se opte.

En todo proceso evaluador, la información recogida es interpretada, pues evaluar no consiste exclusivamente en recoger información, sino en interpretarla y valorarla en función de unos criterios para explicar o comprender la realidad que se analiza. Ello nos lleva a la consideración del papel que juega la emisión de juicios de valor que, junto a la orientación de la toma de decisiones, es probablemente el rasgo más distintivo de la evaluación.

Pero la delimitación del concepto de evaluación, no sólo exige evitar reducir este complejo proceso a la mera acción de recoger información o medir, también es necesario establecer una demarcación con respecto a un concepto tan próximo como el de investigación. Al menos, esta fue una tarea que se acometió antes de realizar una publicación que, como ésta, resume diez años de investigación en evaluación.

Ciertamente, las relaciones existentes entre evaluación e investigación son estrechas (De la Orden, 1985; Tejedor, 1993; Cohen y Franco, 1988; entre otros). Sintéticamente puede expresarse la existencia de un importante denominador común entre ambas actividades, de convenir en que la evaluación exige a menudo parecido grado de control y rigor que la investigación (De la Orden, 1985) y que se distingue de ella por los objetivos que la orientan, más que por los métodos que se emplean (Tejedor, 1993).

De la Orden (1985) cifra en las siguientes las características peculiares y diferenciales de la evaluación en relación a la investigación *sensu stricto*:

- Los juicios de valor acompañan al desarrollo de todo el proceso: desde la selección del problema de estudio hasta la aplicación de metodologías concretas.
- Resulta difícil, y a veces inadecuado, formular hipótesis precisas.
- La replicación es prácticamente imposible, dado que el estudio de evaluación está íntimamente vinculado a un determinado programa o institución, a su contexto espacial y temporal específico.
- La recogida de datos está determinada por la viabilidad del proceso.
- Multitud de variables relevantes sólo son susceptibles de un control superficial, limitándose por tanto el empleo de diseños experimentales.

- Las decisiones últimas relativas a la paralización, sustitución o continuidad de un programa o institución educativa, pueden no ser competencia del evaluador.
- El informe evaluativo debe adaptarse a las exigencias de quien toma las decisiones sobre el programa o centro.

Aunque esta relación de características diferenciadoras de la evaluación no haya sido aplicada de modo estricto a la revisión de investigaciones que aquí se acomete, sí constituye uno de los marcos de referencia considerados.

CAPÍTULO II

LA EVALUACIÓN DE PROGRAMAS Y CENTROS EN ESPAÑA EN EL PERÍODO 1982-1992

La atención recibida por la evaluación educativa en los últimos años en nuestro país ha sido creciente. Buena prueba de ello puede obtenerse del análisis de la legislación educativa o de la revisión de las publicaciones, congresos y seminarios realizados en torno a estos dos temas.

2.1. La evaluación y los períodos de reforma educativa

La importancia creciente de la evaluación suele aparecer vinculada a los procesos de reforma o de innovación en el sistema educativo y, por lo tanto, a la toma de decisiones en política educativa. Esto ha hecho que el ámbito de la evaluación se haya ido ampliando paulatinamente, como demuestra la revisión de la legislación en materia de educación producida en los últimos años.

Ya en la Ley General de Educación de 1970 se contemplaba la evaluación como un medio para valorar, no sólo el rendimiento de los alumnos, sino también el propio proceso educativo; como un instrumento imprescindible para mejorar la calidad del sistema en su conjunto. Analizando la Orden Ministerial de 16 de noviembre de 1970 que regulaba los principios que debían regir los procesos de evaluación educativa, se observa que, indirectamente, algunos principios se referían a la evaluación de programas. En ellos se recomendaba, por ejemplo, la valoración de los métodos y procedimientos educativos empleados, la determinación de la adecuación del contenido de los

programas y su selección de acuerdo con su valor formativo, y la estimación de en qué medida se alcanzaban los objetivos previstos en la programación educativa, contrastando su validez.

Por su parte, el Proyecto para la Reforma de la Enseñanza (1987) consideraba la evaluación como un proceso referido, tanto a los procedimientos que permiten analizar los procesos de enseñanza y aprendizaje, como a los que permiten analizar el funcionamiento del sistema educativo en su totalidad, y el de todos los agentes que intervienen en la acción educativa.

En la década de los 80, los cambios sociales y laborales, los avances científicos y tecnológicos, las disfuncionalidades mostradas por el sistema vigente y la necesidad de adaptarse al contexto europeo, demandaban una reordenación y renovación del sistema educativo español. En estos años todo lo relacionado con la evaluación adquirió un espectacular desarrollo, debido a la necesidad de valorar metodologías y materiales didácticos, programas y estructuras educativas, con el fin de buscar respuestas a las nuevas necesidades y demandas. Es un momento en el que la investigación relativa a la evaluación adquiere una mayor relevancia y atención.

De este modo, y aunque desde la Administración educativa ya se habían venido haciendo estudios evaluativos, es a partir de la planificación, diseño y experimentación de la reforma, cuando los esfuerzos por fomentar la investigación evaluativa se van a ver incrementados. La evaluación se convierte, por tanto, en un instrumento imprescindible para lograr una adecuada implantación de una reforma educativa experimental que, tras la incorporación de las oportunas mejoras, cristalizaría en la Ley de Ordenación General del Sistema Educativo (L.O.G.S.E., 1990).

Así, a partir de 1984, y con objeto de contrastar la reforma que se estaba gestando antes de su aprobación definitiva, el Ministerio de Educación y Ciencia inicia la puesta en marcha de una serie de programas experimentales cuya evaluación permitiría detectar y, sobre todo, corregir, las disfuncionalidades encontradas en esa fase.

En este sentido destacan la Evaluación Externa de la Reforma Experimental de las Enseñanzas Medias y de los Módulos Profesionales, ambas realizadas por el C.I.D.E. e iniciadas, respectivamente, en los años 1984 y 1989, así como la Evaluación del Programa de Integración Escolar (iniciada en 1985).

Superada la fase experimental, la propia L.O.G.S.E. otorga un importante papel a la evaluación, ampliando su ámbito de aplicación y

estableciendo los mecanismos y estructuras responsables de llevarla a cabo. Así, es especialmente en su Título IV “*De la calidad de la enseñanza*”, en el que se plasman las directrices más importantes. En su Artículo 55 se establece que los poderes públicos prestarán una atención prioritaria al conjunto de factores que favorecen la calidad y mejora de la enseñanza, entre otros, la evaluación del sistema educativo.

En su Artículo 62.1, la Ley amplía el ámbito de la evaluación: “*La evaluación del sistema educativo se orientará a la permanente adecuación del mismo a las demandas sociales y a las necesidades educativas, y se aplicará sobre los alumnos, el profesorado, los centros, los procesos educativos y sobre la propia Administración*”.

Una de las novedades más importantes de la reforma en lo que respecta a evaluación, la constituye la creación del Instituto Nacional de Calidad y Evaluación (I.N.C.E.). En el Preámbulo de la L.O.G.S.E. se justifica su creación en los siguientes términos: “*La Ley atribuye una singular importancia a la evaluación general del sistema educativo, creando para ello el Instituto Nacional de Calidad y Evaluación. La actividad evaluadora es fundamental para analizar en qué medida los distintos elementos del sistema educativo están contribuyendo a la consecución de los objetivos previamente establecidos. Por ello, ha de extenderse a la actividad educativa en todos sus niveles, alcanzando a todos los sectores que en ella participan.*”

En su Artículo 62, la Ley establece las funciones del I.N.C.E.: a) *Elaborar sistemas de evaluación para las diferentes enseñanzas reguladas en la Ley y sus correspondientes centros;* b) *Realizar investigaciones, estudios y evaluaciones del sistema educativo y, en general, proponer a las Administraciones educativas cuantas iniciativas y sugerencias puedan contribuir a favorecer la calidad y mejora de la enseñanza.*

Sin duda, se puede considerar que la gestación y generalización de la L.O.G.S.E. ha tenido, tiene y tendrá en el futuro una enorme influencia en el crecimiento de la investigación en evaluación educativa.

2.2. Foros nacionales de discusión en torno a la evaluación de programas e instituciones educativas

Otro de los indicadores de la atención que han recibido en España los temas de la evaluación de programas e instituciones educativas,

puede ser el número de congresos, seminarios o jornadas que se han celebrado en torno a los mismos. Haciendo una revisión de aquellos que han tenido lugar desde 1982, promovidos desde distintas instancias públicas o privadas, se observa que, a pesar del indudable interés que el tema ha despertado y despierta en la comunidad educativa y científica, ésta no parece haber sido la vía principal de canalización del debate.

Ciertamente existe constancia de que este tipo de evaluación se ha abordado más o menos colateralmente en muy diversos foros y encuentros de expertos celebrados en torno a otros temas nucleares como: integración escolar, educación de adultos, necesidades educativas especiales, dirección escolar, reforma del sistema, formación del profesorado, etc. Sin embargo, han sido menos frecuentes las reuniones cuyo tema central de discusión fuera la evaluación de programas y centros. Entre ellas, destacarían las siguientes:

- La Asociación Española para el Desarrollo de la Educación Especial (A.E.D.E.S.), celebró en diciembre de 1987 una reunión científica bajo el título genérico de: *La evaluación de programas educativos*.
- La Asociación Interuniversitaria de Investigación Pedagógica Experimental (A.I.D.I.P.E.) ha celebrado distintos Seminarios relativos al tema:
 - En 1988, el IV Seminario de *Modelos de Investigación Educativa*, en el que se puede destacar la ponencia titulada "Evaluación de programas de orientación: la evaluación del contexto y del diseño" (Alvarez Fernández).
 - En 1990, el V Seminario sobre *Metodología en el diagnóstico y evaluación de los procesos de intervención educativa*.
 - En 1993, el VI Seminario en torno al tema *Investigación sobre diferenciación, educación y orientación: la dimensión metodológica*, en el que se puede destacar, por el tema que trataba, la ponencia titulada "Nuevas aportaciones en la evaluación de programas de intervención pedagógica" (Tejedor, García Valcarcel y Rodríguez Conde).
- El Fórum Europeo de Administradores de la Educación (F.E.A.E.), ha promovido desde 1988 la celebración de dos

Jornadas en relación con la evaluación educativa: *Cómo evaluar el sistema educativo* (San Sebastián, 1990), y *Evaluación de Programas de Formación* (Barcelona, 1991).

- Especialmente importante puede considerarse el Seminario sobre *Las reformas educativas actuales en España, su evaluación: metodología y resultados*, organizado en 1993 por el C.I.D.E., el Departamento de Métodos de Investigación y Diagnóstico en Educación de la U.N.E.D., y el Ilustre Colegio Oficial de Doctores y Licenciados en Filosofía y Letras y en Ciencias de Madrid.
- En septiembre de 1994, la Asociación de la Economía de la Educación (A.E.D.E.) celebró en Barcelona sus III Jornadas en torno al tema de la *Planificación y evaluación de sistemas educativos*, en las que se presentaron numerosas ponencias y comunicaciones relativas a la evaluación de reformas, programas e instituciones educativas universitarias y no universitarias.
- Asimismo, ha ido cobrando creciente importancia la evaluación de la enseñanza universitaria, de lo que da buena prueba la elaboración y puesta en marcha por parte del Consejo de Universidades, de un *Plan Experimental de Evaluación de la Calidad del Sistema Universitario*, que fue presentado en los cursos de verano de la Universidad Internacional Menéndez y Pelayo (1994). Dicha presentación obtuvo una gran repercusión entre los Administradores universitarios y los expertos en el tema, lo que ha dado lugar a un importante debate público sobre el modelo y resultados de la evaluación experimental de la Universidad española.

En definitiva, han sido numerosos y relevantes en estos últimos doce años, los foros cuyo centro de interés ha sido precisamente la evaluación de programas y centros o instituciones educativas.

CAPÍTULO III

LAS EVALUACIONES DE PROGRAMAS Y CENTROS EN EL CONJUNTO DE LA INVESTIGACIÓN FINANCIADA POR EL C.I.D.E.

Antes de proceder a la revisión del contenido de las investigaciones, parece adecuado encuadrar los estudios sobre evaluación de programas e instituciones educativas en el conjunto de la investigación educativa financiada o realizada por el C.I.D.E. Esto nos permitirá calibrar su importancia relativa, así como hacer una estimación de si estos tipos de evaluación han constituido líneas prioritarias de investigación en los últimos años.

3.1. Áreas prioritarias de investigación

Como ya se ha mencionado, y en relación con la reorganización del sistema educativo y con las innovaciones en educación que se estaban llevando a cabo en España en los años 80, la Administración pública adquirió el compromiso de fomentar estudios dirigidos a obtener un mejor conocimiento del sistema y a evaluar su funcionamiento y el de los programas y procesos de reforma, que proporcionan el fundamento para la planificación y la toma de decisiones en política educativa.

Este compromiso se materializa con la formulación en 1989, por parte del Ministerio de Educación y Ciencia, del *Plan Nacional de Investigación Educativa* encuadrado en el *Plan Nacional de Investiga-*

ción Científica y Técnica, que sugiere unas áreas de investigación prioritarias. Entre las cinco áreas que hacen referencia a la evaluación, destaca la destinada a “*evaluación de programas educativos: educación especial, integración escolar, educación compensatoria y de adultos*”.

El organismo encargado de gestionar el Plan Nacional de Investigación Educativa fue el Centro de Investigación, Documentación y Evaluación Educativa el cual, como se explica en el Prólogo de esta obra, desde la creación del Instituto Nacional de Calidad y Evaluación (I.N.C.E.) en 1994, ha pasado a denominarse Centro de Investigación y Documentación Educativa, conservando las mismas siglas (C.I.D.E.). Desde este organismo se ha venido promoviendo la investigación por medio, entre otras actividades, de las convocatorias de los *Concursos Nacionales de Proyectos de Investigación Educativa*, de las *Ayudas a la Investigación Educativa* y de los *Premios Nacionales a la Investigación e Innovación Educativa* (en sus tres modalidades: tesis, investigación e innovación).

Con respecto a los Concursos Nacionales de Proyectos de Investigación, las distintas convocatorias que se vienen realizando desde 1985 han ido delimitando las líneas fundamentales en investigación educativa, algunas de las cuales se ocupan de la evaluación de programas.

Así, el Concurso de 1985 establecía cuatro temas prioritarios, tres de los cuales están relacionados con la evaluación de programas: “1) planteamiento, seguimiento y evaluación de la reforma de las Enseñanzas Medias; 2) evaluación de los usos y resultados de las nuevas tecnologías educativas, en particular el vídeo y los ordenadores; y 3) evaluación del programa de Educación Compensatoria”.

A partir de ese año no se alude expresamente al tema de evaluación de programas hasta la convocatoria de 1992; sin embargo, exceptuando la de 1990, todas abordan esta temática aunque de forma implícita.

En la convocatoria de 1986, que se centra en la relación entre educación y economía, se incorporaron dos temas de estudio que, indirectamente, se referían a la evaluación de programas: “1) Apariencia y realidad de la oferta educativa y su papel en la generación y consolidación de la demanda actual, analizando las derivaciones para la evaluación y planificación de las enseñanzas postobligatorias, y 2) Organización, costes, finan-

ciación y eficiencia del sistema de enseñanza postobligatoria no universitaria”.

El Concurso de 1988 giraba en torno a “los procesos de renovación e innovación educativa” y de las cuatro áreas prioritarias que establecía, hay una que destaca por su relación con la evaluación de programas: “los procesos reales de interacción educativa y formas y recursos necesarios para mejorar su calidad y eficacia”.

Los procesos de renovación e innovación educativa, son también el centro de interés de la convocatoria de 1989, de la cual interesa destacar la propuesta de los siguientes temas de investigación: “1) evaluación de innovaciones metodológicas y desarrollo de instrumentos de evaluación de centros escolares; 2) desarrollo y evaluación sistemática de nuevos materiales didácticos; y 3) proyectos que incorporen el análisis de los efectos de los nuevos medios (tecnologías de la información y la comunicación) sobre el aprendizaje”.

Por último, en el Concurso de 1992, cuya convocatoria establece entre sus áreas de estudio una denominada “El profesorado en las enseñanzas contempladas en la Ley General de Ordenación del Sistema Educativo”, se hace mención expresa de la evaluación de programas en los siguientes términos: “La evaluación de los programas de formación del profesorado en Centros educativos: propuesta de modelos de evaluación; aplicación de los mismos”.

Por su parte, las convocatorias de Ayudas a la Investigación tratan de estimular la investigación educativa en general, sin fijar temas prioritarios, aunque a partir de 1991 se han sugerido algunos temas de interés preferente.

Así, en ese mismo año se definían, entre otras áreas de estudio, “el desarrollo y la evaluación de Proyectos de Centro y de proyectos de innovación”. Asimismo, en la convocatoria de 1992 se señalaban como temas de especial relevancia aquellos que estuvieran relacionados con el proceso de reforma del sistema educativo.

Se puede observar, pues, que la evaluación de programas estuvo presente en ambas convocatorias, si bien referidas, fundamentalmente, a cuestiones relacionadas con la reforma del sistema educativo.

En cuanto a los Premios Nacionales de Investigación e Innovación Educativa, que convoca el Ministerio de Educación bienalmente

a través del C.I.D.E., cuatro son los proyectos premiados en los últimos 10 años que hacen referencia a evaluación de programas:

- *“Diseño, aplicación y evaluación de dos metodologías (expositiva-audiovisual y experimental-integrada) para la enseñanza de la Física en E.G.B”*, de Antonio de Pro Bueno (Premio Nacional 1985).
- *“Análisis de la eficacia del microordenador en la enseñanza de las matemáticas en 7.º de E.G.B. y 2.º de B.U.P.”*, de Antonio Bautista García-Vera (Premio Nacional 1988).
- *“Modelo para evaluar una programación”*, de Mercedes Abril Alonso (Premio Nacional 1988).
- *“Enseñar a comprender: diseño y valoración de un programa de instrucción para formar a los profesores en la enseñanza de estrategias de comprensión de las ideas principales en el aula”*, de Nuria Carriedo López (Premio Nacional 1992).

3.2. La evaluación de programas y centros en el contexto de la investigación educativa impulsada por el Ministerio de Educación y Ciencia

Reviste cierta dificultad establecer una clasificación unívoca de las investigaciones realizadas, ya que cualquier sistema de categorías o áreas temáticas que se utilice no es excluyente. De esta forma, una investigación evaluativa de unos cursos de formación de profesores, puede ser clasificada en el área temática de evaluación, pero también en la de formación de profesores.

La clasificación que se presenta en la Tabla 1, responde a las áreas que el C.I.D.E. utiliza habitualmente para presentar los resúmenes de las investigaciones en sus Catálogos de publicación bienal. En dicha Tabla se presentan las investigaciones educativas financiadas por el Ministerio de Educación a través del C.I.D.E. en el período 1982-1992.

TABLA 1**Investigaciones Financiadas por el M.E.C. clasificadas por áreas temáticas. Período 1982-1992**

| | 1982-84 | 1985-86 | 1987-88 | 1989-90 | 1991-92 ³ | TOTAL |
|--------------|-----------|-----------|-----------|-----------|----------------------|------------|
| 1 | 5 | 4 | 5 | 10 | 66 | 90 |
| 2 | 13 | 10 | 2 | 6 | 38 | 60 |
| 3 | 8 | 14 | 6 | 8 | 31 | 67 |
| 4 | 5 | 4 | — | 5 | 25 | 39 |
| 5 | 13 | 8 | 9 | 13 | 15 | 58 |
| 6 | 7 | 6 | 28 | 6 | 23 | 70 |
| 7 | 10 | 7 | 9 | 5 | 5 | 35 |
| 8 | 7 | 6 | 2 | 5 | 13 | 33 |
| 9 | 5 | 3 | 4 | 4 | 7 | 23 |
| 10 | — | — | 1 | 4 | 3 | 8 |
| TOTAL | 78 | 62 | 65 | 65 | 214 | 484 |

- 1.- Currículum. Programas y contenidos de la enseñanza.
- 2.- Profesorado. Formación de profesores. Profesión docente.
- 3.- Métodos de enseñanza. Medios pedagógicos.
- 4.- Sociología y educación.
- 5.- Psicología y educación.
- 6.- Rendimiento escolar. Evaluación.
- 7.- Administración educativa y planificación. Política y sistema educativo.
- 8.- Educación especial. Educación compensatoria.
- 9.- Formación profesional. Empleo. Orientación
- 10.- Otros temas.

Como se puede observar, el tema “Rendimiento escolar y Evaluación” ha sido objeto del mayor número de investigaciones en el período previo a la aprobación de la reforma educativa, es decir, entre 1987 y 1988 (la L.O.G.S.E. fue promulgada en 1990). Además, también se observa cómo en el último año (1991-92) del período estudiado, el segundo tema con más investigaciones financiadas —realizadas o en curso—, es también el área de “Evaluación”, precedida de la de “Currículum”.

³ Se incluyen investigaciones finalizadas y en curso.

Sin embargo, estos datos pueden no resultar claros para ponderar el peso de la evaluación de programas y centros en el conjunto de la investigación del C.I.D.E. Al comienzo del libro se refería el hecho de que, en este período, ha sido difícil encontrar un número significativo de investigaciones que se centraran en la evaluación de los aprendizajes de los alumnos. De hecho, cuando se evalúa el aprendizaje es como producto de alguna reforma, experimentación o innovación educativa concretas. Si a esto se añade que la mayoría de las evaluaciones de programas y centros incluidas en este libro han sido catalogadas atendiendo a su objeto de estudio (Formación de profesores, Administración educativa, Métodos de enseñanza, Currículum, etc.), parece claro que la forma más adecuada de apreciar el peso que ha tenido la evaluación de programas y centros educativos es realizar una estimación porcentual de la misma.

Así, en el período 1982-92, de un total de 484 investigaciones financiadas en las distintas convocatorias de Ayudas, Concursos y Premios, así como de las realizadas o encargadas por el propio C.I.D.E., 49 consisten en evaluaciones de programas y de instituciones o centros educativos, lo que supone algo más del 10% del conjunto de la investigación educativa financiada por el C.I.D.E. en este período.

Desde una perspectiva global, puede decirse que en estos 10 años las áreas temáticas de currículum y de evaluación (de programas, de centros y de rendimiento), seguidas de la de profesorado, han aglutinado un porcentaje muy importante de la investigación educativa financiada por el Ministerio de Educación y Ciencia.

CAPÍTULO IV

REVISIÓN DE LAS EVALUACIONES DE PROGRAMAS Y CENTROS EN EL PERÍODO 1982-1992

4.1. Criterios utilizados en la revisión y análisis de las investigaciones

Al revisar las investigaciones que se han incluido en esta obra, y dado que todas ellas consisten en evaluaciones de uno u otro tipo, se han tenido en cuenta una serie de criterios que se corresponden con una determinada tipología de la evaluación educativa.

En la determinación de los tipos de evaluación, al igual que ocurre con la definición del concepto mismo, se está aún lejos de lograr un mínimo consenso. Como ya señalara De la Orden (1985b), "*Desde 1975 hasta hoy se han publicado, al menos, ocho tipologías de evaluación diferentes, que responden a otros tantos criterios de clasificación, y arrojan un total de 47 categorías.*"⁴

En el caso de esta obra, la tipología utilizada no pretende en ningún caso ser exhaustiva; los criterios aplicados responden a algunos de los rasgos más reconocibles que pueden caracterizar a una evaluación en educación y que, en gran medida, responden también a los elementos que incorpora la definición de evaluación educativa que se acuñó al principio de este libro: *actividad o proceso sistemático e intencional de identificación, recogida y tratamiento de información relevante a sus objetivos que, tras ser contrastada con algún tipo de pa-*

⁴ DE LA ORDEN HOZ, A. (1985b): *Modelos de evaluación universitaria*. Revista Española de Pedagogía, n.º 169-170, julio-diciembre, pág. 531.

trón o criterio, permite realizar valoraciones que iluminen la toma de decisiones en educación.

Asimismo, es obligado señalar que no se ha pretendido proceder a una revisión científica de las evaluaciones; no consiste este trabajo pues en una metaevaluación de los estudios aquí reunidos. Tan sólo se trata de una aproximación descriptiva a las investigaciones, empleando unos criterios que se espera aportarán claridad e interés a esta revisión de diez años de evaluación de programas y centros financiada por el Centro de Investigación y Documentación Educativa (C.I.D.E.).

La estructura del Capítulo IV se ha realizado clasificando las investigaciones en tres grandes grupos, atendiendo a su objeto de estudio: *evaluación de programas, evaluación de centros, e investigaciones centradas en la elaboración de modelos para la evaluación de programas o centros*. Asimismo, en cada uno de estos tres grandes grupos las investigaciones se han agrupado por temas más específicos. Es en el marco de esta clasificación en el que se ha realizado la revisión y comentario de las evaluaciones, para lo que en cada grupo se han tenido en cuenta los siguientes criterios:

A. *Objeto de la evaluación*

Se ha tratado en este caso de identificar el tema o el ámbito concreto en que se han centrado los autores de cada evaluación: el tipo de programas (curriculares, didácticos, tecnológicos, de perfeccionamiento, etc.) o de centros (por niveles o modalidades educativas, por titularidad, ubicación, etc.).

En los casos en que ha resultado procedente, se comenta también el alcance de la evaluación, es decir, el número de centros, profesores, alumnos, etc., que conformaron las muestras (probabilísticas o no) de cada estudio.

B. *Características definatorias de la evaluación*

Se trata de un conjunto de criterios que, aunque no atañen ni al objeto ni al diseño metodológico de la evaluación, tienen un gran poder de caracterización de una evaluación educativa, por lo que se presentan y consideran conjuntamente. Por otra parte, como es bien sabido,

los criterios que se enunciarán a continuación no son excluyentes entre sí, lo que en ocasiones hace muy difícil y delicada la caracterización de una evaluación si sus autores no han explicitado sus intenciones o acciones en la memoria de investigación. No obstante, y siempre con la debida cautela, se ha procurado describir las evaluaciones en función de estos rasgos, por entender que, explicitadas o no, se trata de opciones que los evaluadores, de hecho, han realizado en algún momento de su trabajo.

b.1. El agente evaluador

En este caso, la atención se centra en el sujeto de la evaluación, en quienes la llevan a cabo, en función de lo que ésta habrá sido *interna o externa*. Evaluaciones internas se han considerado aquellas que han sido realizadas por los propios integrantes de un centro, por los participantes en un programa, etc. Pero, a su vez, las evaluaciones internas han podido adoptar diversas alternativas o direcciones de realización. En el caso en que los papeles de evaluador y evaluado coinciden en las mismas personas, estamos obviamente ante un proceso de *autoevaluación*. Cuando el programa o el centro educativo han sido evaluados por personas distintas a las que se hallaban afectadas por la evaluación, estaríamos describiendo un caso de *heteroevaluación* (por ejemplo, el Equipo Directivo que evalúa una propuesta del Claustro de profesores, o los profesores que evalúan el curso recibido en un Centro de Formación de Profesores o C.E.P.). Por último, cuando los receptores y los ejecutores de una intervención educativa han evaluado conjuntamente dicha intervención, estaremos ante una *coevaluación* (siguiendo el ejemplo anterior, sería el caso de los Asesores del C.E.P. y de los profesores asistentes que evalúan conjuntamente el curso de formación).

Sin embargo, encontraremos también con frecuencia en esta obra evaluaciones que tienen un carácter externo y que, por lo tanto, han sido realizadas sólo por expertos en evaluación pero ajenos en alguna medida al centro o al programa en cuestión.

Pero quizá el caso más frecuente en el grupo de investigaciones aquí reunidas, es el de la *evaluación interna con asesores externos*. Este caso es el más común cuando el propio equipo de investigación estaba compuesto por profesores universitarios y de otros niveles del

sistema educativo (Educación Infantil, Primaria, Secundaria, Educación de Adultos, etc.).

b.2. Objetivos y funciones de la evaluación

Los objetivos y funciones de la evaluación son aspectos fuertemente relacionados. En este caso se ha tratado de recoger, por un lado, aquellos objetivos que el propio equipo de investigadores ha expuesto en su diseño. Por otro, se ha atendido especialmente a la finalidad de la evaluación emprendida, es decir, a si la intención de sus autores era adoptar -o facilitar la adopción-, de decisiones de mejora de la situación estudiada (fuera un programa o un centro) o, por el contrario, se proponían decisiones finales de reafirmación, generalización, selección o eliminación de una determinada reforma, innovación o intervención educativa.

Empleando la tipología enunciada por Scriven (1967), en el primero de los casos estaríamos hablando de evaluaciones con *función formativa*, mientras que la función del segundo sería *sumativa*.

Se puede anticipar ya que la mayoría de las evaluaciones reunidas en esta obra tenían una clara finalidad formativa (de mejora, corrección, perfeccionamiento), lo que no obsta para que algunas hayan dado lugar, finalmente, a adoptar la decisión de generalizar o, por el contrario, de abandonar o sustituir por otro alternativo, un determinado programa, recurso didáctico, etc., al final de un período de aplicación, lo que les otorgaría un carácter también sumativo (de control o sanción en definitiva). En realidad, como afirma García Ramos (1989), en todo proceso global de evaluación resulta complejo separar con claridad la evaluación formativa de la sumativa.

b.3. Dimensiones evaluadas y momento de la evaluación

Aunque en ocasiones puedan confundirse o solaparse estos dos aspectos, la dimensionalidad y el momento de la evaluación pueden ser cosas bien diferentes en la práctica evaluativa.

Al analizar las dimensiones que ha contemplado una determinada evaluación, se ha atendido al tipo de variables o aspectos que los autores han incluido en el estudio: de entrada, de contexto, de proceso o de

producto. Responden a lo que Stufflebeam consideraría cuatro tipos de evaluación y que, de realizarse todas ellas en un mismo estudio, configurarían una evaluación global de un programa o de un centro educativo. Por lo tanto, atendiendo a su dimensionalidad, las evaluaciones han sido caracterizadas de *globales o parciales*, dependiendo de si se han considerado las cuatro dimensiones mencionadas, o tan sólo alguna o algunas de ellas.

Otra cuestión distinta es el momento en que se haya llevado a cabo la evaluación: al inicio (*evaluación inicial*), durante (*evaluación procesual*) o al final (*evaluación final*) de la puesta en marcha o de la creación de un programa o centro educativo.

En este sentido parecía interesante distinguir, por ejemplo, entre la evaluación de un programa que incluye variables de entrada, de contexto y de proceso pero que no recoge ese tipo de información hasta el final de un determinado período de implantación (caso de todos los diseños de tipo *ex-post-facto*), y aquella otra evaluación que recoge información sobre las entradas, el contexto y el proceso del programa, al inicio y durante la aplicación del mismo respectivamente. Ambos modelos son ciertamente válidos, pero no cabe duda que el alcance de las conclusiones que se obtengan puede ser, en muchos casos, distinto. Como es sabido, una de las diferencias más importantes atañe a las posibilidades que estos modelos ofrecen en relación a la toma de decisiones. El segundo modelo permitiría tomar decisiones sobre la marcha, lo que aumentaría su potencial formativo, mientras que las conclusiones del primer modelo de evaluación no pueden incidir sobre un proceso ya ocurrido, o sobre unos resultados ya obtenidos. En este último caso, las decisiones que se adopten —formativas o sumativas— afectarán necesariamente al futuro del programa o del centro en cuestión.

b.4. Patrón empleado para la emisión de juicios de valor

Si la fase central de toda evaluación es la valoración de la información que se ha obtenido, esta información debe haber sido interpretada previamente, para lo cual suelen emplearse distintas instancias de referencia. En caso de que la referencia sea el propio sujeto u objeto de la evaluación (por ejemplo, las características de partida de un programa, los logros educativos de un centro en el pasado, etc.), estaremos empleando la autorreferencia como sistema. En el caso de que las

referencias no sean el propio centro o programa, lo que se conoce como heterorreferencias, nos encontramos con dos posibilidades: la referencia normativa y la criterial. En palabras de De la Orden (1985b), referidas en este caso a la evaluación de alumnos, "*cuando la interpretación se hace en relación a las puntuaciones o calificaciones obtenidas por el grupo a que el sujeto pertenece (...) se habla de referencia normativa. Cuando la interpretación se hace con relación a un patrón de realización preestablecido, se habla de referencia criterial*"⁵.

Aunque la literatura especializada ha dedicado mayor atención a las referencias normativa y criterial en relación a la evaluación de los aprendizajes de los alumnos, éstas también se aplican con el mismo sentido a otros tipos de evaluación.

Así, en la evaluación referida a normas (normativa) la información recogida por el evaluador acerca de un centro o de un programa se habrá contrastado con la realización de un grupo normativo, es decir, con los resultados (o los promedios de los resultados) de otro/s programa/s u otro/s centro/s de la misma población o similares características, o bien con los resultados obtenidos en una prueba estandarizada. De hecho, la mayoría de las evaluaciones que cuentan con un diseño de tipo experimental con grupo de control, y en el que se utilizan pruebas o instrumentos estandarizados, suelen ser de tipo normativo. Así, en el supuesto caso de una evaluación comparada del clima institucional que generan dos modelos de organización de centros, como por ejemplo, una escuela incompleta tradicional y un Centro Rural Agrupado (C.R.A.), en la que se emplea una escala estandarizada para medir el clima escolar, la instancia de referencia empleada habría sido normativa.

Por su parte, en la evaluación referida a criterios (criterial), los resultados o logros de un programa o de un centro se contrastan con un patrón objetivo previamente establecido o definido por los evaluadores, o por los responsables del programa o del centro. En este caso, siempre que se han contrastado los resultados educativos de un programa o de un centro con los objetivos que éstos perseguían intencionadamente, se ha considerado que los investigadores han llevado a cabo una evaluación criterial.

Lo expuesto hasta aquí no quiere decir que no haya evaluaciones que utilizan varias instancias de referencia al mismo tiempo. En esta

⁵ DE LA ORDEN HOZ, A. (1985b), *Ibíd.*, pág. 529.

revisión se incluyen evaluaciones de centros y de programas que así lo han hecho, en su afán por interpretar y valorar de la forma más completa posible la información recogida.

C. *Diseño y desarrollo de la evaluación*

Se ha realizado también una descripción de las investigaciones teniendo en cuenta el diseño y metodología adoptados. Sin pretender abordar en profundidad estos aspectos, se ofrece una panorámica que permitirá apreciar qué enfoque, si el *cuantitativo-experimental-explicativo*, o el *cuantitativo-naturalista-comprensivo* ha predominado a la hora de llevar a cabo evaluaciones de programas y centros educativos. Asimismo, en algunos casos ha parecido relevante mostrar si los estudios se han realizado con muestras más o menos representativas de una determinada población de centros, alumnos, profesores, etc., o bien con muestras incidentales.

Asimismo, en este apartado se describen y comentan brevemente las fuentes de información empleadas por los investigadores (directas o indirectas), el tipo de información recogida (documental, personal, material, etc.), o los procedimientos e instrumentos más relevantes de recogida de información (observación directa, encuestas, aplicación de pruebas, análisis del discurso o de contenido, etc.).

Por otra parte, se ha considerado de gran interés el modo en que se han interpretado los resultados de las evaluaciones, pudiéndose proceder desde una simple *descripción* de los mismos, a una interpretación en términos de *comprensión*, de *explicación*, o de *valoración*. Ciertamente, el modo de concluir una evaluación, la forma en que se haya interpretado la información, puede llevar en ocasiones a reforzar el carácter evaluativo de un estudio o, por el contrario, acercarlo más a la investigación aplicada.

D. *Difusión y uso de los resultados de la evaluación*

En función de este criterio, se ha analizado si en las *memorias* de investigación realizadas por los autores se formalizan vías de presentación de los resultados de la evaluación a las audiencias interesadas, como pueden ser los habituales informes parciales y/o finales de eva-

luación, u otras formas más directas de informar y de presentar conclusiones, como la celebración de seminarios, reuniones, debates, etc..

Se ha pretendido identificar las audiencias a las que, en su caso, se han dirigido los resultados, por la relación que esta actividad tiene con el proceso de toma de decisiones posterior a la evaluación. En este sentido, los objetivos de cada evaluación tienen mucho que ver con la decisión de informar, o no, y a quién. Así, en el caso de una evaluación cuyo objetivo principal es generar teoría o evidencia empírica sobre alguna organización o intervención educativa innovadora, la audiencia a la que parecen dirigirse las conclusiones es, frecuentemente, la comunidad científica. Sin embargo, en el caso de las evaluaciones de marcado carácter aplicado y formativo, las situaciones que pueden encontrarse son más variadas. Por ejemplo, la audiencia potencial de las conclusiones de una evaluación realizada por investigadores particulares sobre un programa institucional, tendría como audiencia potencial a la propia Administración, que es quien puede adoptar decisiones de cambio, de mejora, etc. En otros casos, cuando las decisiones que derivarían de una evaluación pueden ser adoptadas por los propios responsables del programa o centro educativo en cuestión, es interesante saber si éstos han sido o van a ser informados expresamente de las conclusiones del estudio realizado, en el caso de no tratarse de una autoevaluación o de una coevaluación.

Así, el uso que se haya hecho o se vaya hacer de las conclusiones de una evaluación puede, o bien explicitarse en la memoria final del trabajo, o bien obviarse por distintas razones.

Como se dijo con anterioridad, el que se hayan adoptado o no las decisiones pertinentes va a depender, en gran medida, de que los evaluadores tengan capacidad de decisión o de intervención directa en el programa o centro evaluado; de lo contrario, las conclusiones se suelen exponer con dos tipos de intención: exponer y justificar adecuadamente el trabajo realizado, y poner a disposición de los implicados o interesados en dicha evaluación la información que les permitirá, si lo desean, adoptar las correspondientes decisiones, incluyendo, en algunos casos, recomendaciones en este sentido.

Por último, otra forma de difundir los resultados de un estudio evaluativo es que éste sea objeto de una publicación. En el caso de esta obra, tan sólo se ha consignado este hecho cuando ha sido el propio Servicio de Publicaciones del Ministerio de Educación y Ciencia, el C.I.D.E. en concreto, el editor.

E. *Metaevaluación*

Finalizada una evaluación, hay una fase ulterior que, discrecionalmente, sus autores pueden o no acometer: la metaevaluación o evaluación de la propia evaluación. Esta actividad, que puede realizarse con distintos grados de profundidad, de rigor o de sistematicidad, se ha contemplado también en la descripción de las investigaciones reunidas en esta obra como una vertiente más de la panorámica sobre las prácticas evaluativas que se pretende ofrecer.

Son pocas, aunque muy diversas, las situaciones encontradas a este respecto: desde metaevaluaciones previstas, planificadas y llevadas a cabo, a estudios que finalizan proyectando emprender a continuación su autoevaluación, o a evaluaciones que no hacen mención alguna a este proceso. Asimismo, en los casos en que se ha realizado la metaevaluación, el lector podrá observar cómo los procesos y metodologías han sido también muy diversos.

4.2. Evaluación de programas

Objeto y conceptualización de la evaluación de programas

Una caracterización básica de qué se entiende por evaluación de programas educativos puede pensarse que debe comenzar analizando, siquiera someramente, el significado y sentido del término "programa".

Este término es definido en el Diccionario Santillana de Ciencias de la Educación como "*proyecto que expone el conjunto de actuaciones que se desean emprender para alcanzar unos determinados y explícitos objetivos*"⁶.

Cuando el término programa es adjetivado de educativo, las definiciones que se ofrecen son de carácter general y comprensivo. Así Astin y Panos (1971) lo definen como "*...actividad educativa continua diseñada para producir cambios esperados en la conducta de los sujetos que se exponen a ella*"⁷. Por su parte, De la Orden (1985) lo ca-

⁶ Diccionario de Ciencias de la Educación (1983). Madrid, Santillana, pág. 1141.

⁷ ASTIN, A.W. y PANOS, R.J. (1971): "The evaluation of educational programs". En THORNDIKE, L. (Ed.): *Educational measurement*, 2.^a ed. Washington, American Council of Education, pág. 733.

racteriza como "...cualquier curso de acción sistemático para el logro de un objetivo o conjunto de objetivos⁸". Junto a estas y otras definiciones, se proponen ejemplos que recogen un amplio espectro de concepciones.

Para los autores citados, un programa educativo puede referirse a un método particular de interacción, al diseño instructivo específico para el aprendizaje de un tema concreto, a un plan de estudios para la enseñanza general, o a todo el sistema educativo en su conjunto. Desde esta perspectiva, la utilización de un texto, la introducción de la tecnología informática en el sistema educativo, la acción docente de un profesor, y hasta un centro docente concebido como organización funcional para la enseñanza y el aprendizaje, supondrían igualmente para muchos autores formalizaciones concretas de distintos tipos de programas educativos.

Los programas educativos, desde el enfoque predominante en el ámbito anglosajón (centrado en la evaluación de programas), son objeto de una caracterización sensiblemente distinta a la anterior, al aparecer ligados a proyectos institucionales que, si bien pueden poseer diferente envergadura o ámbito de aplicación, no se centran en lo que pudieran considerar niveles "micro" de la enseñanza.

Desde esta perspectiva, un programa educativo puede ser definido como un caso particular dentro del concepto más amplio y general de los denominados Programas de Acción Social (González Soler, 1989; Carballo Santaolalla; 1991, entre otros).

En esta misma línea puede situarse la conceptualización que realiza Alvaro Page (1993), cuando define un programa como la actividad o grupo de actividades que, mediante la aplicación de un plan, están encaminadas a proporcionar un servicio.

Cronbach y colaboradores (1980) proporcionan una concisa definición de programa, caracterizándolo básicamente como un plan o línea de acción permanente que proporciona servicios sociales.

A juicio de Worthen (1985), un programa en un contexto educativo se podría describir como una empresa o intervención cuyos objetivos fuesen la solución de un problema específico o la mejora de algún aspecto del sistema educativo. Generalmente, un programa de estas características estaría sostenido por fondos públicos o privados, ten-

⁸ DE LA ORDEN, A. (1985): "Investigación evaluativa". En DE LA ORDEN, A. (Dir.): *Investigación Educativa. Diccionario de Ciencias de la Educación*. Madrid, Anaya, pág. 133.

dría objetivos específicos, y dispondría de una estructura para la gestión de los procedimientos, el material, las instalaciones y el personal implicado, o de alguno de estos aspectos.

En la práctica, como señala Gairín (1992), un programa suele referirse a un conjunto de actuaciones, más o menos amplias pero coherentes entre sí, orientadas a una misma finalidad y claramente diferenciadas de otras. Por su claridad y concreción, esta es la conceptualización de programa que se ha adoptado en este libro para la inclusión, clasificación y descripción de las investigaciones evaluativas revisadas.

A tenor de las definiciones recogidas, han sido objeto de inclusión en esta obra los procesos evaluativos referidos a líneas de acción e intervención educativa que han sido previamente planificadas, y para las que se han determinado, por tanto, objetivos, actividades, medios y recursos, agentes implicados, etc. El alcance de las mismas puede ser notablemente diverso y, en términos generales, pueden incluirse desde esta óptica tanto los programas de amplia implantación territorial como otros programas específicos de intervención educativa para un centro, etapa, nivel o grupo determinado. En cualquier caso, cabría apuntar al menos tres consideraciones.

En primer lugar, en esta obra se distingue la evaluación de los cursos de acción planificada y orientada a la consecución de objetivos específicos, de aquélla referida fundamentalmente a los escenarios o contextos concretos donde éstas pueden tener lugar. De otro modo se estaría solapando con el ámbito de la evaluación de centros, lo que se hará en otro capítulo de esta obra.

En segundo lugar, se entiende el currículo como un complejo conjunto de programas y, por tanto, la evaluación curricular como un ámbito particular y especializado de la evaluación de programas. No obstante esto debe ser matizado, ya que se ha supuesto, al igual que hace Worthen (1985), que la evaluación del currículum puede calificarse más propiamente de evaluación de programas si ésta se enfoca hacia el cambio o la mejora.

En tercer lugar, y estrechamente relacionado con lo anterior, cabría distinguir el concepto de programa que se está perfilando, del curso ordinario y planificado de los acontecimientos educativos acaecidos en el contexto del aula. Es decir, que el diseño y desarrollo curricular en sus niveles últimos de concreción (programaciones de aula y adaptaciones curriculares) no se entienden estrictamente como objeto de la evaluación de programas. El análisis y evaluación de la in-

tervención educativa a nivel "micro" se suelen realizar desde la perspectiva estrictamente didáctica, mientras que desde la óptica de la evaluación de programas, el objeto de la evaluación lo constituyen más bien el diseño y desarrollo del currículo a niveles mayores de generalidad. Ello supone la existencia de una área de investigación con sus propias particularidades y desarrollo, y con una comunidad científica consolidada y diferenciable.

En su forma más simple, y sencillamente relacionando los conceptos anteriormente analizados, la evaluación de programas consiste en aquellas acciones que se emprenden para juzgar el valor o utilidad de un programa, o de programas alternativos, que trata de mejorar algún aspecto determinado de un sistema educativo (Worthen, 1985).

No obstante, sólo en este nivel de generalidad parece que pueda encontrarse consenso o claridad en la conceptualización. El amplio desarrollo de la evaluación de programas ha tenido por resultado el desarrollo de diversos marcos teóricos y prácticos. En consecuencia, la dificultad de conceptualización se aprecia en las múltiples definiciones con las que cada autor destaca, desde su perspectiva, los rasgos más relevantes.

De acuerdo con Tejedor y colaboradores (1993), las distintas propuestas pueden agruparse en dos tipos básicos de apreciaciones: las que hacen hincapié en la dimensión técnica y las que ponen el énfasis en los considerandos ideológicos. No obstante esto tampoco supone adoptar opciones realmente distintas o enfrentadas, sino focalizadas sobre distintos momentos y aspectos del proceso.

De este modo, De la Orden (1985) entiende que la evaluación de programas en Pedagogía puede definirse como *el proceso sistemático de recogida y análisis de información fiable y válida para tomar decisiones sobre un programa educativo*. Inmediatamente a continuación, apunta cómo el proceso se articula sobre la determinación del valor del programa de acuerdo con determinados criterios.

Para Cook y Shadish (1986) la evaluación de programas es un modelo racional de solución de problemas de carácter psicológico, social y educativo, que requiere que primero los problemas se definan con claridad y después se genere una amplia variedad de soluciones potenciales, algunas de las cuales se aplican y finalmente se evalúan. El conocimiento de las soluciones positivas se presenta a los responsables políticos, o a quienes toman las decisiones, quienes finalmente las utilizarán para adoptar las soluciones políticas que correspondan.

Por su parte, Talmage (1982) señala que, en la literatura especializada, a la evaluación de programas educativos o investigación evaluativa se le asignan como objetivos más frecuentes, los siguientes: proporcionar juicios acerca de la bondad del programa, asesorar a los decisores políticos responsables de las decisiones finales y cumplir una función política. Comprende, por otro lado, la totalidad de actividades de investigación evaluativa, desde las iniciales y referidas al diseño y la planificación, hasta la puesta en marcha y evaluación final del programa.

En todo caso, la conceptualización sobre evaluación de programas ha estado sujeta a la evolución del propio concepto de evaluación. Así, como señala Gairín (1992) *puede detectarse una amplia distancia entre la definición de Ruthman (la evaluación de programas conlleva el empleo de métodos científicos para medir la aplicación y resultados de programas y se orienta a la toma de decisiones) y la más actual ofrecida por Shadish, que entiende la evaluación como el conocimiento empírico del mérito o valor de un programa, con el fin de tomar decisiones.*

En definitiva, la evaluación de programas no constituye sino una ampliación del ámbito de la evaluación tal y como originalmente se concibió, que cobraría en los años 60 y 70 un auge inusitado, surgiendo una nueva modalidad de investigación aplicada que se ha venido denominando investigación evaluativa.

Tejedor (1993) sitúa la investigación evaluativa en lo que algunos han denominado “investigación orientada”, espacio de 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. De hecho, el objetivo de desarrollar guías prácticas para la toma de decisiones, y la exploración y juicio a cerca de la realidad con ánimo de actuar, son rasgos básicos de una práctica investigadora que suma a la investigación aplicada el juicio valorativo.

La investigación evaluativa, pues, se caracteriza no sólo por la inmediatez de su utilidad práctica, común a todo proyecto de investigación aplicada, sino esencial y sustancialmente por ser un proceso que alberga en todas sus fases cuestiones de valor, decisiones valorativas.

4.2.1. Evaluación de reformas y de programas experimentales

— Objeto de la evaluación

Para definir adecuadamente el objeto de las evaluaciones incluidas en este primer grupo, es necesario contextualizar los programas

evaluados dentro de los procesos experimentales de reforma e innovación llevados a cabo en nuestro país en momentos previos a la reforma del sistema educativo establecida por la Ley de Ordenación General del Sistema Educativo (L.O.G.S.E.).

Aguado Odina (1991), apoyándose en la clasificación que realiza Havelock (1973) de los procesos de cambio en los sistemas educativos, afirma que el modelo adoptado por el Ministerio de Educación y Ciencia español para llevar a cabo la reforma responde, en líneas generales, al denominado de *investigación/desarrollo/difusión*, en el que se utilizan teorías y datos básicos para generar ideas, que posteriormente se convierten en prototipos que se someten a prueba, tras lo que son rediseñados y vueltos a comprobar. Cabe considerar que este es el proceso seguido por los siguientes programas institucionales:

- El Plan para la reforma de la Educación Infantil, iniciado por el Ministerio de Educación y Ciencia durante el curso 1985-86.
- La reformulación de las enseñanzas de los Ciclos Inicial y Medio de la Educación General Básica, puesta en práctica con carácter experimental en el curso 1985-86 en diez centros del ámbito territorial gestionado por el Ministerio de Educación y Ciencia. Además, a partir del curso 1984-85 se emprendió el proceso experimental de reforma del Ciclo Superior en 50 centros, ampliándose a más centros en cursos sucesivos.
- La experimentación de nuevos planes y programas de Enseñanzas Medias a partir del curso 1983-84.
- La experimentación de los Módulos Profesionales en algunos centros, a partir del año 1986. Estos Módulos constituyeron la base para el diseño y organización de la actual Formación Profesional Específica, tanto de grado medio como de grado superior.

En esta obra se incluyen tres trabajos evaluativos que tienen alguno de estos procesos experimentales como objeto de estudio o como referente.

Concretamente se trata de la *Evaluación de programas pertenecientes al Plan Experimental para la Reforma de la Educación Infantil* (1991), de Teresa Aguado Odina; y de dos trabajos sobre planes ex-

perimentales de la Educación Secundaria llevados a cabo por profesionales del C.I.D.E. bajo la dirección de Mariano Alvaro Page: *Evaluación Externa de la Reforma Experimental de las Enseñanzas Medias* (1992), y *Evaluación de los Módulos Profesionales: estudio de la Reforma Experimental de la Enseñanza Técnico-Profesional* (1993). Se trata, en los tres casos, de evaluaciones cuyo objeto de valoración son programas curriculares de amplio alcance.

Para este conjunto de estudios puede resultar interesante e ilustrativo señalar el alcance del proceso evaluativo, esto es, su ámbito de aplicación, en la medida que se trata de planes de amplia implantación territorial. En este sentido cabe señalar el carácter más restringido del estudio de Aguado Odina, limitado a la provincia de Zaragoza, frente a la amplitud del ámbito de aplicación de las evaluaciones realizadas por el C.I.D.E.: ámbito estatal en la evaluación de las reformas experimentales de las Enseñanzas Medias y ámbito de gestión del Ministerio de Educación y Ciencia en la de la Enseñanza Técnico-Profesional.

De modo paralelo a los procesos experimentales citados anteriormente, y a lo largo de la década de los ochenta, se pusieron en marcha programas dirigidos a introducir en la escuela aspectos que se consideraban necesarios para mejorar la calidad del sistema educativo y que, de nuevo, deben ser entendidos como iniciativas inscritas en el marco global de la reforma educativa emprendida.

Entre otros programas, se iniciaron de modo experimental los siguientes: *Integración de Alumnos con Necesidades Educativas Especiales*, *Educación Compensatoria*, *Proyectos Atenea y Mercurio*, *Orientación*, *Proyectos de Formación en Centros*, *Programa Prensa-Escuela*, *Educación para la Salud*, *Prevención de Drogodependencias*, etc.

En relación con la experimentación de algunos de los programas mencionados, se llevaron a cabo una serie de investigaciones: *Evaluación del Programa de Integración Escolar de Alumnos con Deficiencias* (1989) y la *Evaluación del Programa Prensa-Escuela* (1989), siendo esta última una investigación contratada directamente por el C.I.D.E. y llevada a cabo por un equipo compuesto por Mar Domínguez Girón, Ana Méndez Miras y Javier Villalaín Benito. Se trata de dos estudios evaluativos sobre unos programas de alcance y ámbito de aplicación muy distintos, siendo el segundo mucho más restringido.

El *Programa de Integración* escolar de alumnos con deficiencias comienza a implantarse en el curso 1985/86, a partir de la aprobación del Real Decreto de Ordenación de la Educación Especial. Dadas sus

características y su carácter experimental, el programa se concibió inicialmente como abierto, ya que cabía esperar que se concretara de muy diversas formas en los centros que se adscribieran al mismo. Durante el período de experimentación, el Programa de Integración operó exclusivamente en los niveles educativos de Preescolar, Ciclo Inicial y Ciclo Medio de Educación General Básica.

El C.I.D.E. financió por espacio de cuatro años un proceso de evaluación sistemático de la implantación de este programa, que fue llevado a cabo por un equipo de técnicos del propio C.I.D.E. y del Centro Nacional de Recursos para la Educación Especial. Este equipo estuvo compuesto por distintas personas a lo largo de sus diferentes etapas, siendo autores del informe M.^º Jesús Aguilera, Katia Alvarez, Mercedes Babio, César Coll, Gerardo Echeita, Marisa Galán, Alvaro Marchesi, Elena Martín y Rosario Martínez Arias (Equipo de Integración, E.P.I.).

Por lo que se refiere al ámbito del estudio, del total de 173 centros implicados en el programa en el curso 1985/86, se seleccionó una muestra aleatoria de 60 centros para realizar el seguimiento.

Por su parte, el *Programa Prensa-Escuela* surge con carácter experimental en el curso 1985/86 impulsado por la firma de un convenio de colaboración entre el Ministerio de Educación y Ciencia y la Asociación de Editores de Diarios Españoles (A.E.D.E.). A través del citado acuerdo, ambas partes se comprometen a desarrollar una serie de iniciativas encaminadas a favorecer y fomentar el uso de la prensa escrita como recurso didáctico en las aulas.

El ámbito de aplicación del Programa Prensa-Escuela lo constituyeron inicialmente los centros educativos públicos y privados que impartieran enseñanzas de E.G.B., B.U.P., F.P., Educación Compensatoria y Educación Permanente de Adultos.

La evaluación del Programa se hizo en el mismo ámbito de su implantación, es decir, en centros situados en el territorio gestionado por el Ministerio de Educación y Ciencia, aunque se excluyeron del estudio las provincias de Ceuta y Melilla y los municipios con menos de 2.000 habitantes. Los tamaños muestrales empleados en la evaluación son de 497 centros (sobre un universo de 8.233) y de 2.043 profesores (sobre un universo calculado de 133.476).

Otros programas experimentales citados anteriormente también han sido evaluados de modo sistemático, aunque no se incluyen en esta obra al no haber sido financiadas dichas evaluaciones por el

C.I.D.E. Este es el caso, por ejemplo, de la evaluación del *Proyecto Atenea* (para la introducción de los ordenadores en los centros educativos), puesto en marcha a partir de 1985 por el Programa de Nuevas Tecnologías de la Información y de la Comunicación (P.N.T.I.C.). Para evaluar los resultados de la fase experimental se emplearon tres líneas concretas de evaluación global del Programa, encabezadas respectivamente por una institución universitaria —el Instituto de Ciencias de la Educación de la Universidad de Murcia—, un organismo internacional —la Organización para la Cooperación y el Desarrollo Económico, O.C.D.E.—, y el propio Programa de Nuevas Tecnologías.

La evaluación de otros programas puestos en marcha en esta misma época, como los realizados en el marco de los Proyectos de Formación en centros, se tratarán en el epígrafe correspondiente a la evaluación de programas de formación del profesorado.

Por último, en este grupo de evaluación de reformas y de programas experimentales se incluye una investigación, actualmente en curso, dirigida por Juan Manuel Escudero Muñoz bajo el título *La disseminación y el desarrollo de la L.O.G.S.E. como reforma educativa: un estudio evaluativo*. En este caso, el programa objeto de evaluación lo constituye la propia L.O.G.S.E., entendida como un complejo conjunto de programas educativos que, a juicio del autor, “tiende a promover cambios en la ordenación del sistema educativo, en el currículo, en la enseñanza, en el papel de los centros y de los profesores, así como en la definición de ciertas funciones pedagógicas”.

El estudio se limita a la región de Murcia y a las etapas de Educación Infantil y Primaria. La población, en consecuencia, está representada por los centros y profesores implicados en la implantación de la L.O.G.S.E. para estos niveles durante el período de estudio.

— *Características definitorias de la evaluación*

Agentes evaluadores

En el epígrafe anterior ya se apunta una caracterización de los agentes responsables de los distintos estudios evaluativos, aspecto que se retoma ahora con el fin de caracterizar mejor las distintas investigaciones.

Cabría considerar como evaluaciones netamente externas los trabajos de Aguado Odina sobre la reforma de la Educación Infantil, y el dirigido por Escudero Muñoz, centrado en la disseminación y desarrollo de la L.O.G.S.E., ya que como es lógico dichos autores no tuvieron

responsabilidad directa sobre el diseño y desarrollo de los programas evaluados.

Por su parte, la Evaluación de la Reforma Experimental de las Enseñanzas Medias, de los Módulos Profesionales y aún del Programa Prensa-Escuela, son estudios dirigidos y realizados por o desde un organismo, el C.I.D.E., que forma parte del Ministerio de Educación y Ciencia, responsable directo del diseño y experimentación de los tres programas. No obstante puede hablarse de una evaluación externa, no sólo porque la evaluación no la realizaron los propios centros implicados, sino porque fue llevada a cabo por técnicos y expertos en evaluación, y no por los responsables políticos o administradores de dichos programas experimentales.

Por el contrario, en sentido estricto no puede hablarse de una evaluación externa en el caso del Programa de Integración. En esta ocasión los diseñadores y, en gran medida, responsables del programa, fueron también los evaluadores de su implantación experimental, en estrecha colaboración con los profesionales de los centros de integración. Este tipo de evaluación presenta como ventaja fundamental el profundo conocimiento del programa y de sus condiciones de aplicación por parte de los evaluadores, lo que sin duda repercutió tanto en el diseño como en las conclusiones de la propia evaluación.

Los estudios evaluativos que aquí se revisan consistieron en heteroevaluaciones llevadas a cabo por técnicos con experiencia en evaluación, junto a expertos en los contenidos y niveles educativos en los que se centraban los programas evaluados.

Solo en un estudio los autores afirman haber emprendido un proceso de autoevaluación, y es el caso de la Evaluación del Programa de Integración Escolar. Sin embargo, dicho proceso se realizó en la fase de intervención y estuvo referido al trabajo de los propios evaluadores, no al programa o a los sujetos evaluados. Por lo tanto, y en la medida en que este proceso es llevado a cabo por el propio equipo de intervención/evaluación, y para su propia autocrítica, supondría un procedimiento más próximo a la metaevaluación que a la autoevaluación.

Objetivos y funciones de la evaluación

Todos los estudios incluidos en este grupo tienen como telón de fondo las necesidades nacidas de la política educativa de la reforma: necesidades de organización, de planificación, pero también, y previa

a las anteriores, necesidad de control y análisis de los procesos de innovación y cambio emprendidos, y aún de debate acerca de los mismos.

Desde esta perspectiva, las evaluaciones auspiciadas por el Ministerio de Educación, de un modo u otro estaban necesariamente orientadas a la toma de decisiones sobre las reformas que finalmente debían adecuarse, aprobarse e implantarse, por lo que el objetivo fundamental era obtener información sistemática y fiable.

Esta orientación a la toma de decisiones aparece explícitamente recogida en el segundo Informe de la Evaluación Externa de la Reforma Experimental de las Enseñanzas Medias. En dicho Informe se afirma que su aportación fundamental consistía en ofrecer datos suficientes, sistemáticos y rigurosos, para que los responsables de la política educativa pudieran valorar el funcionamiento de la reforma experimental, así como los efectos que ésta estaba produciendo en comparación con la enseñanza vigente en ese momento.

Esta misma finalidad, unida a un enfoque fundamentalmente formativo, preside los planteamientos del resto de las evaluaciones.

Así, el objetivo general del trabajo realizado sobre la Reforma Experimental de las Enseñanzas Medias, al margen de otros objetivos parciales y colaterales, estaba constituido por la determinación del mayor o menor rendimiento obtenido en unas determinadas materias por los alumnos escolarizados en el Plan experimental, con respecto a los alumnos de la enseñanza vigente en ese momento.

El otro estudio evaluativo del C.I.D.E. dirigido por Alvaro Page, y centrado en los Módulos Profesionales, tenía dos objetivos fundamentales.

En primer lugar, trataba de recoger la situación que en ese momento presentaban los Módulos Profesionales desde la óptica de los distintos agentes implicados. En definitiva, se proponía perfilar un cuadro descriptivo que permitiera detectar y solucionar las deficiencias encontradas, antes de decidir la implantación o no del programa tal y como estaba diseñado.

En segundo lugar, se deseaba hacer una comparación entre el sistema de Módulos y la vigente Formación Profesional, adoptando un sistema de comparación ya aplicado en la evaluación externa de las Enseñanzas Medias.

Asimismo, el equipo de evaluadores del Programa Prensa-Escuela se propuso determinar, tras un período de experimentación del mismo, los avances logrados en la divulgación y aprovechamiento del Programa por parte de los centros y de los profesores implicados.

Como objetivos genéricos del estudio se proponían los siguientes:

- Conocer la proporción de centros escolares públicos y privados que en aquel momento estaban suscritos al programa Prensa-Escuela.
- Conocer los profesores que hacían uso de la prensa en sus clases.
- Obtener un conocimiento aproximado del uso que se hacía de la prensa en las aulas y de la respuesta de los alumnos a dicha utilización.
- Averiguar el grado de conocimiento que los docentes tenían del Programa Prensa-Escuela, y el tipo de prestaciones del Programa que utilizaban con más frecuencia.

Por su parte, la evaluación del Programa de Integración tenía una doble finalidad: de un lado, la determinación de los logros del programa, relativos tanto a la promoción y desarrollo de los niños integrados, como a la introducción de mejoras en el sistema educativo; de otro, hacer posible un proceso que aportase información y orientaciones a partir de la cuales “modelar” progresivamente el Programa y su puesta en práctica.

Así, los autores se propusieron como objetivos generales de la evaluación del Programa los siguientes:

- Estudiar cómo se estaba concretando el Programa en la realidad.
- Analizar el impacto de la aplicación del Programa sobre el entorno de aprendizaje en el que se llevaba a cabo.
- Analizar el impacto de la aplicación del Programa en los niños y niñas integrados.
- Proponer estrategias de intervención para crear las condiciones que hicieran posible la puesta en marcha de la modalidad o modalidades de integración más adecuadas, en función de las características y demandas de cada centro.

Los autores se refieren a la finalidad iluminativa de esta evaluación, al estar destinada a conocer el funcionamiento del Programa de Integración, cómo influían en éste las distintas situaciones escolares en que se estaba aplicando, qué ventajas y desventajas encontraban en

él las personas directamente interesadas, o cuáles eran las principales críticas que estaba suscitando su implantación. Así, se trata de una evaluación que no tenía un mero carácter descriptivo de la puesta en marcha del Programa (*¿qué estaba pasando?*) sino también explicativo (*¿por qué estaba pasando?*).

Las dos características atribuidas a los estudios evaluativos anteriores —orientados a la toma de decisiones y con carácter formativo— son también aplicables al trabajo de Aguado Odina, consistente en la evaluación del Plan experimental de Educación Infantil.

Como objetivos generales de este estudio, se mencionan la valoración de las condiciones de aplicación del Plan y de las características de los proyectos desarrollados dentro del mismo, así como de los resultados conseguidos durante la fase de experimentación. Más concretamente, los objetivos fijados, al margen ahora de los de carácter instrumental, son los siguientes:

- Analizar y valorar comparativamente los diferentes proyectos de Educación Infantil desarrollados en el marco del Plan Experimental para la Reforma de la Educación de esta Etapa, y los proyectos no pertenecientes a dicho Plan, en orden a identificar las variables de entrada y de proceso que diferenciaban a ambos, y que se podían relacionar con los resultados obtenidos por los alumnos.
- Analizar los efectos diferenciales de los programas cuando se adopta como criterio el grado de consecución de los objetivos generales establecidos en la Reforma de la Educación Infantil, en función de distintas variables relevantes.

Además, en el estudio se presentan dos objetivos complementarios, calificados expresamente como objetivos propios de las evaluaciones formativas orientadas a la toma de decisiones:

- Proporcionar datos acerca de la situación en que se encontraba el proceso de reforma, y sobre las características estructurales y procesuales que lo diferenciaban de otras propuestas más tradicionales.
- Mejorar el modelo educativo y los proyectos desarrollados, atendiendo a las variables (estructurales y procesuales) identificadas como más influyentes en los resultados obtenidos por los alumnos.

Finalmente, la caracterización general del diseño de evaluación del trabajo dirigido por Escudero Muñoz (*La diseminación y el desarrollo de la L.O.G.S.E. como reforma educativa: un estudio evaluativo*), se corresponde con lo que suele denominarse evaluación del proceso, de puesta en marcha de un programa.

El propósito general de esta investigación es la descripción y comprensión de un conjunto de fenómenos relacionados con la diseminación, implantación y desarrollo inicial de la L.O.G.S.E. Desde el modelo adoptado, se asume que la difusión, diseminación y adopción inicial de una reforma, constituye, de hecho, una fase inicial de la puesta en práctica de la misma.

Recogiendo nuevamente una orientación formativa y orientada a la toma de decisiones, entre los objetivos principales de esta evaluación se señalan los siguientes: proponer líneas de actuación para establecer una política de apoyo al desarrollo progresivo de la L.O.G.S.E. en la región murciana en lo que atañe a ciertos factores socioadministrativos y de asesoramiento, así como a otros factores más directamente relacionados con el funcionamiento interno de los centros y con el quehacer pedagógico de los profesores.

Dimensiones consideradas y momento de la evaluación

De los estudios revisados, al menos tres emprenden la recogida de datos del punto de partida (entradas), de los procesos y de los productos del programa evaluado. No obstante guardan algunas diferencias significativas entre sí.

El trabajo de Aguado Odina sobre la reforma de la Educación Infantil recoge, como datos de entrada, variables tanto personales (relativas a los alumnos, a la familia y a los profesores) como contextuales (características estructurales y dinámicas del centro educativo). La evaluación de la dimensión de proceso incluye variables relativas al ámbito familiar (estilo educativo, actitud básica, relación con el centro, etc.) y escolar (modelo educativo, objetivos, adecuación al Diseño Curricular Base, etc.). Finalmente como variables criterio (de producto) se agrupan informaciones procedentes de los ámbitos cognoscitivo-verbal, afectivo y psicomotor de los alumnos. Lógicamente, estas últimas variables, dependientes de acuerdo con su función dentro del diseño, están referidas en todos los casos a los alumnos que siguieron los programas.

Este trabajo de evaluación se desarrolló durante los dos últimos cursos de la etapa de experimentación del Plan (1988/90), previa a la generalización de los programas experimentales. Debe entenderse pues, como una evaluación final. Parece diseñarse así, una evaluación primeramente centrada en los productos y articulada a partir de la adopción de un diseño clásico de investigación, como se verá más adelante. En este modelo, a partir de las variables de entrada se determinan asociaciones causales con los resultados de los programas, a pesar de que dichas variables sirvan a una descripción/explicación de los programas evaluados desde perspectivas distintas a sus productos.

También el estudio dirigido por Alvaro Page en torno a la Evaluación de la Reforma Experimental de las Enseñanzas Medias comparte algunas de las características ya mencionadas. En primer lugar, se trata de una evaluación final (sin evaluación inicial propiamente dicha), llevada a cabo tras unos años de desarrollo del programa. Igualmente se incluye información de los grupos de variables consideradas de entrada (personales, familiares, actitudinales) y contextuales y se estudia su influencia sobre los productos. Además se incluye alguna variable de las catalogadas como de proceso —las variables denominadas didácticas en el estudio—, pero medidas indirectamente mediante las valoraciones emitidas por los propios alumnos acerca de la dinámica de la clase. Por ello, parece que la caracterización de este diseño como de proceso-producto debería hacerse con prudencia.

Los productos educativos en los que se centra el estudio, son los siguientes: rendimiento, aptitudes, expectativas académico-profesionales, actitudes cívico-sociales, reacciones emocionales ante el estudio e informaciones relativas a componentes didácticos. Expresado de una forma más sistemática, las variables empleadas fueron:

- Como variable independiente: tipo de enseñanza o grupo.
- Como variables moduladoras o de control: datos personales, familiares, contextuales y relativos a la trayectoria académica del alumno.
- Como variables dependientes: rendimiento, tasas de repetición y de abandono.
- Variables con doble función (dependientes y de control): aptitudes, actitudes cívico-sociales, reacciones emocionales ante el estudio, variables didácticas o pedagógicas y aspiraciones profesionales y educativas de los alumnos.

Por su parte, en la evaluación del Programa de Integración, aunque también se consideran dimensiones de entrada, proceso y producto, destacan algunas características que hacen de ésta una evaluación notablemente distinta a las dos anteriores.

La investigación se desarrolla durante los cuatro primeros años de implantación del Programa, a partir de septiembre de 1985.

Dadas las características ya mencionadas del Programa —su carácter abierto y experimental— se pretendió analizar distintas formas de llevar a cabo el proceso de integración de los alumnos. Quiere decirse que el proceso de evaluación discurrió simultáneamente al proceso de puesta en marcha del programa. En definitiva, se trata de una evaluación de las condiciones iniciales, que se mantiene durante el desarrollo y el final de una etapa de funcionamiento del programa. Por lo tanto, en este caso se puede hablar con toda propiedad de una evaluación global del Programa, es decir, de sus dimensiones de entrada, de contexto, de proceso, y de los resultados obtenidos.

Las dimensiones evaluadas, denominadas por los autores “sistemas”, fueron concretamente tres: el centro, el profesor y los alumnos, tanto integrados como pares.

Para la primera fase del estudio, y para el sistema centro, el tipo de variables consideradas fueron: las características generales del centro (rural-urbano, público-privado, número de unidades, etc.), experiencia previa en integración, grado de acuerdo para inscribirse en el Programa, proyecto educativo del centro, y recursos humanos y materiales.

Por lo que se refiere al profesorado, se recogió información relativa a su formación, experiencia, edad, etc.; a sus actitudes hacia la integración; a los procesos instructivos o estilo de enseñanza, y a sus opiniones pedagógicas.

Finalmente, y en relación a los niños integrados, se obtuvieron datos de edad, curso, necesidades especiales, condiciones familiares, modalidad de integración, variables aptitudinales y de desarrollo cognitivo y social, y de adquisición y rendimientos escolares. Asimismo se analizaron las actitudes de sus padres y su historia educativa y escolar. En lo que respecta a los niños pares las variables consideradas fueron similares.

En la fase más cualitativa de la evaluación, se incluyeron variables de naturaleza diversa, pertenecientes a los siguientes “sistemas” o dimensiones: centro, ambiente inmediato de aprendizaje, profesor, necesidades educativas de los alumnos, modalidad de integración y familia.

Por otro lado, la evaluación de los Módulos Profesionales (C.I.D.E.) tiene unas características distintas y se realizó de acuerdo a unos objetivos también diferentes.

Dado que se trataba de conocer cómo eran percibidos los Módulos Profesionales por los agentes implicados, qué grado de satisfacción mostraban éstos, cuáles eran las ventajas reales que ofrecían los Módulos, así como de proceder a su comparación con la enseñanza vigente, las dimensiones y variables consideradas se corresponden en todos los casos con tres grandes ámbitos de evaluación: alumnos, profesores y empresas vinculados a los dos programas de enseñanza (el experimental y el vigente).

Para evaluar los Módulos Profesionales se operativizaron un conjunto de dimensiones relativas tanto a los agentes implicados (incluyéndose variables personales, familiares, contextuales, etc.), como a diversos aspectos del programa propiamente dicho, ya fueran estrictamente curriculares o de relación, dotación, organización, planificación, etc.

Al margen de los datos empleados para caracterizar a las muestras, algunas de las dimensiones estrictamente referidas al programa, que se emplearon para la comparación de los dos tipos de enseñanza, fueron las siguientes: metodología didáctica, evaluación de prácticas, valoración de la formación recibida, información acerca de los Módulos, evaluación de la prueba de acceso, evaluación del área del Proyecto, evaluación de las clases de expertos, intervención de docentes y empresas en la organización de las prácticas, información sobre el plan de estudios en la empresa y sobre la aceptación de alumnos en prácticas, opinión de las empresas sobre la preparación mostrada por los alumnos de Módulos y por los de Formación Profesional en la fase de prácticas, etc.

Por otro lado, y puesto que el estudio se inició en 1989 —transcurridos tres años desde la implantación experimental del programa—, se trata de una evaluación con algunos aspectos de evaluación procesual, pero fundamentalmente final. Al mismo tiempo, la comparación entre un programa establecido y uno de carácter experimental tenía naturalmente como objetivo principal la evaluación de resultados.

El estudio, pues, parece centrado en los resultados arrojados por el programa modular en este período; resultados no vinculados en este caso al ámbito cognoscitivo o del rendimiento académico, sino al de las actitudes y valoraciones subjetivas de los sujetos implicados en el desarrollo del programa evaluado.

Por su parte, la evaluación del Programa Prensa-Escuela, cuyo objetivo fundamental era la apreciación de cómo se estaba llevando a cabo dicho Programa en los centros, se centró en los siguientes factores:

- Recursos de prensa escrita disponibles en los centros y vías de adquisición.
- Perfil del profesor como lector de prensa.
- Conocimiento del Programa Prensa-Escuela en los centros.
- Utilización de la prensa en el aula (cuál, con qué frecuencia, para qué materias, con qué función, etc.).
- Interés y participación de los alumnos en las clases cuando se empleaba este recurso didáctico.

Estos factores podrían quizá ser clasificados, de acuerdo con los criterios clásicos, como dimensiones de entrada, de proceso y de producto. En todo caso, es preciso apuntar la naturaleza indirecta de los datos recabados, y el carácter final que cabe atribuir a una evaluación llevada a cabo en un solo momento y tras el período de experimentación del programa.

Por último, y por lo que se refiere al estudio sobre la diseminación y el desarrollo de la L.O.G.S.E. (Escudero Muñoz), cuyo objetivo es evaluar el proceso de puesta en marcha del complejo y amplio programa objeto de análisis, los principales factores considerados en el diseño son los siguientes:

- Vías de acceso a la información sobre las nuevas propuestas de la L.O.G.S.E.
- Formación y tipo de asesoramiento y apoyo recibido en los procesos de contacto y desarrollo inicial de la reforma en las aulas.
- Elaboración del Proyecto Curricular y del Proyecto Educativo en los centros.
- Preparación y planificación del curso.
- Elaboración de Programaciones cortas.
- Coordinación entre los docentes del centro.
- Tipo de tareas realizadas en las aulas.
- Estructura del tipo de actividades creadas por los profesores y alumnos.
- Valoraciones generales del profesorado acerca de la L.O.G.S.E., del tipo de formación recibida, de las implica-

ciones para su centro, del desarrollo de capacidades para su puesta en marcha, etc.

Cabe apuntar que la evaluación esta diseñada para llevar a cabo el seguimiento del proceso de implantación de la reforma en curso, proponiéndose recoger y valorar la información de modo simultáneo al mismo (evaluación de proceso).

Patrones empleados para la emisión de juicios de valor

Salvo el trabajo de Aguado Odina (Evaluación del Plan Experimental para la reforma de la Educación Infantil) y la evaluación del Programa de Integración (VV.AA.), en los estudios evaluativos incluidos en este grupo no se formulan explícitamente juicios de valor acerca de los programas evaluados. Por otro lado, todos los estudios adoptan una perspectiva comparativa —del programa de reforma en fase experimental con el programa de enseñanza vigente—, por lo que pueden ser consideradas evaluaciones normativas, lo que ocurre también si consideramos los instrumentos utilizados. Los tests de referencia criterial son escasísimamente usados, quizá por la dificultad y el coste que entraña su elaboración *ad hoc*.

— Diseño y desarrollo de la evaluación

El modelo de evaluación adoptado en estas investigaciones no siempre es explicitado por los autores en sus estudios. No obstante, el análisis de los diseños permite reconocer en ellos determinados enfoques.

Al menos tres de los estudios agrupados como evaluaciones de reformas o programas experimentales adoptan marcos de investigación evaluativa próximos a los modelos basados en la experimentación, incluyendo la comparación con otro programa. Es el caso de los dos estudios del C.I.D.E. dirigidos por Alvaro Page —Evaluación externa de la Reforma Experimental de las Enseñanzas Medias y Evaluación de los Módulos Profesionales— y la investigación de Aguado Odina sobre el Plan Experimental para la Reforma de la Educación Infantil.

El primero de los estudios citados, sobre las Enseñanzas Medias, adopta un enfoque eminentemente cuantitativo (como suele ocurrir

con todas las evaluaciones realizadas desde el modelo experimental), a partir de un diseño calificado de cuasi-experimental y de carácter longitudinal.

Los autores señalan cómo el modelo al que tienden en su estudio sigue el patrón del modelo input-output, pero también el de proceso-producto, al incluir alguna variable de las catalogadas como didácticas, referida al marco del aula. Igualmente puntualizan que la calificación del modelo como de proceso-producto ha de hacerse con ciertas reservas, puesto que sólo la percepción por parte de los alumnos del método utilizado por el profesor en clase hace referencia al proceso.

En el estudio se ha trabajado con tres grupos claramente definidos:

- a) Experimental, compuesto por sujetos que cursaban las enseñanzas de reforma.
- b) Control interno, configurado por sujetos que cursaban las enseñanzas entonces vigentes, pero en centros experimentales de la reforma.
- c) Control externo, constituido por alumnos que cursaban también las enseñanzas vigentes, pero en centros no experimentales.

En este trabajo no se han llevado a cabo selecciones ni asignaciones totalmente aleatorias de los sujetos. Sin embargo, mediante la utilización de procedimientos estadísticos se lleva a cabo una equiparación *a posteriori* entre los grupos experimental y de control, por lo que se refiere a todas las variables excepto la experimental, lo que se entiende aproxima el diseño a los diseños experimentales.

Este estudio se ha caracterizado también como longitudinal, pues se ha hecho un seguimiento de los alumnos participantes aplicándoles pruebas en diferentes momentos de su vida académica. En concreto se ha trabajado con dos cohortes de alumnos, que comenzaron las Enseñanzas Medias en los cursos 1984-85 y 1985-86, respectivamente.

Las muestras empleadas (representativas y suficientes) dan idea de las dimensiones de este estudio. Así, por ejemplo, para la "Generación III", evaluada al final del primer y segundo ciclos, y considerando los tamaños muestrales de ambos niveles conjuntamente, el grupo experimental constaba de 3.030 sujetos; el de control interno, de 2.910, y el de control externo, de 8.538.

Los procedimientos empleados para la recogida de información incluyeron tanto pruebas estandarizadas existentes en el mercado, como pruebas elaboradas específicamente y validadas como parte del propio estudio evaluativo (cuestionarios, pruebas de aptitudes y pruebas de rendimiento en algunas materias fundamentales).

La evaluación de los programas pertenecientes al Plan Experimental para la Reforma de la Educación Infantil, llevada a cabo por Aguado Odina, presenta, desde el punto de vista del diseño y del modelo, las mismas características que el anterior.

Se trata de un diseño cuasi-experimental, jerárquico de seis grupos intratamientos, transversal (se toman medidas en momentos dados del proceso) y longitudinal (se realizan medidas en seis momentos diferentes a lo largo de los dos cursos correspondientes al segundo ciclo de Educación Infantil). Además, se incluye el análisis de tres dimensiones: contexto, procesos y productos.

La muestra sobre la que se realiza el estudio está constituida por 184 alumnos (130 en centros de reforma y 54 en centros ordinarios) y por 10 profesores pertenecientes a seis colegios públicos (cuatro integrados en el Plan Experimental para la Reforma de la Educación Infantil, y dos ajenos al mismo). Los criterios empleados para su selección fueron: a) pertenecer a centros situados en Zaragoza capital, b) incluir todos los centros participantes en el Plan, c) disponer en el grupo de control (no reforma) de centros representativos de modelos tradicionales de Educación Infantil y equiparables en cuanto a ubicación (nivel socioeconómico), y d) selección en cada centro de 35 alumnos del mismo grupo-clase.

Los dos grupos básicos de comparación (si bien no únicos) son los centros con programas desarrollados en el marco del Plan Experimental y los ajenos al mismo.

Además, para la evaluación de los efectos de un programa complementario para padres desarrollado a lo largo de la investigación, el diseño central se complementa con un diseño experimental.

Tres son las principales fuentes de información utilizadas en la evaluación: personales (distintos responsables de la Administración, inspectores de zona, directores de los centros, equipos de apoyo, etc.); documentales (proyectos educativos y planes de centro, programaciones, instrumentos de evaluación, expedientes de los alumnos, etc.); y comportamentales (test aptitudinales y actitudinales, observaciones directas de conducta, interacciones sociales, etc).

Los procedimientos y técnicas empleados en la recogida de la información son notablemente variados, y están constituidos tanto por pruebas estandarizadas para la medida de las variables criterio, como por instrumentos elaborados *ad hoc* para las variables independientes. De este modo se emplearon tests y pruebas de aptitudes, escalas y listas de observación, instrumentos para la evaluación del desarrollo cognitivo, para la identificación del modelo educativo del centro y del grado de adecuación de su programa al Diseño Curricular Base, cuestionarios/entrevistas para padres y profesores, entrevista/guión sobre actitudes hacia la escuela para niños, inventario para la observación del centro, reuniones, grabaciones en audio y video, etc.

Por su parte, el trabajo sobre la experimentación de los Módulos Profesionales (C.I.D.E.) adopta, como los dos anteriores, los esquemas y diseños propios de la investigación empírica tradicional, incluyendo la comparación con otro programa (el de Formación Profesional) desde un enfoque cuantitativo.

De acuerdo con sus objetivos generales (conocer cómo se percibían los Módulos Profesionales, qué grado de satisfacción mostraban los agentes directamente implicados en ellos, cuáles eran las ventajas reales que ofrecían), las fuentes empleadas en todos los casos son las personas implicadas en la puesta en marcha de los programas. Concretamente se consultó a los siguientes sectores:

- Los profesores encargados de desarrollar los programas (Módulos y Formación Profesional)
- Los alumnos que seguían los programas.
- Las empresas implicadas en los programas.

Los tamaños de las muestras empleadas en el estudio fueron los siguientes: 675 alumnos, 329 profesores y 143 empresas de Módulos Profesionales; 283 alumnos, 116 profesores y 73 empresas de Formación Profesional.

Las informaciones requeridas a los tres colectivos seleccionados fueron de naturaleza tanto descriptiva, como valorativa. En todos los casos se trata de información indirecta acerca del desarrollo de los programas, tal y como se deduce de las fuentes de información empleadas.

Las valoraciones requeridas a los alumnos incluían satisfacción con los estudios, valoración de la formación recibida, evaluación de la prueba de acceso, del área de proyecto y de las clases de expertos. De

los docentes, se recaba información, además, sobre la satisfacción con la labor docente, la preparación de los Módulos, la intervención en la organización de prácticas en la empresa, etc. De las empresas el tipo de información que se recoge se refiere además de a informaciones descriptivas, a cuestiones igualmente valorativas sobre las prácticas, sobre la formación de los alumnos, etc.

El cuestionario es el instrumento de recogida de información utilizado en la Evaluación de los Módulos Profesionales. Se construyeron instrumentos *ad hoc* destinados a alumnos, profesores y empresas que acogían a alumnos en prácticas, tanto de Módulos como de Formación Profesional.

Para los tres estudios que han sido descritos hasta ahora, y por lo que se refiere al análisis de la información obtenida, cabe destacar el alto grado de afinidad en el modelo de explotación de los datos. Los tres estudios proceden al cálculo de estadísticos descriptivos e inferenciales.

Sin embargo, si para estas mismas evaluaciones consideramos el modo de interpretación de los resultados de los análisis realizados, aparecen algunas diferencias relevantes.

Así, en los trabajos llevados a cabo por el Servicio de Evaluación del C.I.D.E. y dirigidos por Alvaro Page, las conclusiones formuladas son las que estrictamente proporcionan los índices estadísticos empleados, asumiendo pues un carácter estrictamente descriptivo, como de hecho se señala en el propio informe: "No se hacen en este trabajo juicios de valor sobre la bondad de la Reforma evaluada..."⁹. Esta observación es aplicable a la totalidad del proceso de evaluación de las Enseñanzas Medias, pero también a la Evaluación de la Enseñanza Técnico-Profesional.

El primer estudio es entendido por los autores como una medida u observación sistemática realizada sobre la extensión o calidad de la consecución de un logro. A este tipo de evaluación se atribuye dos fines principales: guiar el desarrollo y mejora del proceso educativo, y/o describir la consecución de un logro o rendimiento en un momento dado.

Sin embargo, la evaluación del Plan Experimental de la Educación Infantil realizada por Aguado Odina, si realiza valoraciones a partir de la información obtenida. El carácter de las conclusiones es primordialmente explicativo (como corresponde al modelo adoptado),

⁹ ALVARO PAGE, M. y col. (1990): *Evaluación Externa de la Reforma Experimental de las Enseñanzas Medias (II)*. C.I.D.E. Ministerio de Educación y Ciencia, pág. 21.

y aunque no se proporciona una valoración general de los programas del Plan evaluados, son puestos de manifiesto los puntos débiles y fuertes de los mismos.

Con los tres estudios anteriores guarda una cierta afinidad de enfoque la primera fase de la Evaluación del Programa de Integración Escolar (VV.AA.), caracterizada por los propios autores como cuantitativa, y aun la segunda. No obstante, en tanto que este trabajo cuenta con un diseño elaborado desde un modelo evaluativo netamente distinto, merece consideración aparte.

Efectivamente, en el informe final la evaluación es caracterizada por los autores como un modelo global de evaluación formativa e iluminativa. El estudio no reúne las características de una investigación experimental en sentido estricto y, sin renunciar a ello, se adopta una estrategia de validación convergente. De este modo se utilizan de forma conjunta y coordinada metodologías diferentes con el fin de superar las limitaciones en la capacidad explicativa de cada una de ellas.

La muestra empleada estuvo constituida como sigue:

- Del total de 173 centros implicados en el Programa en el curso 1985/86, se seleccionó aleatoriamente una muestra de 60 centros para realizar con ellos el seguimiento. De estos, continuaron hasta el final 52. El número de niños integrados en los 60 centros de los que se poseen datos del primer año fue de 379. Asimismo se seleccionó aleatoriamente una muestra de 379 niños pares, procedentes de aulas en las que habían sido integrados los niños anteriores.
- De los centros inicialmente seleccionados, se escogieron tres centros característicos y paradigmáticos de los tres modelos de centro (tres formas específicas de aplicar el Programa), a los que se unió posteriormente un centro más.

La metodología adoptada se concreta en las tres distintas fases de la investigación (con diferentes procedimientos, diseño y temporalización):

- Primera fase. Constituye un estudio extensivo, realizado mediante procedimientos estadísticos. El diseño es longitudinal, ya que se prolongó tres años, desde el curso 1985/86 hasta el curso 1987/88, y se procedió a la medición periódica de las variables y dimensiones identificadas como componentes principales del programa. Incluye también numerosos diseños parciales.

- Segunda fase. Se centra en la medición de algunos de los componentes del programa en dos centros de la muestra. Esta fase no tenía por objeto establecer una comparación global, sino aportar informaciones que pudieran servir de complemento a los análisis de la primera fase.
- Tercera fase. Consistió en un análisis en profundidad de ciertos casos, mediante métodos propios de la investigación etnográfica. Esta fase incluye la utilización de dos procedimientos: a) análisis cualitativo, y b) proceso de intervención. La única diferencia entre ambos reside en que en el segundo se produjo una intervención sistemática previamente planificada de acuerdo con las hipótesis directrices de la evaluación. Esta última fase adopta los parámetros propios de la investigación-acción, entendiendo por tal el estudio de una situación social con el fin de mejorar la calidad de acción dentro de la misma. De hecho, los autores afirman que la finalidad última del estudio es conseguir un cambio social a través de la intervención.

A lo largo de las distintas fases se emplearon fuentes de información tanto personales, como documentales y comportamentales. Los distintos tipos de fuentes adquirirían un mayor protagonismo dependiendo del momento evaluativo, pero se empleaban de forma combinada.

Los procedimientos empleados para recoger la información necesaria en la primera fase (cuantitativa), son notablemente variados, de acuerdo con el tipo de informaciones que era preciso recoger. De este modo se aplicaron cuestionarios (para el análisis del funcionamiento de la escuela, de aspectos curriculares e instructivos, de opiniones pedagógicas, de actitudes de los padres ante la integración, etc.); tests o pruebas estandarizadas de inteligencia, de aptitudes y pedagógicas; escalas de observación, hojas descriptivas, informes cualitativos, etc.

Las técnicas cualitativas utilizadas son variadas, e incluyen entrevistas grabadas semiestructuradas y abiertas; observación participante; análisis de documentos escritos; recogida de materiales; reportaje fotográfico del centro; elaboración de diarios y notas de campo; encuestas, cuestionarios y reuniones.

Cabe destacar el hecho de que, en muchas ocasiones, los procedimientos utilizados servían a un doble fin, tanto de recogida de información como de realimentación de la acción (fase de intervención).

A partir de la información recogida en la fase cuantitativa del estudio, se aplicaron diversas técnicas de análisis estadístico de los datos, tanto de tipo exploratorio y descriptivo como inferencial.

En el análisis cualitativo se utilizaron técnicas propias de este enfoque (agrupación en categorías, análisis de opinión, contrastación de resultados parciales en procesos generados desde la triangulación, etc.). Concretamente, se utilizó como guía para el análisis el modelo teórico de la investigación, permitiendo la elaboración de un informe individual de cada centro, en el que se presentaban los resultados de la evaluación del desarrollo del Programa en el centro. Se analizaron pormenorizadamente también los valores y características que adoptan los distintos sistemas implicados, siguiendo el orden y las directrices generales conjeturadas desde el modelo. Además de esta descripción, se formularon las relaciones de influencia entre las variables, emitiéndose una valoración sobre las mismas. Para el proceso de intervención se emplearon técnicas de análisis similares.

A partir de los resultados obtenidos, se llevó a cabo una formulación de conclusiones que trataba de reflejar el análisis y, por tanto, la interpretación convergente que el propio equipo de investigación hizo de los datos obtenidos. Dado que se aspiraba a realizar una evaluación de carácter formativo e iluminativo, se obviaron las explicaciones detalladas e irrelevantes para adoptar en las conclusiones una perspectiva global y significativa.

En términos generales, puede decirse que los resultados son interpretados en términos descriptivos, pero también explicativos y de comprensión de la realidad. Igualmente, y como se preveía hacer, se emiten juicios de valor fundamentados, tanto sobre aspectos concretos como globales del programa.

Desde el punto de vista metodológico parece también asumirse un enfoque mixto en el estudio de Escudero Muñoz, actualmente en curso, sobre *la diseminación y el desarrollo de la L.O.G.S.E.*

El proceso de evaluación está dirigido por las hipótesis de trabajo, consideradas en un sentido amplio y no estrictamente estadístico.

Se ponen en juego técnicas cuantitativas bajo criterios de significatividad estadística en un diseño longitudinal que se extiende a lo largo de dos años. En el muestreo se utilizan como criterio de estratificación las seis zonas correspondientes a otros tantos Centros de Profesores y, dentro de ellas, se hace una afijación proporcional. Debido al diseño longitudinal previsto, se establecen dos muestras equi-

valentes para dos momentos de medida distintos en el tiempo. Para los análisis más cualitativos se seleccionan sujetos/centros prototípicos, sobre los que se pretende realizar un estudio de casos.

Si bien en este estudio no se explicita estrictamente el modelo de evaluación adoptado, sí se presentan el marco teórico y las asunciones metodológicas que comporta desde el punto de vista de la investigación evaluativa diseñada.

Más allá de los modelos preocupados por los procesos de cambio y reforma, y orientados básicamente a tomar decisiones de implantación, en este caso se adopta una perspectiva de investigación que sostiene la necesidad de entender y analizar los fenómenos de diseminación de los cambios educativos en estrecha relación con el desarrollo de los mismos.

Como se señala en el propio trabajo, la investigación sobre el cambio educativo desde esta perspectiva trata de contribuir a la comprensión del conjunto de condiciones sociales, institucionales, comunitarias, profesionales, administrativas, etc., que, en un momento determinado, caracterizan la realidad de los centros y profesores, pues todo ello parece interaccionar con las propuestas de reforma. Desde esta perspectiva, los cambios no sólo son adoptados, sino adaptados, reconstruidos y desarrollados en la práctica, debido a la existencia de condiciones y procesos que median entre lo que las reformas postulan y lo que efectivamente llega a realizarse.

Las fuentes de información son tanto personales como comportamentales. Las primeras están constituidas en todos los casos por los agentes implicados de modo directo o indirecto en la puesta en marcha de la L.O.G.S.E. (Equipos Directivos, profesores no implicados formalmente y profesores implicados formalmente en su aplicación). Como fuentes comportamentales se emplean la observación directa de conductas del profesor y un cuestionario descriptivo-valorativo.

En consecuencia, las informaciones que está previsto recoger son tanto cuantitativas como de naturaleza cualitativa; así como información directa acerca del programa e información indirecta acerca del mismo.

Para recoger la información, los procedimientos previstos son, por tanto: a) Cuestionario sobre el desarrollo inicial de la L.O.G.S.E., destinado a los profesores de la región; b) Entrevistas semi-estructuradas, realizadas a equipos directivos y a profesores implicados y no implicados formalmente en el desarrollo de la Reforma; y c) Observaciones

de aula, obtenidas mediante la técnica del relato y orientadas formalmente por una concepción ecológica de la enseñanza, inspirada en el modelo de W. Doyle.

El tratamiento de la información de naturaleza más cuantitativa (procedente del cuestionario), se plantea mediante técnicas estadísticas, mientras que con la información de naturaleza cualitativa (estudio de casos: entrevista y observación directa), se establecerán categorías de contenido y se elaborarán matrices funcionales, temáticas, etc.

Se entiende que los datos recogidos a partir de las observaciones, constituyen un conjunto de evidencias empíricas que son leídas e interpretadas por el investigador a la luz de los presupuestos que, según la L.O.G.S.E., deben inspirar el desarrollo de la enseñanza. A partir de tales informaciones, será posible determinar el empleo que, de las propuestas de la reforma, están haciendo los profesores en su enseñanza, y en qué medida y por qué, la implantación de la misma está traduciendo o no en procesos de reconstrucción de la propia enseñanza por parte de los docentes.

Por último, consideraremos el enfoque adoptado en la Evaluación del Programa Prensa-Escuela. Este estudio, si bien no se puede alinear con ningún modelo clásico de evaluación, es fácilmente clasificable dentro de aquellas evaluaciones que se construyen sobre la adaptación de la técnica de encuesta como instrumento con el que se obtiene la información a partir de la cual "describir/explicar/valorar" un programa. El enfoque metodológico, en consecuencia, es cuantitativo-estadístico.

En un diseño como el que se arbitra para este estudio, es básico el procedimiento de muestreo. Concretamente se consideraron dos universos: 1) centros de enseñanza públicos y privados de E.G.B., B.U.P. y F.P.; 2) el conjunto de profesores que desempeñan su actividad docente en dichos niveles (objeto central del estudio).

El diseño muestral asume cada uno de los niveles educativos como muestras independientes, y para cada uno se arbitra un procedimiento polietápico estratificado por conglomerados y afijación proporcional. Las variables de estratificación empleadas son: Comunidad Autónoma, tamaño del hábitat y tipo de enseñanza (pública/privada). Las unidades primarias seleccionadas de forma aleatoria proporcional al número de profesores adscritos, o de forma aleatoria simple, están constituidas por los centros. Las unidades secundarias, seleccionadas por combinación de criterios aleatorios y de cuotas, fueron los profesores.

Las fuentes de información empleadas fueron, en todos los casos, personales. Concretamente, los informantes fueron los centros seleccionados en el muestreo (a través de su director), adscritos o no al Programa Prensa-Escuela, y los profesores potenciales a los que éste se dirige.

En ambos casos, se recogieron informaciones indirectas acerca del programa. Unas tienen un carácter fundamentalmente descriptivo y otras más valorativo. Igualmente se recogió la opinión de directores y profesores acerca de la posible realización de una evaluación cualitativa del Programa, que finalmente no llegó a realizarse.

Para la recogida de información se emplearon dos cuestionarios diseñados a tal efecto. Uno de ellos dirigido a los centros (a cumplimentar por el director), y el segundo, dirigido a los profesores. Se trata en ambos casos de cuestionarios de respuesta cerrada. Aunque en algunas ocasiones permitían las respuestas abiertas, estas son tratadas también cuantitativamente. Los cuestionarios se aplicaron mediante entrevista personal.

A partir de los datos recabados, se lleva a cabo un análisis descriptivo para todas las variables incluidas en el cuestionario, así como una tabulación cruzada de las variables consideradas principales en el estudio (Comunidad Autónoma, hábitat, nivel educativo, tipo de centro y adscripción o no al Programa) con las correspondientes al tema “la prensa en el aula”. Una vez sistematizada y analizada la información, se ofrece una descripción de las respuestas, sin entrar en consideraciones explicativas o valorativas del Programa.

— *Difusión y uso de los resultados de la evaluación*

Tal y como se apuntaba al inicio, las evaluaciones sobre reformas y programas experimentales incluidas en este grupo, auspiciadas en buena parte por el Ministerio de Educación, nacen de las necesidades de contraste, retroalimentación y control del importante conjunto de innovaciones emprendidas por la Administración educativa a lo largo de la pasada década. En relación a las respuestas y conclusiones que los estudios comentados han proporcionado, resultan ilustrativas las palabras del entonces Director del C.I.D.E., Alejandro Tiana, en el prólogo al II Informe de la Evaluación del Plan Experimental de las Enseñanzas Medias: “... ha sido precisamente el análisis de ésta y

otras experiencias el que ha permitido avanzar a la hora de elaborar un proyecto general de transformación del sistema educativo. La puesta en marcha de un proceso controlado de cambio ha permitido su experimentación y reorientación antes de proceder a su generalización” (Alvaro Page et. al., 1990: pág. 18).

Por lo que se refiere a la difusión realizada de los resultados obtenidos en las investigaciones, los informes finales de las evaluaciones correspondientes a la Reforma Experimental de las Enseñanzas Medias, de las Enseñanzas Técnico-Profesionales y del Programa de Integración, fueron publicados por el C.I.D.E.

En concreto, los informes de las distintas fases de la Evaluación Experimental de las Enseñanzas Medias fueron publicados en tres volúmenes en los años 1988, 1990 y 1992¹⁰. Al mismo tiempo, se procedió al envío a cada uno de los centros participantes de los resultados obtenidos por sus alumnos en el estudio. Por su parte, el informe de la evaluación de los Módulos Profesionales se publicó en 1993¹¹.

En el informe correspondiente a la Evaluación del Programa de Integración Escolar¹², publicado en 1990, se explica cómo a lo largo de la investigación se devolvió a los centros, a los equipos psicopedagógicos e incluso a los padres implicados en el trabajo, información relativa a la marcha del Programa y a la propia evaluación. La valoración llevada a cabo por este equipo de la difusión de resultados y conclusiones realizada entre los sujetos directamente implicados en el programa, pone de manifiesto la relevancia de estos procesos de difusión en los estudios evaluativos: en todas las ocasiones el envío de información a los interesados se ha valorado muy positivamente, al tiempo que se consideraba negativo que no se hiciera con más frecuencia y continuidad.

Sin embargo, los ámbitos en los que se toman las decisiones que afectan al destino de los programas o reformas experimentales, no

¹⁰ ALVARO PAGE, M. et al (1988): *Evaluación externa de la Reforma Experimental de las Enseñanzas Medias (I)*. C.I.D.E. Ministerio de Educación y Ciencia.

ALVARO PAGE, M. et al (1990): *Evaluación externa de la Reforma Experimental de las Enseñanzas Medias (II)*. C.I.D.E. Ministerio de Educación y Ciencia.

ALVARO PAGE, M. et al (1992): *Evaluación externa de la Reforma Experimental de las Enseñanzas Medias (III)*. C.I.D.E. Ministerio de Educación y Ciencia.

¹¹ ALVARO PAGE, M. et al (1993): *Evaluación de los Módulos Profesionales: Estudio de la Reforma Experimental de la Enseñanza Técnico-profesional*. C.I.D.E. Ministerio de Educación y Ciencia.

¹² VV.AA. (1990): *Evaluación del Programa de Integración Escolar de Alumnos con Deficiencias*. C.I.D.E. Ministerio de Educación y Ciencia.

siempre suelen coincidir con aquellos ámbitos sobre los que los evaluadores tienen capacidad de intervención. Este es el caso del estudio sobre el Plan Experimental para la Reforma de la Educación Infantil, realizado por Aguado Odina. Sin embargo, el informe, fiel a su orientación formativa, finaliza con un capítulo de recomendaciones para la toma de decisiones, clasificadas en dos tipos:

- Recomendaciones orientadas a potenciar objetivos y acciones ya contenidos en el Proyecto para la Reforma de la Educación Infantil, cuya relevancia y eficacia han sido puestas de manifiesto en el estudio.
- Orientaciones y sugerencias destinadas a los equipos implicados en el diseño y desarrollo de programas específicos.

Naturalmente nada se puede apuntar todavía de la difusión y uso de los resultados del trabajo en curso de Escudero Muñoz sobre la diseminación y el desarrollo de la L.O.G.S.E.

— *Metaevaluación*

La evaluación del Programa de Integración es el único estudio de los incluidos en este grupo en el que se contempla algún elemento próximo a lo que supone realizar una evaluación de la evaluación emprendida, si bien ésta no constituye un procedimiento previsto o sistemáticamente diseñado desde el inicio del trabajo. No obstante, su realización dota a este estudio de importantes elementos para su comprensión y utilización como modelo de trabajo evaluativo.

4.2.2. Evaluación de programas curriculares de aplicación restringida

4.2.2.1. Evaluación de programas curriculares para distintas áreas y niveles educativos

— *Objeto de la evaluación*

En este grupo de estudios evaluativos, se han incluido tres trabajos que, centrados en el ámbito curricular, no tienen en cambio como

referente la evaluación de amplios procesos de experimentación o innovación. Además, en general poseen un carácter notablemente más restringido que el de los estudios tratados en el apartado anterior. Así, si en aquel se incluyen estudios que operan sobre planes curriculares de aplicación a toda una etapa educativa o a varias, y con un ámbito de evaluación notablemente amplio, ahora se analizan estudios que están centrados en programas para áreas concretas del currículo o en la consideración acotada de una etapa del sistema educativo.

De este modo, se incluyen en este grupo dos evaluaciones que tienen por referente la etapa de Formación Profesional. La primera de ellas, *Estudio sobre la Formación Profesional en Asturias: valoración de resultados* (1989), está dirigida por Miguel Ángel Caldevilla Fernández. El programa evaluado en este estudio consiste en el currículo de la Formación Profesional, y su ámbito de estudio es la provincia de Asturias.

Por su parte, el trabajo dirigido por Miguel Ángel Zabalza Beraza *Evaluación del programa de Prácticas en Alternancia de la Comunidad Autónoma Gallega* (1991) se centra en el análisis y valoración de un programa de implantación territorial aplicado en Galicia, a partir del curso 1986-87, para los alumnos del último curso de Formación Profesional de segundo grado. El programa de Prácticas en Alternancia, de considerable implantación (56 centros gallegos en la primera convocatoria), constituye un elemento integrado de modo opcional en la oferta formativa proporcionada por los centros en este nivel educativo, y constituye el objeto de evaluación.

Por último se incluye la investigación dirigida por Felipe Martín Hernández bajo el título *Evaluación del currículum y de la didáctica de la Filosofía en 3.º de B.U.P.* (1991). El programa objeto de evaluación es el currículo oficial establecido para esta materia del Bachillerato Unificado y Polivalente en todo el Estado. El ámbito en que se ha desarrollado el estudio ha sido la Comunidad Autónoma de Madrid (los centros docentes públicos y privados concertados, constituyen la población de referencia).

— *Características definitorias de la evaluación*

Agentes evaluadores

En los tres trabajos citados se diseña y lleva a cabo una evaluación externa. Se considera que este carácter externo es aplicable a la eva-

luación del currículo de Filosofía (Martín Hernández), a pesar de la posible filiación de sus autores a la profesión docente en el área y nivel de interés. Tanto si consideramos como unidad de análisis el Departamento o el centro educativo, y dado el alcance del estudio, los evaluadores no forman parte de las instancias directamente implicadas en el desarrollo y puesta en marcha del programa, considerado en su conjunto. En los dos estudios restantes, la evaluación es pilotada y desarrollada por expertos procedentes del ámbito universitario.

Igualmente, se trata en los tres casos de heteroevaluaciones que no incluyen entre sus procedimientos la autoevaluación.

Objetivos y finalidad de la evaluación

En las tres evaluaciones analizadas se aprecian notables diferencias en cuanto a los objetivos perseguidos por cada una de ellas.

El estudio de Caldevilla Fernández se propone valorar la eficacia de la Formación Profesional en la sociedad asturiana desde la perspectiva del “producto docente”. Se trata por lo tanto de evaluar el grado y modo de consecución de los dos grandes objetivos o finalidades que se atribuyen a esta etapa del sistema educativo: proporcionar formación para la inserción laboral efectiva, y proporcionar formación que posibilite, en su caso, la continuación de estudios. Por otro lado, el trabajo tiene una clara vocación de informar la toma de decisiones en el terreno de la política educativa y de la ya entonces anunciada reforma para esta etapa (a la que se hacen numerosas referencias). Se trata de determinar puntos fuertes y débiles, causas de disfuncionalidades, etc., que permitan la proposición de modificaciones y mejoras del programa.

El otro estudio referido a la Formación Profesional, concretamente al Programa de Prácticas en Alternancia en la Comunidad Autónoma gallega, aparece sin embargo centrado en los procesos y los principios de funcionamiento, y no en la evaluación de resultados. Así, entre los fines básicos de la evaluación se señalan: disponer de datos (más allá de las opiniones o impresiones generales) sobre cómo estaban funcionando en realidad las prácticas y determinar de qué manera se podía mejorar su integración en el proceso formativo normal de los alumnos, de manera que no resultase una experiencia “fuera de programa”.

Los propios autores entienden la investigación como un proceso de obtención de información relevante con respecto a las demandas

planteadas y, por lo mismo, orientada a la toma de decisiones y con un carácter esencialmente formativo.

Por su parte, el estudio de Martín Hernández sobre el currículum y la didáctica de la Filosofía, coincide básicamente con el anterior en la naturaleza de sus objetivos y finalidades. La investigación se ordena, en términos generales, a conocer y valorar qué es lo que efectivamente se está haciendo en esta asignatura del Bachillerato: qué metas se está pretendiendo conseguir, qué materia concreta se están impartiendo, qué tipo de selección curricular se está llevando a cabo y con qué criterios, cómo o mediante qué procedimientos didácticos discurre la tarea docente, etc. En definitiva, el estudio pretendía aportar elementos de juicio e informar, desde una perspectiva de evaluación formativa, la toma de decisiones en relación a la Filosofía en el nuevo Bachillerato, a partir del conocimiento y ponderación de las prácticas actuales y de la opinión y valoración de los profesores de esta asignatura.

Dimensiones evaluadas y momento de la evaluación

El trabajo de Caldevilla Fernández constituye una evaluación de producto o de los resultados educativos de los alumnos de Formación Profesional en el ámbito territorial de Asturias. Más concretamente, los principales indicadores empleados en la evaluación de resultados son los siguientes:

- Actividad académica posterior a la terminación de la Formación Profesional (otros estudios medios, estudios superiores, etc.).
- Inserción profesional posterior a la terminación de la Formación Profesional.
- Relación de la actividad laboral realizada con los estudios cursados y utilidad subjetiva.
- Necesidades de formación complementaria.

Por su parte, el estudio de Zabalza Beraza sobre las prácticas en alternancia en la Formación Profesional gallega, supone una evaluación global del funcionamiento o puesta en marcha del programa. Para la evaluación de la dimensión procesual en que se centra el trabajo, se incluyen factores tanto organizativos (horarios, momento del curso en

que se realizan las prácticas, voluntariedad, relación con el trabajo actual, etc.), como de gestión (selección de empresas, contraprestaciones, etc.) y estrictamente didáctico-curriculares (objetivos, relación de las prácticas con las otras materias del currículo, existencia y papel de los tutores, evaluación de los alumnos de prácticas, etc.). En el momento en que se lleva a cabo el estudio, el programa cuenta con algo más de dos años de rodaje.

Finalmente, el estudio evaluativo sobre el currículo y la didáctica de la Filosofía en el tercer curso del B.U.P., centrado igualmente en los elementos procesuales del programa curricular, comprende como dimensiones generales a evaluar: las concreciones curriculares que se están realizando en las aulas y Seminarios didácticos de los centros (objetivos, selección de contenidos, opciones metodológicas, ajuste al programa oficial y percepción del mismo, etc.) y el grado de cohesión y coordinación en el funcionamiento de los propios Seminarios (estabilidad de sus miembros; elaboración, ejecución y evaluación del currículum; participación en tareas de formación, de investigación y elaboración de materiales, etc.).

Patrones empleados para la emisión de juicios de valor

El estudio evaluativo de la Formación Profesional en Asturias, tal y como se ha adelantado, constituye una evaluación construida fundamentalmente sobre el contraste de consecución de objetivos, que funcionan en este sentido como los criterios de bondad. En el trabajo de Zabalza Beraza, parece que a la hora de interpretar y comentar los resultados obtenidos, se adopta fundamentalmente la perspectiva de la competencia profesional del evaluador para emitir juicios de valor sobre el programa de prácticas en alternancia (modelo del juicio experto). Por su parte, el trabajo de Martín Hernández tiene un carácter fundamentalmente descriptivo, sin realizar valoraciones en la interpretación de los resultados obtenidos en la evaluación del currículo de Filosofía.

— *Diseño y desarrollo de la evaluación*

En ninguno de los tres estudios se explicita estrictamente el modelo de evaluación adoptado.

En el trabajo dirigido por Zabalza Beraza sobre el Programa de Prácticas en Alternancia se proporcionan algunas características que, atribuidas al proceso de evaluación, contribuyen en gran medida a la caracterización del marco evaluativo de referencia:

- Proceso ampliamente participativo, en el que intervienen todos los protagonistas del programa que se pretende evaluar.
- Investigación orientada a la toma de decisiones.
- Enfoque cualitativo en la obtención, tratamiento e interpretación de los datos obtenidos.
- Diversidad en la naturaleza de las informaciones recogidas, de las fuentes de información empleadas y de los instrumentos usados para su recogida.

Los otros dos trabajos incluidos en este grupo adoptan el esquema propio de los estudios evaluativos basados en la técnica de encuesta, al considerarse el procedimiento de recogida de la información más necesario para responder a la totalidad de preguntas derivadas de los objetivos formulados en la evaluación. Se trata por tanto de dos estudios cuya estrategia se centra, básicamente, en la localización de las fuentes de recogida de información y en el procedimiento de acceso a las mismas.

Retomando el trabajo de Zabalza Beraza, y por lo que se refiere a las fuentes de información empleadas, éstas fueron en todos los casos personales. Concretamente se recabó información de las autoridades educativas de la Consellería en el ámbito de la Formación Profesional; de los responsables específicos del programa (coordinadores provinciales de empleo, directores de centros adscritos al programa de prácticas en alternancia, etc.); de representantes de las empresas (propietarios o responsables de formación); de los tutores de las empresas; de los profesores de Formación Profesional que eran tutores de prácticas, y de los propios alumnos. Las informaciones recogidas sobre el programa de prácticas son en todos los casos indirectas y de carácter subjetivo (opiniones, expectativas, valoraciones y propuestas de mejora de los implicados en el programa).

Dos son las técnicas de recogida de información empleadas. En la primera fase del estudio (contacto con los protagonistas), se organizaron dos reuniones —entrevistas colectivas—, que tenían una función exploratoria y se estructuraron de forma abierta en torno a un tema

central. Para la segunda fase de la investigación evaluativa se construyeron cuestionarios destinados a profesores-tutores de centro, para tutores de empresa y para los alumnos.

Las muestras de estos tres colectivos se seleccionaron considerando la representación equilibrada de todas las especialidades de Formación Profesional, de las diversas zonas de la Comunidad, y de los distintos tipos de centros. El tamaño final de las muestras fue de 448 alumnos que realizaban prácticas en ese momento, 35 profesores tutores en sus respectivos centros y 39 tutores en las empresas.

De acuerdo con los dos tipos de procedimientos empleados para la recogida de datos, con los que se obtuvo información de diversa naturaleza (cualitativa y cuantitativa), se emplean dos técnicas distintas para el análisis de los datos. A las reuniones exploratorias se les aplica análisis de contenido, identificando ocho grupos temáticos que a su vez son analizados por bloques afines. Por otro lado, con las respuestas a los cuestionarios se realiza análisis de frecuencias. Puesto que éstos instrumentos incluían algunas preguntas abiertas, se llevan a cabo análisis cualitativos (de tipo semántico) como paso previo a la tabulación.

Finalmente, la interpretación de los resultados se hace en términos de comprensión y explicación, realizando algunas interpretaciones valorativas.

Las fuentes de información empleadas en el estudio sobre la Formación Profesional en Asturias, son tanto personales como documentales. Se recaba información de antiguos alumnos de F.P. que acabaron su formación entre 1979 y 1982. El procedimiento empleado fue la técnica de encuesta por correo. Además, se revisaron las solicitudes de empleo, cursadas como consecuencia de ofertas del Instituto Nacional de Empleo (I.N.E.M.), a quienes estaban en posesión del título de Técnico Especialista (F.P.2) o de Bachillerato (B.U.P.). Se contó, por lo tanto, con dos muestras, ambas incidentales. Una de ellas estaba representada por el número de cuestionarios devueltos por la población de referencia (474 encuestas, que suponen el 50% de los casos); y la segunda, por el conjunto de casos revisados (612) en los registros del Instituto Nacional de Empleo. La información recabada sobre el Programa es, pues, tanto directa como indirecta.

Los datos recogidos son tratados estadísticamente, realizándose en todos los casos análisis descriptivos. A partir de los resultados hallados, se procede al comentario de los resultados desde un punto de vista claramente valorativo.

En el estudio evaluativo sobre la filosofía en el Bachillerato, dirigido por Martín Hernández, la muestra estuvo constituida por 60 Institutos de titularidad pública y 170 profesores (muestra considerada representativa de la población dado el ámbito de aplicación del trabajo). Para su selección se considera importante la inclusión de Institutos de Bachillerato de todas las circunscripciones territoriales de los Centros de Profesores (C.E.P.) de la Comunidad de Madrid, para así representar convenientemente a los centros ubicados en el centro y en la periferia. Se contó además con 20 centros privados-concertados, con un total de 26 profesores, que no se consideran una muestra representativa de la población de referencia.

En este trabajo se emplea como fuente de información a los propios responsables de la puesta en práctica del programa curricular: Jefes de Seminario y profesores de la materia y nivel de interés. En consecuencia, los datos recabados acerca del programa son de carácter indirecto, y subjetivo, tanto cuando se requiere información descriptiva, como valorativa o propositiva.

Los procedimientos empleados para la recogida de información, y en parte adelantados anteriormente, consistieron en un cuestionario a los profesores y entrevistas mantenidas con el Jefe de los Seminarios Didácticos de aquéllos Institutos donde se aplicó el cuestionario individual, o con otro u otros miembros del mismo ante la ausencia de aquél.

La información que realmente da lugar a la obtención de conclusiones procede de la aplicación del cuestionario, que es objeto de análisis estadísticos de tipo descriptivo. Los análisis realizados permiten la adscripción de los profesores a categorías correspondientes a distintos enfoques del currículo. Se realiza una descripción e interpretación de resultados sin emplear pruebas de significación estadística, obteniéndose un informe descriptivo, en el que no se avanzan interpretaciones explicativas ni valoraciones globales del programa, sino que las respuestas de los encuestados se interpretan en términos de coherencia/incoherencia con un enfoque determinado del currículo.

— *Difusión y uso de los resultados de la evaluación*

El trabajo dirigido por Zabalza Beraza, de acuerdo con la información disponible, nace vinculado a una de las necesidades puestas de manifiesto en la primera reunión de expertos y profesionales de la For-

mación Profesional, auspiciada por la Administración educativa gallega sobre las prácticas en alternancia. A pesar de ello, y de la clara orientación formativa de la evaluación llevada a cabo, parece claro que el uso efectivo de sus resultados excede al ámbito de actuación de los evaluadores, que no pueden en principio sino proporcionar a los decisores políticos los *informes que correspondan*. No se ofrecen datos acerca de este último aspecto en la memoria, que no informa de las vías previstas para la difusión de resultados. Estas consideraciones, son igualmente aplicables a los otros dos estudios revisados.

El estudio de Martín Hernández incluye entre sus objetivos iniciales la presentación —a los responsables de tomar las correspondientes decisiones—, de propuestas útiles para la elaboración de un diseño curricular coherente en el área de Filosofía para el nuevo Bachillerato, a partir de los propios datos de la investigación y de las sugerencias explícitas del profesorado.

Por su parte, los autores del trabajo sobre la Formación Profesional en Asturias identifican tres campos en los que los resultados del estudio podrían ser útiles: en el proceso de reforma del currículo de Formación Profesional, en el ámbito de la Orientación escolar y profesional, y en el análisis sociológico del “producto” de la Formación Profesional.

— *Metaevaluación*

Por lo que puede apreciarse en las memorias de investigación, ninguno de los tres estudios incluidos y analizados en este grupo de investigaciones sobre evaluación de programas curriculares para distintas áreas y niveles educativos, prevé o lleva a cabo un proceso de metaevaluación.

4.2.2.2. Evaluación de programas instructivos específicos

a. *Evaluación de programas desarrollados en Educación Primaria y Secundaria*

a.1. Evaluación de programas instructivos centrados en los procesos didácticos

En este grupo se incluyen aquellas investigaciones que han centrado todo su interés en el diseño y la evaluación de programas orien-

tados a la instrucción en determinadas destrezas cognitivas o actitudinales, y que, por lo tanto, se centran en los procesos didácticos de distintas áreas del currículo. Son las siguientes:

- *Un programa para instruir en la comprensión de textos expositivos en la etapa Secundaria Obligatoria* (Sánchez Miguel, 1989).
- *Desarrollo de programas para la mejora de la comprensión lectora y los procesos de razonamiento* (Alonso Tapia, 1991).
- *Diseño e implementación de un programa sobre habilidades cognitivas en el área de Matemáticas* (Garrido Gil, 1992).
- *La investigación del alumno desde la perspectiva de las Ciencias Sociales* (Rivera Otero, 1992).
- *Diseño y evaluación de programas de educación ética* (Martínez Martín, 1991).
- *Exploración de la situación bilingüe en los escolares de Melilla. Propuestas y estrategias de intervención educativa* (Mesa Franco, 1994).

Aunque todas las evaluaciones que aquí se incluyen están centradas en programas instructivos, sin embargo son muy diversos los ámbitos a los que se refiere cada una de ellas. De modo que esta revisión va a recoger, sobre todo, aquellos aspectos que caracterizan los procesos evaluadores de estos programas.

— *Objeto de la evaluación*

Las evaluaciones de programas que se incluyen en este grupo, cubren un amplio espectro de temas, dos programas para instruir en la comprensión de textos expositivos en la Educación Secundaria Obligatoria; un programa de entrenamiento metacognitivo para la mejora de la comprensión lectora y de los procesos de razonamiento; otro para mejorar el dominio del lenguaje en una situación de bilingüismo; un programa para desarrollar unidades didácticas vinculadas a la educación ética; la evaluación de diversas metodologías dentro de la enseñanza de las Ciencias Sociales; y la evaluación de un programa sobre habilidades cognitivas en el área de Matemáticas.

La evaluación dirigida por Sánchez Miguel (1989) se centra en comprobar la utilidad de un programa diseñado para instruir en la

comprensión de textos expositivos llevado a cabo en su medio escolar normal. Para ello se somete a una muestra de cuatro grupos de alumnos de 6.º y 7.º de E.G.B. a dicho programa, y se mide su nivel de comprensión lectora antes (pretest) y después de la aplicación del mismo (postest).

La investigación realizada por el equipo de Alonso Tapia (1991) tiene por objeto la evaluación de la eficacia de un programa de entrenamiento metacognitivo orientado a la mejora de la comprensión lectora y de los procesos de razonamiento a través de la comparación de los resultados de tres condiciones de instrucción: método tradicional, método sólo con modelado y programa de entrenamiento metacognitivo.

Garrido Gil y sus colaboradores (1992) llevan a cabo un estudio que tiene por objeto evaluar un programa sobre habilidades cognitivas en el área de Matemáticas, experimentándolo con una muestra de 90 estudiantes de 6.º de E.G.B. de dos centros públicos de Murcia, y comparando las diferencias en sus resultados con un grupo de control de las mismas características.

Por su parte, el estudio de Rivera Otero y su equipo evalúa mediante un estudio comparativo la influencia de un programa de enseñanza de las Ciencias Sociales, centrado en la investigación de los alumnos, en el rendimiento en dicha área. La muestra que se sometió al programa estuvo constituida por 157 alumnos de 8.º de E.G.B. de tres centros diferentes de Madrid, constituyendo tres grupos experimentales y tres de control con los que se procedió a estimar diferencialmente la eficacia del programa.

El objeto de la evaluación de Martínez Martín (1991) radica en el análisis del grado de eficacia de un programa de educación ética diseñado por los investigadores, aplicable a los niveles educativos de Primaria y Secundaria Obligatoria. Asimismo, se realiza un estudio analítico de distintos métodos o estrategias para la educación moral:

- Estrategias de autoconocimiento y autoexpresión: clasificación de valores y ejercicios de autoexpresión.
- Estrategias para el desarrollo del juicio moral: discusión de dilemas morales y reconocimiento de alternativas y previsión de consecuencias.
- Estrategias de comprensión conceptual: ejercicios de análisis y comprensión conceptual y filosofía para niños.

- Estrategias para el desarrollo de la perspectiva y la empatía: estudio de casos, role-playing y juegos de simulación.
- Estrategias para el desarrollo de la capacidad de diálogo, argumentación y acuerdo: debates, análisis de valores y resolución de conflictos.
- Estrategias orientadas al desarrollo de las competencias autorreguladoras, las habilidades sociales y la prosociabilidad.

Tras ello, los autores seleccionaron y adaptaron minuciosamente y por edades estas estrategias con el fin de elaborar su propio programa, destinado a los diversos niveles educativos de la Enseñanza Obligatoria. Tras haber efectuado un estudio sectorial en el que se interrogó a los alumnos de distintos grupos de edad sobre los temas moralmente relevantes para ellos se aseguró el consenso entre el profesorado, los alumnos y la opinión de los expertos con respecto al modelo y a los materiales diseñados. Para concluir, se realiza un esbozo de unos módulos curriculares y de unas unidades didácticas para la educación moral.

Por último, la investigación de Mesa Franco (1994) tiene por objeto evaluar un programa de intervención educativa para compensar y mejorar el dominio lingüístico del castellano. Se estudia el problema tradicional con el que se encuentra la población escolar de origen musulmán en Melilla, al vivir en lo que se denomina un medio lingüístico disglósico (en el que la lengua materna y la lengua oficial y de la escuela no siempre coinciden, puesto que hay niños que se expresaban en un dialecto árabe, el chelja). Esta confrontación lingüística de algunos niños/as tiene lugar al ingresar en Preescolar, fenómeno más pronunciado en las clases socioeconómicas desfavorecidas.

Este tema ya fue estudiado con anterioridad por este mismo equipo de investigadores (Mesa, 1989), mediante la comparación del nivel de comprensión verbal de alumnos de 3.º y 5.º de E.G.B. (Primer ciclo de Educación Primaria), en dos tipos de colegios; bilingües-musulmanes y monolingües-no musulmanes. En él se encontraron grandes deficiencias en la comprensión verbal de los alumnos bilingües frente a los que empleaban el castellano como lengua materna. De esta misma investigación se dedujo que los estudiantes seguían hablando chelja en su familia, y utilizaban el castellano sólo en el colegio o con amigos y hermanos de su misma edad. El estudio fue aplicado a una muestra de 281 escolares de Infantil y Primaria de Melilla de origen beréber, y pertenecientes a un estrato sociocultural medio y medio-bajo.

— *Características definitorias de la evaluación*

Agente evaluador

En relación a la investigación dirigida por Sánchez Miguel, un dato a tener presente es la existencia de un precedente que consiste en la evaluación iniciada en 1986 por este mismo autor y su equipo, —todos ellos pertenecientes al Departamento de Psicología de la Universidad de Salamanca—, que estudiaba los procedimientos para instruir en la comprensión de textos. Dicha investigación se dio por concluida en 1988, aunque la trayectoria investigadora iniciada por estos autores les llevó a elaborar una segunda evaluación —que se recoge aquí—, y que viene a recopilar los datos extraídos de la anterior a fin de avanzar formulando y proponiendo un programa para la instrucción en la comprensión de textos expositivos en la etapa de la Educación Secundaria Obligatoria.

Dichas investigaciones —cuyo interés queda focalizado en los programas de comprensión lectora— pueden considerarse en gran medida evaluaciones internas, dado que son los propios investigadores los que en calidad de expertos evalúan el programa que diseñan, con el objetivo de estimar posibles diferencias significativas en el aprendizaje de los alumnos.

Por su parte, la evaluación de Alonso Tapia y colaboradores, desarrollada desde el departamento de Psicología de la Universidad Autónoma de Madrid, se diseñan unos programas para la mejora de la comprensión lectora y los procesos de razonamiento. La preocupación por el fomento de la lectura y, por ende, de la comprensión lectora, en tanto factores que intervienen directamente en los aprendizajes, suscitó numerosos intentos por mejorar los mecanismos y las estrategias que conducen a incrementar la asimilación y la comprensión. Esta evaluación, al igual que las anteriores, es interna ya que el equipo investigador es el responsable del diseño de los programas, y, a su vez, de la evaluación de la eficacia de los mismos.

La investigación dirigida por Garrido Gil se propone dotar a los sujetos de estrategias de reconocimiento de la utilidad de su propia capacidad matemática, a través de un programa sobre *habilidades cognitivas* en el campo de los conceptos matemáticos. Consiste también en una evaluación interna, por las mismas razones que en los casos de las evaluaciones antes mencionadas.

Con respecto al estudio del equipo de investigadores del C.E.P. de Ciudad Lineal, dirigido por Alicia Rivera, se lleva a cabo una evaluación planteada desde distintos enfoques de la enseñanza de las Ciencias Sociales aplicables al contexto de enseñanza-aprendizaje dentro del aula. Pretenden una reflexión experimental ante la renovación del currículo las Ciencias Sociales en la Educación Secundaria Obligatoria, haciendo especial hincapié no sólo en el aprendizaje de conceptos, sino también de procedimientos y actitudes. Dicha evaluación es de carácter interno, puesto que es el mismo equipo el que diseña, aplica y evalúa una nueva metodología de enseñanza de las Ciencias Sociales.

Con la investigación coordinada por Martínez Martín se intenta dotar a los profesores de Educación Primaria y Secundaria Obligatoria de los instrumentos y recursos precisos para favorecer el estudio consciente y sistemático de la educación ética, contemplada como un tema transversal que impregna todas las áreas del currículum. En este sentido puede considerarse una evaluación externa, ya que los principales agentes son el equipo de expertos universitarios que firma la investigación, aunque sin duda se ha contado con la opinión y colaboración de profesores de los distintos niveles educativos no universitarios a los que se dirige el programa.

Por último, la investigación de Mesa Franco posee un carácter de evaluación externa, ya que los investigadores son los expertos que diseñan y promueven dicha evaluación, pero no son protagonistas de la enseñanza de los alumnos/as de los centros de Melilla.

Objetivos y finalidad de la evaluación

Sánchez Miguel y su equipo trabajan en una de las líneas de investigación más desarrolladas dentro de la Psicología de la Instrucción: el entrenamiento de habilidades de comprensión lectora. Entre sus objetivos destacan el entrenamiento de sujetos de nivel de rendimiento bajo para alcanzar una determinada actividad cognitiva, y por ende, el mismo rendimiento lector que los sujetos más capaces. Se considera que el propio programa de instrucción se convierte en una descripción verosímil de las diferencias evolutivas entre dos niveles de capacidad. Se parte de la idea vigotskiana de que, al incorporar los esquemas de interpretación y los universos simbólicos que conforman el legado de cada cultura, el sujeto conforma sus propios esquemas mentales.

El supuesto de partida es pues la necesidad de considerar las propiedades y mecanismos de la mente que hacen posible la capacidad de comprensión, análisis y síntesis de lo leído. En resumen, los supuestos teóricos de partida son los siguientes:

1. La capacidad para operar con los textos escritos es uno de los objetivos básicos que se deben alcanzar en la Educación Obligatoria.
2. Es preciso describir el proceso de actividad espontánea de nuestra mente mientras comprende el lenguaje habitual escrito.
3. Es preciso aislar el conjunto de estrategias que intervienen en el proceso de interpretación del material escrito.
4. Es posible enseñar a utilizar las estrategias mentales.
5. Los alumnos/as de menos capacidad precisan una enseñanza explícita de esas estrategias mentales.

En definitiva, su investigación se compromete a elaborar una descripción de la actividad mental espontánea, a detectar qué aspectos son más costosos para los sujetos discapacitados, y a mostrar cómo pueden enseñarse estas actividades en el aula, con lo que la finalidad de esta evaluación es claramente formativa.

Así, el objetivo prioritario es estimar las diferencias en la comprensión de la información escrita entre los distintos niveles de competencia, para poder determinar las actividades en las que se debe instruir. Tras comprobar esas diferencias, se abordan las dificultades que tienen aquellos escolares que demuestran un adecuado dominio de la lectura cuando la meta es extraer el significado general de lo que se lee, pero que, sin embargo, fracasan cuando se les exige un esfuerzo que requiere un despliegue de actividades conscientes para obtener el significado. Para todo ello parten de una definición operativa de tres niveles de competencia y de la búsqueda de procedimientos de instrucción acordes con esos tres niveles.

Este trabajo de Sánchez Miguel se propone completar los resultados obtenidos por él mismo en una investigación anterior, que consistía en poner a prueba dentro de las aulas la utilidad de ciertos programas de instrucción en la comprensión de textos que habían sido estudiados a nivel teórico y aplicados fuera del contexto escolar normal. Este nuevo trabajo (1989) pretendió completar y dar continuidad

al primero, por lo que su objetivo consistió en establecer una relación entre el rendimiento escolar y los tres niveles de competencia que se definirían inicialmente.

Con respecto a la investigación de Alonso Tapia, también abunda en el tema de las evaluaciones coordinadas por Sánchez Miguel. Sin embargo, su pretensión radica en el desarrollo y valoración de un programa de entrenamiento y regulación de la comprensión lectora. Se centra en el análisis del papel de la metacognición y del razonamiento en la comprensión lectora, así como en el modo de facilitar el desarrollo metacognitivo para mejorar el aprendizaje. En este sentido, se plantean como objetivos los siguientes:

1. El diseño de programas de entrenamiento en habilidades metacognitivas en relación con la comprensión lectora y el razonamiento. Las habilidades consideradas en este caso son: supervisión, toma de conciencia y autorregulación de la propia actividad mental mediante el empleo de estrategias generales y específicas adecuadas. La comprensión lectora y el razonamiento se conceptualizan en términos de valoración crítica de la información de carácter argumental contenida en los textos.
2. Evaluar la eficacia de tales programas a través de la comparación con otros de corte más tradicional, a partir de distintas medidas de comprensión lectora y de razonamiento y de la transferencia de los resultados a otras variables.

Así pues, además de su carácter sumativo (persigue la estimación del valor de cada programa para un propósito, y seleccionar aquellos más eficaces) el objetivo de esta investigación es doble. Por una parte se trata de evaluar la viabilidad del entrenamiento de las capacidades metacognitivas y su repercusión sobre la generalización y transferencia de los procesos de razonamiento inductivo y deductivo que tratan de entrenarse, y sobre el desarrollo cognitivo en general. Por otra, este objetivo clave desde el punto de vista teórico, acompaña a un objetivo final de carácter aplicado: el desarrollo de estrategias, materiales y procedimientos de instrucción concretados en un curso que, en la medida en que resulte validado, pueda ser útil en el ámbito escolar, ayudando a prevenir el fracaso o contribuyendo a compensarlo. Y es que la demostración de la viabilidad del entrenamiento metacognitivo y de su efectividad, podría contribuir, sin duda, a reorientar los procesos de enseñanza.

En orden a la consecución de estos objetivos, hasta el momento han sido desarrollados dos programas de entrenamiento: uno principal en el que el procedimiento instructivo posee la orientación metacognitiva que propugnan los investigadores, y otro paralelo, semejante en cuanto a estructura y contenido, pero con el "modelado" como base de la instrucción. El interés de este segundo programa estriba en la posibilidad de obtener datos comparativos más rigurosos y precisos sobre la eficacia relativa del procedimiento aplicado en el programa principal. Además se han diseñado cuatro técnicas de evaluación para medir su impacto.

Por su parte, los objetivos de la investigación coordinada por Garrido Gil, se alejan bastante de los perseguidos por los estudios centrados en las técnicas instructivas de la comprensión lectora. Este equipo de expertos diseña y pone en marcha un programa sobre habilidades cognitivas en el área de Matemáticas.

Suponen que las Matemáticas cada vez más van formando parte del repertorio estratégico que todo sujeto necesita para su desenvolvimiento cotidiano. En este sentido la solución de problemas matemáticos constituye una estrategia cada vez más necesaria que nos informa del uso que un sujeto puede darle a su capacidad matemática.

Dentro de la Psicología de la instrucción, hay un movimiento que trabaja en la enseñanza de habilidades cognitivas a través del currículum. Es este el marco en el que se ubica este trabajo, en un intento por encontrar respuestas a problemas planteados desde la realidad educativa, tales como el desarrollo de la capacidad lógica, la eficacia y dominio del vocabulario, la resolución de problemas, etc. La investigación se encuentra dividida en dos partes: una teórica y otra empírica. La primera consiste en un estudio sobre las bases teóricas de la Psicología de la instrucción en el área de las Matemáticas, tras lo que se propone un modelo de intervención con el que, posteriormente, se pretenden corroborar los supuestos teóricos en la práctica, que es precisamente el objeto del trabajo empírico.

Por otra parte, a tenor de lo que se desprende de los objetivos y del proceso de esta evaluación, su finalidad es formativa, al perseguir la orientación de la adopción de decisiones de mejora del programa.

El equipo integrado por Alicia Rivera Otero y el resto de profesores adscritos al C.E.P. de Ciudad Lineal (Madrid) centran su interés en el estudio de diversas metodologías de enseñanza-aprendizaje en el aula. El objetivo principal de su estudio es la evaluación comparada de

dos métodos didácticos. Un método, el tradicional, está basado en la transmisión de conocimientos como productos ya elaborados y cerrados por parte del profesor, que el alumno debe memorizar de forma repetitiva. Frente a éste, se dispone un método alternativo centrado en la investigación del alumno, de corte constructivista, donde aquél se constituye en verdadero artífice del aprendizaje y descubre los conocimientos por sí mismo llegando a una memorización y a una asimilación comprensivas. Esta evaluación comparada de dos métodos tiene una finalidad sumativa, puesto que de sus conclusiones se derivará la preferencia por uno de los dos métodos que se comparan.

Los objetivos que persigue la investigación coordinada por Martínez Martín son el diseño de módulos curriculares de educación moral, la organización de la escuela como comunidad democrática, y la elaboración de programas de formación de los docentes en estos aspectos curriculares. De todo ello se desprende una intención formativa muy clara, puesto que la valoración del programa tiene por objeto mejorarlo al mismo tiempo que se construye, así como mejorar la propia enseñanza-aprendizaje de la educación moral.

Por su parte, la evaluación emprendida por Mesa Franco y su equipo estaba orientada a la detección de las deficiencias que presentan los alumnos/as de extracción musulmana con lengua materna *chelja*, en su dominio de la lengua castellana oral y escrita, y la influencia de éste en el rendimiento de sus tareas escolares durante el primer curso de la Educación Primaria. También operativizan otros objetivos:

1. Conocer la naturaleza y peculiaridades del bilingüismo *chelja*-castellano de los escolares musulmanes de Melilla.
2. Conocer y valorar la influencia de la escolarización en la comprensión verbal y en el dominio del lenguaje oral y escrito castellano en estos niños bilingües.
3. Conocer y valorar la influencia de la escolarización en Educación Infantil o Educación Preescolar en el rendimiento escolar y en el dominio del lenguaje de estos niños bilingües en el primer curso de E.G.B. o de Educación Primaria.

A partir de este diagnóstico previo se desarrolla y evalúa un programa de intervención educativa para compensar y mejorar el dominio lingüístico y el rendimiento escolar de estos niños. Como parece desprenderse de lo anterior, así como de las conclusiones del estudio, la

finalidad de la evaluación es de tipo diagnóstico en primer término, así como propedéutico.

Dimensiones evaluadas y momento de la evaluación

Aunque en todas las investigaciones de este grupo se consideran dimensiones bastante distintas, sin embargo también todas ellas coinciden en considerar variables de entrada como parte del análisis previo en el que se intenta homogeneizar las muestras que se utilizan.

La evaluación del programa diseñado por Sánchez Miguel orientado a la instrucción en la comprensión de textos expositivos, es de tipo global, al contemplar las siguientes dimensiones en su diseño:

- De entrada, puesto que se agrupa a los sujetos en función de unas características que los homogenizan.
- De contexto, al someter a todos los alumnos al programa en circunstancias similares.
- De proceso, detectándose los logros y las dificultades que se encuentran los sujetos a lo largo de la aplicación del programa.
- De producto, ya que se evalúan los resultados que han alcanzado los sujetos tras seguir el programa de comprensión de textos expositivos.

Con respecto al momento en que se lleva a cabo la evaluación cabe decir que se trata de un estudio en tres momentos: una evaluación inicial en la que se somete a los alumnos de E.S.O. a un pretest; una evaluación procesual que ha servido para perfilar qué aspectos eran más o menos convenientes en orden a conseguir los objetivos propuestos y poder actuar en consecuencia (de ahí su carácter formativo) y, por último, una evaluación final que permitió estimar la eficacia del programa experimentado a través de un postest.

El programa de entrenamiento metacognitivo para la mejora de la comprensión lectora y los procesos de razonamiento, diseñado por Alonso Tapia y su equipo (1991), consiste en una evaluación parcial (según el modelo presagio-producto) en la que se consideraron variables de entrada, dado que se someten a un pretest a los sujetos que van a experimentar los distintos programas, y se midieron variables de producto para detectar los logros de cada uno de los sujetos sometidos

a cada condición de instrucción: método tradicional, método con modelado y el programa de entrenamiento metacognitivo.

La evaluación del programa sobre desarrollo de habilidades cognitivas en el área de las Matemáticas, elaborado por el equipo de Garrido Gil, puede considerarse prácticamente global, al incluirse tres de las cuatro dimensiones tradicionales: las entradas, puesto que a todos los estudiantes implicados en la investigación se les aplicaron las mismas pruebas de aptitudes previas (test de factor "G" de Cattell, escala 2 A de aptitudes académicas y de competencia matemática); de proceso, detectando los logros y problemas que va produciendo sobre la marcha la aplicación del programa en relación con la competencia en la resolución de problemas; y de producto, analizando la situación de los alumnos en el postest en las mismas variables iniciales.

La evaluación del programa de enseñanza de las Ciencias Sociales centrado en la investigación del alumno y realizada por Rivera Otero y su equipo, puede definirse como parcial, ya que se sólo se consideraron variables correspondientes a las dimensiones de entrada y de producto (midiéndose los resultados obtenidos por los alumnos del grupo experimental y de control antes y después del proceso de enseñanza de una unidad didáctica de Ciencias Sociales).

El diseño y la evaluación de programas de educación ética coordinado por Martínez Martín (1991) puede considerarse una evaluación parcial (exclusivamente del contenido de un programa), y que no se ciñe a un diseño cuasi-experimental con pretest y postest, sino que extrae aquellas estrategias válidas que inspiran los distintos constructos teóricos sobre educación ética existentes, para diseñar a partir de ahí un programa propio.

La evaluación del programa de intervención educativa para compensar y mejorar el dominio del castellano en los escolares bilingües de Melilla (Mesa Franco y colaboradores), también puede ser definido como una evaluación cuasi-global, al contemplar variables de tres dimensiones, faltando la de proceso. Se miden pues, variables de entrada que definen a cada grupo de sujetos seleccionados (unos de origen beréber y bilingüe, y otros de origen europeo y monolingüe). Se analizan asimismo aquellas variables de contexto que pueden ser determinantes, tales como su extracción socioeconómica y cultural, para, por último, estimar cuáles han sido los resultados tras la intervención educativa.

Patrones empleados para la emisión de juicios de valor

Cinco de las seis evaluaciones que se han incluido en este grupo son criteriosales, ya que los resultados de los distintos programas se contrastan con una instancia de referencia o patrón objetivo previamente definido por los responsables del programa.

Sin embargo, el programa de entrenamiento metacognitivo de Alonso Tapia (1992) utiliza varias instancias de referencia al mismo tiempo; una referencia criterial relativa a los objetivos previamente definidos en cada programa; y una referencia normativa por cuanto se evalúan los resultados de unos programas comparándolos con otros, difiriendo entre sí en las condiciones de instrucción. Así, las comparaciones se establecen entre el método tradicional, el método sólo con modelado y el propio programa de entrenamiento metacognitivo.

— *Diseño y desarrollo de las evaluaciones*

Se ha de dejar patente que, si bien en contadas ocasiones los investigadores explicitan el diseño evaluativo que adoptan en sus estudios, éstos pueden identificarse con relativa facilidad.

Cinco de los seis estudios que se han incluido en este grupo adoptan diseños experimentales o, más concretamente, cuasi-experimentales. Sólo una investigación, dirigida por Martínez Martín, responde más a un diseño cualitativo, a pesar de utilizarse también técnicas como la encuesta, que aportan datos cuantitativos relevantes para el análisis.

En la evaluación coordinada por Sánchez Miguel y su equipo centrada en probar la utilidad de programas concretos de instrucción en comprensión de textos, se comparan los resultados obtenidos por el programa diseñado a tal efecto con los obtenidos por el grupo de control. El estudio se realiza con muestras de estudiantes pertenecientes a 6.º y 7.º de E.G.B. —final de la Educación Primaria y principio de la Secundaria Obligatoria—, adoptándose un diseño pretest-postest para controlar variables extrañas, y comprobar las mejoras debidas a la eficacia del programa experimental. En concreto se llevó a cabo un análisis de:

- 1) La relación entre el nivel de competencia inicial de los sujetos en la comprensión de textos y el tipo de programas de instrucción que requieren.

- 2) La interdependencia entre las preguntas sobre el texto que se plantean a sí mismos los sujetos y el reconocimiento de las características organizativas de los textos.
- 3) La influencia que sobre la composición escrita ejerce la instrucción en el reconocimiento de la organización de los textos.

La evaluación contemplaba:

- Preguntas referidas a las ideas periféricas.
- Preguntas referidas a algún detalle concreto del texto, aunque remitidos a ideas centrales.

Las fuentes de información utilizadas para recabar los datos relevantes para el análisis han sido directas.

Por otro lado, en la investigación dirigida por Alonso Tapia los instrumentos que sirvieron para recoger información relevante fueron diseñados por los propios evaluadores: Prueba básica de razonamiento deductivo, Prueba de razonamiento condicional, y Test D-A.

La investigación de Garrido Gil coincide con las anteriores en la metodología adoptada y en los planteamientos evaluadores que en ella se aplican. En primer lugar, se lleva a cabo un estudio de las bases teóricas de la Psicología de la Instrucción en relación con la solución de problemas de Matemáticas. En segundo lugar, se realiza un estudio empírico cuyo objetivo es evaluar la eficacia del programa.

El estudio tiene un diseño experimental para el contraste de la hipótesis sustantiva que afirma que el programa diseñado sobre habilidades y estrategias de pensamiento aumenta la competencia matemática y la eficacia en la solución de problemas, los procesos de razonamiento y las aptitudes escolares.

El estudio es longitudinal teniendo lugar a lo largo de dos cursos, con 90 alumnos de 6.^º de E.G.B. de dos centros públicos de Murcia distribuidos en un grupo experimental y otro de control. Al primero se le aplica un programa para favorecer el desarrollo de habilidades cognitivas en el aprendizaje matemático, basado en los materiales del Proyecto Instrumental Harvard y en el Programa de Enriquecimiento Instrumental de Feuerstein; el grupo de control siguió la metodología que se venía empleando.

Las variables incluidas en el diseño son: la inteligencia general (medida mediante el test de factor "G" de Cattell, Escala 2-A); las ap-

titudes académicas (verbal, de razonamiento y de cálculo, evaluadas mediante el Test de Aptitudes Escolares de Thurstone, Escala 2), la competencia matemática, y el nivel de adquisición de habilidades y estrategias de resolución de problemas (evaluadas, respectivamente, con una Prueba de Competencias Matemáticas, elaborada al efecto y con unas pruebas específicas de resolución de problemas).

La evaluación siguió las siguientes fases: entrenamiento del profesorado implicado, evaluación (pretest) de la situación de los sujetos en las variables definidas, aplicación del Programa de Habilidades Cognitivas, evaluación del proceso de aplicación del programa en relación con la competencia para la resolución de problemas, y análisis de la situación en el postest en las mismas variables, utilizando idénticas pruebas a las de la situación pretest.

En el estudio de Alicia Rivera Otero y su equipo (1992) se pretende poner de manifiesto que la metodología empleada en la enseñanza de las Ciencias Sociales influye en el rendimiento y en el autoconcepto del alumno, así como en el clima social del aula. Las variables objeto de estudio fueron, según su posición en el diseño, las siguientes:

Independiente:

- Estrategia didáctica, que consta de dos niveles: método tradicional de enseñanza y método basado en la investigación del alumno.

Dependientes:

- Información adquirida. Se comprueba mediante una prueba objetiva elaborada para tal fin.
- Nivel de razonamiento, medido mediante una prueba elaborada al efecto y valorada por el sistema de jueces.
- Autoconcepto. Se comprueba mediante la prueba de Piers-Harris.
- Clima social de aula, medido con la Escala C.I.E.S.-C.E.S.

VARIABLES CONTROLADAS:

- Conocimientos previos, apreciados mediante un pretest.
- Inteligencia general, medida con la prueba factor "G" de Cattell.

La muestra utilizada era incidental y estaba constituida por alumnos de 8.º de E.G.B. pertenecientes a tres centros. En cada uno de ellos se ha dividido a los alumnos en dos grupos: uno de control y otro ex-

perimental. En los grupos de control se ha utilizado un método expositivo tradicional y en los experimentales un método de enseñanza basado en la investigación del alumno dentro de un enfoque constructivista del aprendizaje escolar.

Todos los alumnos trabajaron sobre la misma unidad didáctica: "La revolución industrial", impartida en 15 sesiones tanto en el grupo de control como en el experimental.

Por su parte, la metodología adoptada en la evaluación del estudio de Martínez Martín a diferencia de las que se han analizado hasta el momento, no se ocupa en la comparación del rendimiento de distintos grupos sometidos a diferentes programas. Este diseño tiene un carácter más cualitativo e interpretativo. Esta investigación se caracteriza por el uso y validación de técnicas y recursos que permiten la creación y/o aplicación de un currículum de educación ética que aporte recursos informativos, de aprendizaje, y técnicas favorecedoras del desarrollo moral, en un paradigma investigador de carácter cualitativo, centrado en el modelo de investigación-acción. Se trata de proporcionar a los docentes unos recursos para que analicen su propia práctica y reflexionen sobre su acción, con la finalidad de que mejoren las situaciones de carácter pedagógico que se refieren a la educación ética.

El estudio propuesto por Mesa Franco se divide en dos partes; en primer lugar hay una fase de exploración para conocer la naturaleza y peculiaridades del biligüismo tamazight-castellano que presentan muchos de los escolares de Melilla de origen beréber, así como para valorar la influencia de la escolarización (tanto en E.G.B. como en Educación Infantil), y de la clase social, en la comprensión verbal del castellano de estos escolares. En una segunda fase, se pone en marcha un programa destinado a compensar y mejorar el dominio de esta lengua, lo que se llevó a cabo en un aula de Educación Infantil.

El proceso metodológico se caracterizó por una primera fase de recogida de datos para hacer un diagnóstico individual de la situación, a través de la exploración de las habilidades psicolingüísticas de los niños de origen beréber, en comparación con los niños de habla castellana. Para ello se utilizó una muestra de 285 sujetos, de los que 166 eran de origen beréber y bilingüe y 199 de origen europeo y monolingüe. Se eligieron tres cursos, uno de Educación Infantil y dos de Primaria para la selección de los sujetos: preescolar, primero y tercero de E.G.B. Como instrumento de recogida de datos se utilizó el Test de Illinois de Aptitudes Psicolingüísticas (I.T.P.A.).

Durante la segunda fase se trabajó con los alumnos de dos grupos de Educación Infantil, uno de las cuales sirvió de grupo experimental por lo que a lo largo de cinco meses con él se llevó a cabo el programa de intervención. Se realizaron medidas posttest y se constataron las posibles mejoras habidas en las variables dependientes: comprensión auditiva, asociación auditiva, expresión verbal e integración gramatical.

— *Difusión y uso de los resultados de la evaluación*

En cualquiera de las investigaciones que se han incluido en este grupo, se constata cómo los resultados se formalizan a través de un informe final que se circunscribe a los resultados obtenidos después de efectuar el estudio comparativo de rendimientos tras aplicar los programas diseñados.

De este modo, la evaluación dirigida por Sánchez Miguel se limita a enumerar los aspectos que contribuyen a optimizar los resultados obtenidos a partir de la puesta en marcha del programa diseñado. Los principales resultados obtenidos son los siguientes: los sujetos con bajo nivel de comprensión necesitan un análisis de las estructuras locales para descubrir la organización global, mientras que los sujetos con buen nivel de comprensión pueden comenzar los procesos de comprensión lectora a partir de la organización global del texto. Por otro lado, después de la aplicación de los programas, los sujetos de pobre comprensión aumentan notablemente su rendimiento, mientras que los de buena comprensión apenas aumentan su rendimiento. El proceso de instrucción en comprensión de textos favorece la mejor organización interna de las redacciones.

En términos generales, el estudio indica la aplicabilidad con resultados satisfactorios de los programas de instrucción en comprensión de textos a la situación habitual del aula en el contexto escolar normal.

A pesar de que no existe una mención expresa de los propios autores que garantice una generalización de su experiencia, podemos intuir que, a la vista del trabajo efectuado sobre programas para la instrucción en la comprensión de textos expositivos, esto servirá como marco de referencia de nuevas aplicaciones dentro de esta misma línea.

De modo similar proceden Alonso Tapia y sus colaboradores. Sus conclusiones muestran, por un lado la eficacia del entrenamiento explícito y sistemático en el desarrollo de habilidades de supervisión y

regulación de la comprensión, mediante la reflexión sobre los procesos intervinientes (metacognición), aunque no en todas las medidas consideradas en el estudio. Sin embargo, no aparece una relación clara entre los cambios en las estrategias metacognitivas entrenadas y la mejora en la comprensión; no obstante, sí es evidente una tendencia en esa dirección. Por otro lado, el entrenamiento metacognitivo aplicado a procesos de razonamiento resulta igualmente eficaz, aunque a veces no haya llevado a cambios psicológicamente relevantes (debido quizá a la no inclusión de contenidos del currículum escolar). Todo ello recomienda, por tanto, una modificación de dichos entrenamientos para su integración en el contexto más propiamente escolar. La investigación de Alonso Tapia supone un comienzo en el diseño y aplicación de programas de entrenamiento cognitivo, extrapolable a otras áreas del currículo.

Para Garrido Gil y su equipo de investigadores, los resultados no confirman su hipótesis en todos los aspectos estudiados. Así, en cuanto a los procesos de razonamiento (factor "g"), no se aprecian diferencias significativas entre los grupos experimental y de control, ni en las puntuaciones totales, ni en ninguna de las subescalas. Respecto a las aptitudes académicas, sólo en la subprueba de razonamiento aparecen diferencias significativas a favor del grupo experimental. En la prueba de competencia matemática, existen diferencias en la puntuación total y en casi todas las parciales. Para finalizar, se plantean unas conclusiones abiertas sobre perspectivas de futuro para la enseñanza de las Matemáticas a partir de los resultados de esa investigación.

El trabajo de Martínez Martín apunta a nuevos diseños de programas específicos para los distintos niveles educativos orientados a la educación moral. Los investigadores se proponen seguir trabajando con profesores/as en esta misma línea para confeccionar materiales curriculares válidos. Los propios autores conciben este trabajo como un ensayo sistemático sobre la utilización de técnicas y estrategias para la educación moral; el diseño definitivo de módulos curriculares y unidades didácticas es un trabajo a realizar en el futuro. Asimismo se incorpora un vídeo como ilustración práctica y didáctica de la enseñanza de conocimientos procedimentales y de contenidos en el aula, que se puede considerar un recurso muy importante para la divulgación de las conclusiones de esta investigación.

Este trabajo, ha sido galardonado con el premio de Pedagogía Rosa Sensat de 1989, y próximamente será objeto de dos publicaciones independientes pero autosuficientes: el diseño curricular por una

editora catalana, y la fundamentación teórica por la Editorial Luis Vives.

Por su parte, la evaluación realizada por Rivera Otero y su equipo (1992) ha obtenido las siguientes conclusiones:

- No existen diferencias significativas entre el grupo de control y el experimental en la cantidad de información adquirida por los alumnos/as.
- El método basado en la investigación del alumno favorece la comprensión de conceptos más que en el método tradicional.
- Existen diferencias significativas entre el grupo de control y el experimental en cuanto a la mejora del autoconcepto, siendo mejor en el grupo experimental.
- El clima social del aula se ha visto favorecido con la estrategia didáctica basada en la investigación del alumno.

Los resultados de la evaluación efectuada por Mesa Franco y su equipo se centran en las cuatro hipótesis planteadas:

- Las puntuaciones obtenidas en las pruebas de lengua por los niños monolingües son más altas que las de los niños bilingües. Las diferencias entre ambos grupos son mayores en las pruebas relacionadas con la comprensión y expresión verbal.
- Los niños bilingües de extracción socioeconómica más favorable obtienen mejores resultados que los niños bilingües de nivel socioeconómico bajo.
- La escolarización temprana (en Educación Infantil) así como la participación en los programas de intervención educativa para compensar deficiencias, inciden positivamente en el rendimiento de los alumnos. El grupo experimental mejora considerablemente sus niveles de competencia, tanto en comparación con su situación de partida, como cuando se compara con la evolución que ha tenido el grupo control.
- Además, otra conclusión relevante es que el interés de los padres aparece muy relacionado con los progresos y la rapidez en el dominio de las habilidades lingüísticas, por lo que recomienda a las familias aumentar su colaboración con la acción educativa de la escuela.

Aunque no se concreta el tipo de difusión que se pretende hacer de los resultados, de sus conclusiones se puede inferir el interés que tiene

emprender una estrategia educativa común entre el centro y la familia, que contribuya a la mejora y dominio de las habilidades lingüísticas.

a.2. Evaluación de programas instructivos centrados en el uso de medios y recursos didácticos

— *Objeto de la evaluación*

Tres son las investigaciones que conforman este grupo:

- *“Investigación acerca de los recursos sonoros como materiales didácticos en las áreas curriculares de Lengua/Literatura y Ciencias Sociales”* (Julio Carmona Cerrato, 1994).
- *“Eficacia didáctica de una metodología en el diseño de unidades didácticas de Física con ordenador”* (M.^a Carmen Pérez de Landazábal, 1991).
- *“Análisis de la eficacia del microordenador en la enseñanza de las Matemáticas en 7.º de E.G.B. y 2.º de B.U.P.”* (Antonio Bautista García-Vera, 1987).

Todas ellas evalúan unas determinadas metodologías que utilizan recursos didácticos específicos (nuevas tecnologías de la información y de la comunicación) como material de apoyo.

La primera de ellas (Carmona Cerrato) se ocupa de comprobar la eficacia de una serie de recursos sonoros en las áreas de Lengua y Ciencias Sociales. La población escolar sobre la que se realiza la investigación es una muestra de 270 alumnos de 5.º y 8.º de Enseñanza General Básica, de 1.º de B.U.P. y de 1.º de Formación Profesional, pertenecientes a centros rurales y urbanos de la provincia de Badajoz. La selección de los cursos se ha hecho teniendo en cuenta la equivalencia respecto a la nueva distribución de ciclos educativos que contempla la Reforma.

Pérez de Landazábal evalúa una metodología encaminada a la obtención de cambios conceptuales en el alumno que se enfrenta con la Física por primera vez, utilizando como medio didáctico el ordenador. La muestra está constituida por 443 alumnos: 258 de 2.º de B.U.P., 96 de 3.º de B.U.P. y 89 de C.O.U.

Por último, Bautista García-Vera evalúa cuatro métodos de enseñanza o estrategias didácticas. Concretamente, se trata de unos méto-

dos de enseñanza basados en la resolución de problemas y en la utilización del ordenador en el aprendizaje de las Matemáticas. Se experimentan cuatro estrategias de enseñanza, dos de ellas desarrolladas en medios computacionales (CAI-tutorial y resolución de problemas utilizando LOGO) y otros dos en ambientes no computacionales (clásico y de resolución de problemas sin apoyo informático).

— *Características definitorias de la evaluación*

Agentes evaluadores

Los estudios dirigidos por Pérez de Landazábal y García Vera se pueden considerar evaluaciones externas. Así, el primero de ellos cuenta con un equipo de investigación formado por profesionales adscritos al Consejo Superior de Investigaciones Científicas; mientras el de Bautista García-Vera está compuesto por profesores de Universidad, de C.P. y de Institutos.

Por otra parte, la investigación de Carmona Cerrato es una evaluación interna ya que son precisamente los responsables del programa (el director del Centro de Recursos y un profesor de apoyo de dicho centro), quienes llevan a cabo la evaluación.

Objetivos y funciones de la evaluación

El estudio de Carmona Cerrato se propone analizar la utilidad y eficacia que tiene para la promoción de determinados aprendizajes, el empleo de una metodología para el desarrollo curricular de las áreas de Lengua y Ciencias Sociales en los niveles de 5.º y 8.º de E.G.B., 1.º de B.U.P. y 1.º de F.P., metodología que se sustenta en la aplicación de una serie de recursos didácticos sonoros (lecturas ambientadas y programas de radio originales).

El equipo investigador basa su propuesta en el contraste de una conjetura formulada tras una larga experiencia docente en observar el tipo de problemas que presentan los alumnos en el proceso lector y en el aprendizaje de las Ciencias Sociales; por ello se emprende la comprobación de dicha hipótesis y, posteriormente, se elaboran y aplican unas pruebas para comprobar la utilidad de dichos recursos didácticos sonoros en la mejora del aprendizaje.

Con respecto al área de Lengua/Literatura se han elaborado modelos lectores que consisten en la grabación sonora de textos escritos,

adaptados a los alumnos, ambientados de forma adecuada con música y efectos especiales. En el área de Ciencias Sociales, se diseñan unidades didácticas bajo el formato de guión radiofónico, cuidando especialmente la calidad técnica del montaje.

La función de la evaluación es básicamente sumativa, puesto que se pretende, no la mejora de los propios recursos didácticos, sino su adopción o no en función de la eficacia demostrada en la enseñanza a través de un diseño del tipo pre/postest.

El objetivo principal del trabajo de Pérez de Landazábal es analizar si la Enseñanza Asistida por Ordenador, que incorpora la experimentación en el laboratorio, resulta eficaz en la modificación de las ideas previas que tienen los alumnos sobre ciertos conceptos físicos. Asimismo, se quiere comprobar si este tipo de enseñanza promueve el desarrollo de destrezas científicas y de actitudes favorables hacia el aprendizaje de la Física. Esto se concreta en los siguientes objetivos:

- Conocer los errores conceptuales más frecuentes en el alumnado.
- Diseñar unidades didácticas de Física y contrastar experimentalmente su validez para producir un cambio conceptual (evaluación de la eficacia didáctica). Dichas unidades didácticas pueden ser utilizadas como un recurso más en un entorno multimedia. También pueden emplearse para la recuperación de estudiantes con dificultades en conceptos concretos, dado que proporcionan al alumno la posibilidad de estudiar a su propio ritmo.
- Analizar la eficacia global de la metodología de la Enseñanza Asistida por Ordenador en función del rendimiento de los alumnos.

La evaluación cumple una función básicamente sumativa, ya que trata de comprobar al final de un período de aplicación, si el programa empleado mejora o no el aprendizaje de los conceptos de Física.

El objetivo principal del estudio dirigido por García-Vera es aportar una serie de orientaciones a los profesores de Matemáticas de E.G.B. y B.U.P. sobre las posibilidades del ordenador en la enseñanza de su asignatura. Se enmarca en el ámbito de la didáctica de las Matemáticas y su finalidad es analizar la eficacia de dos métodos de enseñanza en el aprendizaje de unos contenidos concretos de Geometría y

Cálculo, en la adquisición y desarrollo de estrategias para resolver problemas y de habilidades básicas para programar ordenadores. Los dos métodos estudiados son: a) el método tradicional centrado en la resolución de problemas sin utilizar el ordenador como recurso; y b) el experimental o enseñanza para la resolución de problemas asistida por ordenador.

Dada la función que cumple la evaluación del uso de este recurso didáctico en la enseñanza de las Matemáticas, se puede considerar una evaluación sumativa del mismo, más encaminada a la valoración de su pertinencia, que a la mejora de sus propias características sobre la marcha.

Dimensiones evaluadas y momento de la evaluación

Las tres investigaciones incluidas en este grupo siguen el modelo clásico de evaluación de input-output, por lo que las variables consideradas son las obtenidas en las medidas previas realizadas (variables de entrada) y en las medidas tomadas tras la aplicación de las metodologías o recursos didácticos que son objeto de evaluación. En algunos casos también se ha realizado un análisis del proceso de aplicación de las mismas que ha servido para matizar la interpretación de los resultados.

En concreto, para evaluar la utilidad del programa, el equipo de Carmona Cerrato llevó a cabo un diagnóstico inicial de la situación en el que se contemplaron variables de entrada referidas a los problemas de los alumnos en aspectos relacionados con el proceso lector: lectura silábica con falta de ritmo y entonación, deficiente velocidad lectora, bajo nivel de comprensión, o escaso gusto por la lectura.

Asimismo, se pretendía dar una respuesta didáctica a algunas deficiencias detectadas en el proceso de aprendizaje de las Ciencias Sociales: dificultad para desligarse de las experiencias concretas y alcanzar un nivel de abstracción suficiente, problemas relacionados con la conceptualización del espacio y del tiempo histórico, poco rigor en los planteamientos y escasa curiosidad científica, o pobre desarrollo de la capacidad de participar responsable y críticamente.

Tras aplicar la metodología de enseñanza diseñada al efecto, se realizó una evaluación final de los resultados obtenidos en una serie de pruebas que medían el rendimiento en las siguientes dimensiones de las dos áreas contempladas:

- *Área de Lengua:* comprensión lectora, velocidad lectora, vocabulario, imaginación creadora, asociación de palabras.

- *Área de Ciencias Sociales*: nociones espacio-temporales, conceptos básicos, técnicas de trabajo.

De este modo, se puede considerar que este estudio contempla las dimensiones de entrada, de proceso y de producto del programa metodológico que se experimenta.

El estudio de Pérez de Landazábal, como se ha mencionado, se centra en el cambio conceptual del alumno con respecto a los conceptos físicos de calor y de temperatura, en la actitud hacia la Física y en el desarrollo de destrezas científicas. Respecto al primer elemento del rendimiento del alumno, se consideraron como variables de entrada sus ideas previas respecto al calor y a la temperatura. Para ello se realizó una evaluación inicial anterior al uso de la metodología de Enseñanza Asistida por Ordenador. Tras este proceso, se realizó una evaluación final que trataba de estimar el cambio conceptual producido.

Respecto a la actitud hacia la Física, también se llevaron a cabo una evaluación inicial y final recabando la opinión del alumno sobre el trabajo realizado con el ordenador, así como su opinión sobre otros aspectos generales: gusto por las clases, afición por los programas científicos de televisión, afición por la realización de experimentos científicos, elección de la Física como optativa en 3.º de B.U.P. y C.O.U., autoestimación como estudiante de Física, o dificultad percibida de la asignatura.

Por último, la adquisición de destrezas científicas se analiza antes y después de aplicarse la metodología en cuestión, analizando la capacidad de los alumnos para controlar variables en la contrastación de hipótesis científicas.

A la vista de las medidas tomadas, se puede decir que se trata de una evaluación de entradas y de producto, en la que se controlan algunas variables de proceso.

Igualmente, en la investigación de García-Vera para medir el efecto sobre el rendimiento en Matemáticas de las cuatro estrategias didácticas, se lleva a cabo una evaluación inicial y final, como se deriva del diseño pretest-postest empleado.

Patrón empleado para la emisión de juicios de valor

Las investigaciones que se pueden considerar normativas son aquellas en las que el criterio de comparación es precisamente un grupo normativo: en este caso el grupo de control, al que no se le aplica el recurso, estrategia o metodología didáctica. Sin embargo, cuando el

criterio de comparación son las ideas, actitudes o el rendimiento previo de los alumnos (antes de recibir una nueva metodología de enseñanza), como ocurre en los diseños con un solo grupo y con medidas pretest-postest, puede considerarse que la evaluación utiliza una instancia de autorreferencia para interpretar los resultados. Este es el caso de las investigaciones dirigidas por Bautista García-Vera y por Carmona Cerrato. La primera de ellas, con un diseño pretest/postest, mide el rendimiento de los alumnos (concretamente el aprendizaje de conceptos matemáticos y el desarrollo de estrategias de resolución de problemas) antes y después de aplicar las cuatro estrategias de enseñanza cuya eficacia se quiere comprobar. Asimismo, Carmona Cerrato somete a comprobación experimental una metodología centrada en una serie de materiales didácticos, mediante la aplicación de pruebas y tests antes y después de utilizar dicha metodología.

Por su parte, en el estudio dirigido por Pérez de Landazabal se llevan a cabo dos tipos de evaluaciones: una evaluación criterial, analizando si se había producido algún cambio conceptual en cada uno de los alumnos; y una evaluación normativa al incluir en su diseño un grupo de control y un grupo experimental.

— *Diseño y desarrollo de la evaluación*

Las tres estudios de este grupo adoptan un modelo de investigación con medidas pretest/postest y con un enfoque metodológico cuantitativo.

Carmona Cerrato en su trabajo "*Investigación acerca de los recursos sonoros como materiales didácticos en las áreas curriculares de Lengua/Literatura y Ciencias Sociales*" trata de someter una serie de materiales didácticos a la oportuna comprobación con objeto de contrastar la hipótesis de partida. La selección de los sujetos se realiza con un muestreo estratificado en función de los siguientes criterios: entorno escolar (centros del área rural y del área urbana), y nivel educativo (E.G.B., B.U.P. y F.P.).

Las fases de la investigación las resume el equipo investigador como sigue:

- Aplicación de pruebas para la detección de dificultades de aprendizaje en las dos áreas objeto de estudio (Lenguaje y Ciencias Sociales).

- Baremación de las pruebas aplicadas con la finalidad de conocer la situación inicial o el nivel de partida de los alumnos.
- Confección del material curricular, es decir, de los recursos sonoros (lecturas ambientadas y guión radiofónico original) que conforman la metodología con la que se trabajó con las áreas mencionadas.
- Entrenamiento del profesorado en el conocimiento de dicha metodología.
- Aplicación de la misma.
- Análisis de los resultados finales para conocer la incidencia que el material empleado y la nueva metodología habían producido.

Para fijar los niveles de partida de los 270 alumnos de la muestra se aplicaron las siguientes pruebas: de comprensión lectora (el test de Lázaro Martínez y el test "Cloze" de Suárez y Meara, a alumnos de 5.º de E.G.B.); de velocidad lectora (en 5.º de E.G.B. se aplicó la prueba de Pérez González sobre la lectura de un texto de 295 vocablos, mientras que en 8.º de E.G.B. se aplicó la prueba de M.A. Furones sobre la lectura de un texto con 652 vocablos); de vocabulario (Test de García Hoz en 5.º y 8.º de E.G.B. y en 1.º de B.U.P., con sus respectivas baremaciones).

La imaginación creadora se analizó a través de la existencia o inexistencia de interacción entre las cuatro variables consideradas (comprensión y velocidad lectora, vocabulario e imaginación). Por último, la asociación de palabras se midió con el Test de García Hoz fundamentado en la cantidad de vocablos generados por una palabra-estímulo (se aplicó en 5.º y 8.º de E.G.B.).

En cuanto al área de Ciencias Sociales y en lo que se refiere a las nociones temporales de los alumnos de 1.º de B.U.P. y de 1.º de Formación Profesional, se adoptó la prueba de J.L. Sánchez y R.M. Urea para comprobar la madurez de la capacidad de abstracción temporal.

Los conceptos básicos se analizaron a partir de una prueba elaborada al efecto, que consistió en un cuestionario cerrado compuesto por 60 ítems de elección múltiple de respuesta. Asimismo, los investigadores elaboraron un cuestionario sobre técnicas de estudio basado en una selección de trabajos propuestos en los libros de texto de Ciencias

Sociales. Su objetivo era fijar el nivel de dominio que tenían los alumnos de los procedimientos más necesarios para el aprendizaje de esta materia.

El análisis de los datos fue de tipo estadístico (descriptivo y correlacional), pero también se hizo un análisis cualitativo de la información que la observación y el control del proceso permitieron obtener. La interpretación de los resultados de rendimiento de los alumnos se hace en términos más valorativos que explicativos.

El diseño del estudio de Pérez de Landazábal es de tipo experimental y con carácter longitudinal (realizado a lo largo de dos cursos escolares), con grupo de control, y muestreo aleatorio simple. Como ya se apuntó, se han realizado dos tipos de evaluaciones:

- Una evaluación referida a criterio en la que se comprobó si se había producido algún tipo de cambio conceptual en los alumnos. Se trata en este caso de un diseño pretest/postest.
- Una evaluación normativa con un diseño de dos grupos (control y experimental). También aquí la variable dependiente fue el cambio conceptual experimentado por los alumnos.

Se utilizó el cuestionario como técnica para recoger la información tanto en lo referente al cambio conceptual (medida pre/postest) como en el análisis de las actitudes hacia la Física, mientras que para evaluar las destrezas científicas se aplicó una prueba específica. El análisis de los resultados es fundamentalmente de tipo estadístico-descriptivo.

Igualmente, la investigación de Bautista García-Vera es un estudio experimental con diseño pretest/postest.

La muestra está dividida en dos grupos según el nivel educativo: 193 alumnos de 7.º de E.G.B., y 454 alumnos de 2.º de B.U.P. A su vez, se han subdividido cada una de estas muestras en cuatro grupos por ser cuatro los métodos de enseñanza analizados en la investigación. Se llevó a cabo un muestreo sectorial por nivel educativo (7.º de E.G.B. y 2.º de B.U.P.), se utilizaron fuentes directas de información, y se realizó un análisis de la misma con las técnicas estadísticas adecuadas.

Las variables contempladas, según su función en el diseño son las siguientes:

- *Variables independientes*: las cuatro estrategias didácticas: a) tradicional; b) resolución de problemas sin ordenador; c) ense-

ñanza asistida por ordenador mediante programas tutoriales; d) resolución de problemas en ambientes computacionales utilizando el lenguaje LOGO.

- *Variable dependiente*: rendimiento de los alumnos en Matemáticas. Se operativiza en las siguientes adquisiciones: a) aprendizaje de conceptos matemáticos sobre Geometría y Aritmética (se comprueba mediante la aplicación de unas pruebas *ad hoc*); b) estrategias de resolución de problemas (se observan a través de unos protocolos contruidos para tal fin); c) introducción a la programación de ordenadores (mediante pruebas en las que los alumnos tienen que confeccionar ordinogramas para resolver con ordenador una serie de problemas propuestos).
- *Covariables* que pueden incidir en el rendimiento: conocimientos previos (medidos con un test elaborado para este fin); inteligencia (Test "factor g" de Gattell); razonamiento numérico (Prueba N.A. del D.A.T.); aptitud espacial (escala S.R. del D.A.T.); razonamiento abstracto (escala R.A. del D.A.T.); percepción de diferencias (test de "Formas idénticas"); tener/no tener ordenador en casa (ficha de datos personales). También se contemplaron otras covariables de interés como la actitud y predisposición hacia el uso del ordenador.

Para evaluar los conocimientos de los alumnos antes y después de emplear las estrategias didácticas, se elaboraron dos pretest y dos posttest para cada una de las variables dependientes y posteriormente se realizó un análisis estadístico descriptivo, de forma global por método de enseñanza, y de forma conjunta por niveles (7.^º de E.G.B. y 2.^º de B.U.P.). Asimismo, se llevó a cabo una interpretación descriptiva de los resultados en función de la suerte corrida por las hipótesis de partida.

— *Difusión y uso de los resultados de la evaluación*

La investigación de Pérez de Landazábal supone una interesante revisión bibliográfica sobre las dificultades de enseñanza y aprendizaje del concepto de "calor", y se puede decir que es ésta la principal aportación del estudio. Los investigadores no encontraron diferencias significativas en los cambios conceptuales producidos entre los alumnos del grupo de control y del grupo experimental.

Por otra parte, los resultados de las pruebas finales del estudio de Carmona Cerrato demuestran que la utilización de los recursos sonoros diseñados, mejoraban los aprendizajes en las áreas de Lengua y de Ciencias Sociales; se confirma pues la hipótesis inicial de que los recursos sonoros aplicados como materiales didácticos de apoyo, favorecen una motivación especial y un proceso de aprendizaje capaz de paliar las deficiencias existentes en las dos áreas curriculares estudiadas. Asimismo, hay que destacar el esfuerzo realizado en la elaboración de los materiales curriculares, tanto en su soporte exclusivamente auditivo como en los soportes gráficos que acompañan a los textos. En este sentido, y para el área de Lengua, se han recopilado un total de 25 textos dirigidos a 5.º de E.G.B., y 10 textos para 8.º de E.G.B. En el caso de la Ciencias Sociales se elabora un guión radifónico con distintas unidades didácticas, que constituyen una interesante aportación para el estudio de los hechos históricos. Se incide en la conveniencia de que experiencias similares puedan ser realizadas por los maestros en sus aulas.

Por último, de la evaluación realizada por Bautista García-Vera y su equipo para analizar la eficacia de distintos métodos de enseñanza en el aprendizaje de unos contenidos de Geometría y Cálculo y en la adquisición y desarrollo de estrategias para resolver problemas y para programar ordenadores, se desprende que la educación no puede rehuir la utilización de un instrumento como el microordenador que amplía enormemente las posibilidades de aprendizaje.

a.3. Evaluación de programas instructivos contruidos sobre modelos teóricos definidos

— Objeto de la evaluación

En un intento de definir el marco en el que se ubican las investigaciones integradas en la categoría de evaluación de programas instructivos diseñados sobre determinados modelos teóricos, parece lógico contextualizar su naturaleza y perfilar el alcance de las mismas.

Las evaluaciones de programas que se han integrado en este grupo son aquellas que toman como punto de partida un modelo específico y que, apoyadas en determinados constructos teóricos, han originado el diseño de programas, unidades, instrumentos didácticos, etc., los cua-

les han sido sometidos a un proceso de evaluación de su efectividad en el transcurso de su aplicación.

Si bien es cierto que no cabe agruparlas en función de la fundamentación teórica que las sustenta dada su diversidad, sí puede hacerse por la existencia de elementos comunes que, desde el punto de vista metodológico, van a poner de manifiesto la eficacia de los programas que se pretenden evaluar de un modo u otro.

Las evaluaciones de programas que conforman este grupo son las siguientes:

- *Enseñanza de la Literatura española contemporánea en el Ciclo Superior de la E.G.B., tomando como base el método de tipos (Miguel Delibes), y el Sistema de Unidades Modulares de Aprendizaje (Alvarez González, 1986).*
- *Entrenamiento de habilidades cognitivas y enriquecimiento motivacional: Nuevas Tecnologías para la educación compensatoria (Alonso Tapia, 1986).*
- *Estudio experimental basado en la Teoría de Lakatos acerca del progreso de la Ciencia y sus implicaciones didácticas (Ramos Fernández, en curso).*
- *Diseño y evaluación de una propuesta curricular de aprendizaje de la Geometría Plana basada en el Modelo de Razonamiento de Van Hiele (Gutiérrez Rodríguez, 1991).*

La tónica predominante de estos estudios es el carácter restringido de las evaluaciones que se llevan a cabo, en tanto se trata de programas de una implantación acotada y suscrita a ámbitos muy reducidos. El alto grado de experimentación que parecen compartir todas ellas, condiciona en cierto modo su generalización, dado que dichas experiencias parten de unos postulados que implican unos requisitos previos muy concretos.

— *Características definitorias de la evaluación*

Agentes evaluadores

La investigación de Alvarez González (1986), que tiene por objeto evaluar un programa de enseñanza individualizada de la Literatura española contemporánea (el Sistema de Unidades Modulares de Apre-

dizaje) en el primer ciclo de la Educación Secundaria Obligatoria, consistió en una evaluación interna del programa al ser efectuada por los propios profesores expertos en la aplicación de la metodología S.U.M.A. quienes, en calidad de experimentadores, deseaban poner de manifiesto la existencia de diferencias significativas en el aprendizaje de los sujetos sometidos a las distintas metodologías. En esta evaluación el equipo ha contado con el asesoramiento externo de un experto en el tema (el profesor Fernández Huerta de la U.N.E.D.).

Por su parte, el equipo de investigadores que llevó a cabo el trabajo centrado en el *entrenamiento de habilidades cognitivas y enriquecimiento motivacional*, desarrollado en el ámbito de la educación compensatoria, fue coordinado y dirigido por el profesor Jesús Alonso Tapia (1986). Dicho equipo estaba integrado por expertos en evaluación, cuya trayectoria investigadora se enmarca dentro de una línea definida de la evaluación de la inteligencia, el pensamiento conceptual, la motivación, etc. Por eso, al igual que la anterior, esta evaluación se puede considerar interna, dado que son los propios investigadores los que proponen, llevan a efecto y evalúan el programa.

En la tercera de las evaluaciones contempladas en este grupo (Ramos Fernández), iniciada en 1991 y actualmente en curso, se presenta una propuesta teórica y experimental derivada de un modelo sistémico de evolución conceptual.

Dicha evaluación se propone concluir que, a nivel grupal, los alumnos/as, lejos de aprender necesariamente el esquema conceptual formalmente correcto que se les pretende enseñar, retienen los conceptos estructurándolos en una diáspora de esquemas, lo cual se constata a través de la medición del incremento del parámetro "dispersión conceptual". Esto tiene lugar como consecuencia de la relativa inconsecuencia de los esquemas que "construyen" los alumnos, puesto que se percibe una cierta ausencia de concordancia entre lo que los estudiantes definen de modo explícito (lo que verbalizan) y lo que aplican implícitamente.

Los investigadores que evalúan la propuesta curricular orientada a propiciar el aprendizaje de la Geometría Plana (Gutiérrez Rodríguez), aunque dirigidos desde el departamento de Didáctica de las Matemáticas de una Escuela de Formación del Profesorado, son todos ellos profesores de Enseñanzas Medias. Esto hace que la evaluación que llevan a cabo sea interna, puesto que son los propios diseñadores de la propuesta curricular, quienes tratan de estimar el grado de efectividad de la misma, aunque cuenten con asesoramiento externo.

Objetivos y funciones de la evaluación

Alvarez González (1986), desde el discurso de la enseñanza individualizada, transcribe sus principios a una de las metodologías didácticas que asumen sus presupuestos teóricos básicos: el Sistema de Unidades Modulares de Aprendizaje (S.U.M.A.). El contexto teórico de la investigación se fundamenta en la enseñanza a medida, partiendo de la convicción de que los sujetos poseen unas características distintivas que constituyen su individualidad, y de que la enseñanza debe procurar atender a esas diferencias.

Se desea comprobar la eficacia del S.U.M.A. cuando se utiliza una Unidad Modular-Tipo, en el área de Lengua de 7.º de E.G.B. (1.º curso del 1.º ciclo de la Educación Secundaria Obligatoria).

El objetivo principal de la evaluación era estimar la eficacia del método S.U.M.A. a través de los resultados de los aprendizajes de los alumnos que lo seguían, en comparación con aquellos que seguían una metodología convencional.

La intención de los investigadores parece orientarse a la adopción de la metodología S.U.M.A. en la enseñanza a distancia de determinadas áreas del conocimiento, en la medida en que mejore el rendimiento de los alumnos y quede constancia de su eficacia respecto de la enseñanza tradicional. En este sentido se puede considerar que se trata de una evaluación sumativa que trataba de facilitar la decisión de adopción o no del programa.

Tras la organización de un módulo literario adoptando el esquema heurístico del S.U.M.A., y en función del rendimiento lingüístico-literario de los alumnos, se pretende constatar la eficacia del programa. Además, al determinar las denominadas áreas correctivas o de recuperación del alumno, se propicia, a juicio de los autores, un sistema metodológico útil y eficaz para la enseñanza a distancia.

La investigación de Alonso Tapia partió de presupuestos distintos a la anterior. Tenía como objetivo prioritario crear un conjunto de materiales, técnicas y procedimientos integrados en un curso para propiciar el enriquecimiento cognitivo y motivacional de los alumnos del Ciclo Superior de la E.G.B., con especial hincapié en aquellos que tienen unas necesidades educativas especiales por pertenecer a un colectivo desfavorecido culturalmente. Además pretendían evaluar la efectividad de los programas disponibles a tal efecto.

De los objetivos de esta investigación evaluativa dirigida por el profesor Alonso Tapia se desprende una doble orientación, formativa

y sumativa. Formativa porque no renuncia a la mejora de alguno de los programas, pero con un carácter sumativo, en la medida que al final de la aplicación de la metodología se realiza el balance de la eficacia de los instrumentos considerados mejores para el entrenamiento de habilidades cognitivas y el enriquecimiento motivacional. En función de esto se afirma la idoneidad o no de cada programa para los propósitos deseados. Además, se efectúa una evaluación preliminar selectiva (sumativa por tanto) de todos los programas de enriquecimiento existentes, contrastando los resultados parciales obtenidos en investigaciones precedentes, tras lo que se eligen sólo aquellos seis que parecen ofrecen una mayor garantía con respecto al objetivo educativo prioritario de mejora del rendimiento y la motivación.

Los autores de la investigación coordinada por Ramos Fernández afirman, como primera hipótesis, que la relativa consecuencia o inconsecuencia de los esquemas conceptuales de los alumnos guarda un cierto paralelismo con el enfoque de la ciencia defendido por Imre Lakatos (1983). Este autor sostiene que el núcleo firme de una teoría científica se mantiene debido a la existencia de un “cinturón protector compuesto por ideas auxiliares” que la apuntalan y defienden de las discrepancias y evidencias experimentales en su contra.

Otras dos hipótesis de investigación sostienen que la resistencia de los alumnos al cambio conceptual correlaciona positivamente con el grado de protección de los conceptos previos que poseen, y que el cambio conceptual se produce cuando el alumno encuentra un esquema conceptual capaz de facilitarle la explicación de los problemas que con sus esquemas iniciales no lograba resolver. A modo de síntesis, los objetivos perseguidos en esta investigación son:

- Constatar hasta que punto la teoría de Lakatos puede aplicarse al ámbito del aprendizaje de conceptos científicos, intentando identificar y caracterizar posibles pautas acerca de cómo “protegen” los estudiantes sus esquemas conceptuales previos.
- Diseñar y valorar un modelo de intervención didáctica en el que, partiendo de las ideas anteriores, junto con las secuencias evolutivas de los patrones de aprendizaje de los sujetos y la consecuencia/inconsecuencia en la aplicación de los esquemas conceptuales.

Al igual que en la investigación dirigida por Alonso Tapia, en este caso la evaluación del programa persigue, tanto su perfeccionamiento

(función formativa), como la estimación de su valor para el propósito que se persigue (función sumativa).

Por su parte, la investigación coordinada por Gutiérrez Rodríguez se originó tras la constatación sistemática por parte de los profesores de los problemas de comprensión de los contenidos matemáticos que presentaban los alumnos de Secundaria. En un intento de salir al paso a este tradicional problema, se diseñaron, por un lado, unas unidades didácticas para la enseñanza de determinados contenidos de la Geometría Plana y, por otro, unos tests que sirvieran de instrumentos para evaluar el nivel de razonamiento en el que, según el modelo de Van Hiele, se ubican los alumnos antes y después de haber estudiado con el material de enseñanza objeto de experimentación. En este caso, la función sumativa de la evaluación de esta metodología es clara.

Dimensiones evaluadas y momento de la evaluación

La investigación de Alvarez González (1986) contempla variables de entrada, dado que mediante distintas pruebas (de aptitud, creatividad y adaptación social y escolar) se logra agrupar a los sujetos que van a formar parte del grupo experimental; y variables de producto o de los resultados obtenidos medidos en términos de valoración de los conocimientos adquiridos en función de la metodología a la que han sido sometidos. Se trata pues de una evaluación parcial, que responde al modelo presagio-producto, y de carácter final, puesto que es al concluir todo el proceso de aplicación de la metodología S.U.M.A. cuando tiene lugar.

La evaluación de los programas de habilidades cognitivas y enriquecimiento motivacional para la educación compensatoria (Alonso Tapia), es de tipo global puesto que el diseño contempla dimensiones de entrada (variables descriptivas de los sujetos), de contexto (procedencia sociocultural de los alumnos), del proceso del curso que se imparte, y de producto, puesto que se miden los resultados mediante unos cuestionarios que valoran el éxito de los programas. Su evaluación se produce en varios momentos: al inicio, cuando se comprueba la situación de los sujetos antes de ser sometidos al programa; durante la aplicación de los programas de habilidades cognitivas y de enriquecimiento motivacional, y al final, al llevarse a cabo medidas sobre los resultados del programa una vez concluido un período de aplicación.

El estudio experimental basado en la teoría de Lakatos acerca del progreso de la ciencia y sus implicaciones didácticas (Ramos Fernández, en curso), puede considerarse una evaluación global, puesto que en su diseño se propone la medición y control de variables iniciales, de contexto, de proceso —que servirán para el diseño de unidades didácticas para trabajar en la línea del afianzamiento y cuestionamiento de las ideas previas de los alumnos—, y de producto. La evaluación se lleva a cabo en tres momentos: al inicio, durante y al final de la aplicación del programa.

El diseño y evaluación de una propuesta curricular de aprendizaje de la Geometría Plana basada en el modelo de razonamiento de Van Hiele, coordinada por Gutiérrez Rodríguez, constituye una evaluación parcial al contemplar solamente variables de entrada y de producto. Además se trata de una evaluación final, dado que los resultados son medidos en términos que determinan el nivel de razonamiento de Van Hiele alcanzado por los estudiantes, al terminarse la aplicación de las unidades de aprendizaje de la Geometría Plana.

Patrón empleado para la emisión de juicios de valor

Al interpretar la información que se ha obtenido tras efectuarse las distintas evaluaciones que en este grupo se incluyen, se han empleado distintas instancias de referencia.

Al tratarse en este caso de un grupo de evaluaciones de programas instructivos contruidos sobre modelos teóricos definidos, se puede decir que todas ellas emplean preferentemente como instancia de referencia, criterios prefijados en los propios modelos teóricos, para poder interpretar y valorar la información recogida sobre los respectivos programas. En este sentido se puede decir que todas ellas llevan a cabo evaluaciones criteriosales.

— Diseño y desarrollo de la evaluación

En la investigación efectuada por Alvarez González se adopta un modelo de evaluación que incluye un enfoque metodológico cuantitativo de los resultados de los alumnos en función del empleo de diversas metodologías de enseñanza de la Literatura.

La evaluación de la eficacia del Sistema de Unidades Modulares de Aprendizaje (S.U.M.A.), se apoya en un modelo de evaluación de

producto, es decir, en la medición del rendimiento de los sujetos. Para hacer posible la comparación entre grupos se utilizan procedimientos estadísticos descriptivos e inferenciales.

Debido al limitado control de variables extrañas, la metodología adoptada en este estudio se ajusta al diseño cuasi-experimental de dos grupos. Se ha efectuado un muestreo doble: por un lado, de los productos literarios que se han seleccionado (por la técnica de catas distribuidas proporcionalmente entre algunas obras del autor literario elegido) y, por otro, de los escolares que intervinieron en la experimentación del programa. Los 90 sujetos de la muestra se clasifican en dos grupos:

- * Grupo experimental: sujetos sometidos a la tecnología del S.U.M.A. (45 alumnos).
- * Grupo de control: sujetos que siguen el sistema convencional de enseñanza (45 alumnos).

La evaluación se centra en contrastar la hipótesis que afirma que la aplicación de la tecnología del programa S.U.M.A. contribuye a generar resultados significativamente más positivos que la metodología convencional.

La información se recogió a través de pruebas diseñadas específicamente para valorar los conocimientos en el área de Literatura de los sujetos, al inicio y al final de la experimentación.

Con el fin de controlar las posibles diferencias de rendimiento debidas a las aptitudes de los sujetos que integran ambos grupos, se recogieron datos sobre las aptitudes, el razonamiento abstracto, la adaptación y creatividad de los sujetos. Los instrumentos usados a tal efecto fueron:

- El Test de Pronóstico Académico (A.P.T.) para medir las aptitudes (numérica, verbal y razonamiento abstracto).
- El Cuestionario de Adaptación Escolar (C.A.E.).
- El Inventario de Adaptación Social que permite el diagnóstico diferenciado de sujetos adolescentes y preadolescentes.
- Un Test de creatividad, a través del que se constata la fluidez de ideas e imágenes, la flexibilidad en las respuestas, la originalidad de las soluciones, etc.

El análisis estadístico de los datos confirma la hipótesis enunciada: *“La tecnología didáctica centrada en Unidades Modulares de Aprendizaje (S.U.M.A.), con módulos literarios que promueven aprendizaje independiente, autoevaluación y recuperación discente, obtiene resultados superiores al sistema de enseñanza convencional”*.

Los resultados de la evaluación vienen a subrayar que la tecnología didáctica centrada en Unidades Modulares de Aprendizaje (S.U.M.A.) de la Literatura, promueve un aprendizaje independiente, y facilita la autoevaluación y el progreso del alumno con dificultades.

Por otro lado, Alonso Tapia y su equipo, comienzan su investigación efectuando una revisión de los programas de entrenamiento cognitivo existentes para, posteriormente, hacer una valoración de los mismos y de sus perspectivas de éxito.

Una parte importante del trabajo consiste en la puesta en marcha de una serie de estudios experimentales encaminados a analizar los efectos que la aplicación de dichos programas ha suscitado.

El modelo de evaluación queda plenamente explicitado, y consiste en analizar los instrumentos más idóneos para potenciar y desarrollar las habilidades cognitivas y aumentar la motivación de los alumnos, como estrategia de educación compensatoria.

En la evaluación se adopta un diseño experimental, y se lleva a cabo la experimentación de los seis subprogramas seleccionados, para lo cual se ha seleccionado una muestra de 60 sujetos (asignándose aproximadamente 30 sujetos al grupo experimental y 30 sujetos al grupo de control para cada programa). Se han utilizado seis subtest y los materiales de aplicación de los programas. Tras la aplicación de los diversos programas de entrenamiento de habilidades cognitivas y enriquecimiento motivacional, se obtuvieron las puntuaciones de los sujetos en los instrumentos diseñados para tal fin, y se procedió a los análisis estadísticos pertinentes con el fin de estimar la efectividad de los programas, medidos en términos de mejora de la inteligencia, de las habilidades cognitivas y de la motivación.

Para la construcción, estandarización y validación de los instrumentos de evaluación referentes al programa de enriquecimiento motivacional se elaboraron tres cuestionarios que responden a la teoría sobre la motivación de logro, los estilos atributivos y las expectativas de control, y a las interacciones entre estas variables. Las muestras a las que se aplicaron estaban compuestas por 1202 alumnos de la 2.^a etapa de E.G.B.

Asimismo, se efectúa un estudio diferencial del entrenamiento atribucional y del entrenamiento en autoinstrucciones frente al fracaso en una situación de logro. En éste se constata que los resultados más destacables del entrenamiento atribucional y en autoinstrucción son que se aprenden tanto las respuestas instrumentales como las de atribución; que los efectos de la intervención están en función de las experiencias previas de fracaso; y que los efectos del entrenamiento se generalizan entre tareas.

Otra parte del diseño se ocupa del análisis o contraste entre el "entrenamiento en conductas instrumentales"/"el entrenamiento en atribuciones"/el grupo de control, midiéndose los resultados en función del nivel de dificultad en resolución de problemas de anagramas y *tangram*; que responden a diseños cuasi-experimentales de ejecución pre y post-tratamiento.

Los resultados obtenidos a través de los cuestionarios diseñados especialmente para cada uno de los programas aportan la información acerca de la mejora de las habilidades o capacidades de los sujetos (inteligencia, creatividad y motivación). Los resultados indican que el entrenamiento con los programas del Proyecto Inteligencia es efectivo, pudiéndose, en parte, predecir la ganancia de los sujetos a partir del nivel previo en el propio programa, pero no a partir del nivel previo en el resto de los programas.

Por su parte, el modelo adoptado por el equipo que realizó el estudio experimental basado en la teoría de Lakatos acerca del proceso de la Ciencia y de sus implicaciones didácticas (Ramos Fernández), es el de una evaluación de entrada-proceso-producto (la observación y seguimiento de la metodología se efectuó a lo largo de un curso completo), de tipo experimental con medidas pretest/postest.

El grupo experimental lo constituían alumnos sometidos al nuevo plan de intervención didáctica, y el grupo de control, quienes seguían la metodología didáctica convencional. Las muestras estaban constituidas por alumnos de 7.^º de E.G.B., 1.^º y 3.^º de B.U.P. y 1.^º y 3.^º de Medicina. Se pretendió hacer un seguimiento de cuáles eran las ideas previas de los estudiantes del último curso de E.G.B., como eran afianzadas en los de 1.^º y 3.^º de B.U.P., y cómo prevalecían en los distintos cursos universitarios, siempre dentro de las áreas científicas afines.

El tratamiento al que se sometió al grupo experimental consistió en unas unidades didácticas orientadas al afianzamiento y cuestiona-

miento de las ideas previas de los alumnos para después, y tras confrontar en un debate las distintas opiniones, buscar la contrastación experimental de las mismas.

La evaluación de una propuesta curricular de la Geometría plana basada en el Modelo de Razonamiento de Van Hiele, coordinada por Gutiérrez Rodríguez, se ajusta a un diseño cuasi-experimental, con muestra incidental, de un grupo al que se le aplicaron un pretest y un postest.

En un primer momento se presenta el marco teórico que sustenta la propuesta, según el cual se formulan cinco niveles de razonamiento geométrico, que van desde el puramente visual propio de los niños de los primeros cursos de la Educación Primaria, hasta el lógico-formal que desarrollaría un matemático.

El modelo de Van Hiele sugiere una enseñanza cíclica, en la que una parte de la Geometría se retoma para completar y mejorar su comprensión, utilizando unas formas de razonamiento más sofisticadas que cuando se inició su estudio.

Partiendo de estos presupuestos teóricos, se confeccionaron unas unidades de enseñanza de la Geometría secuenciadas en cinco niveles. El trabajo consistió pues en la evaluación de la eficacia de dichas unidades, en tanto procedimientos didácticos que facilitarían la consecución de dos tipos de objetivos: los relacionados con las estructuras de razonamiento y los relacionados con el aprendizaje de los conocimientos geométricos.

Para ello, mediante pretest-postest se identificaron las diferencias en el rendimiento de los sujetos, aplicando el correspondiente instrumento de evaluación que consistió en una prueba elaborada y validada al efecto, para medir los niveles de razonamiento matemático a partir del modelo de Van Hiele.

La muestra que sirvió para contrastar la efectividad del programa la constituían 128 alumnos de cinco Institutos de Formación Profesional (6 grupos de 1.º curso de F.P.) de la provincia de Valencia.

Un elemento importante a considerar en este estudio es la necesidad de evaluar el propio material diseñado, lo que el equipo de investigación se ha planteado como objeto de un próximo trabajo.

— *Difusión y uso de los resultados de las evaluación*

La generalización de los resultados del estudio de Alvarez González (1986) sobre la efectividad de la metodología S.U.M.A., está muy

determinada por la posibilidad de asumir el gasto que esto supondría. En el propio trabajo se deja constancia de esta dificultad, ya que en cada aula debería haber tantas Unidades Modulares como temas a tratar en las diversas áreas de aprendizaje del Ciclo Superior de E.G.B. Asimismo, se necesitaría personal especializado en Tecnología Modular (S.U.M.A.) para la elaboración adecuada de cada módulo de enseñanza-aprendizaje.

Sin embargo, se señala que el S.U.M.A. puede ser un recurso eficaz (tecnología didáctica) para los profesores de Literatura española, ya que, a partir de la Unidad Modular-Tipo elaborada, se podrían crear y desarrollar otras, con otros autores literarios significativos que, coordinadas, pudieran dar una visión general de la Literatura española actual.

Los resultados de la evaluación que ha efectuado el equipo del profesor Alonso Tapia (1986), consistentes en el diseño y evaluación de un conjunto de materiales, técnicas y procedimientos complementarios para potenciar el enriquecimiento cognitivo y motivacional de los alumnos con necesidades educativas especiales de la E.S.O., ponen de manifiesto la utilidad práctica que tiene el M.A.P.E. para los profesores de este tipo de alumnos. Los estudios de validación realizados parecen demostrar su eficacia como predictor del rendimiento escolar, que incrementa sensiblemente la validez de las predicciones que parece conseguirse con medidas de la inteligencia mediante el test de Factor "G" de Cattell.

Con respecto a la utilidad del cuestionario E.A.T. (de estilos atributivos) apoyado en las teorías de las "escuelas de causalidad multidimensionales y multiatributivas" de Lefcourt, Von Baeyer Ware y Cox (1979), junto con los trabajos del propio investigador que coordina el estudio (Alonso Tapia, 1985), la evaluación pone de relieve que su utilización conjunta con el test de inteligencia general de Catell permite mejorar notablemente la predicción de los resultados académicos en relación con lo que suponía la utilización exclusiva de este test.

Las conclusiones generales de la evaluación de Alonso Tapia (1986) indican que no parece aconsejable ni obligada la utilización de los programas en un paquete común, dado que las habilidades entrenadas son de niveles de dificultad muy distintos.

Asimismo, el programa en general se revela como efectivo aunque la duración del entrenamiento parece insuficiente para consolidar los aprendizajes complejos que el programa intenta lograr. De todo

ello se infiere que los autores de este programa de entrenamiento de habilidades cognitivas y enriquecimiento motivacional orientadas a la educación compensatoria han estimado conveniente someter a revisión algunos aspectos del mismo con el fin de hacer posible un mayor incremento en las mejoras propuestas.

No puede hacerse una estimación de la utilidad y difusión de los resultados del estudio dirigido por Ramos Fernández, ya que se haya todavía en curso.

Por otro lado, de las conclusiones generales de la evaluación de la propuesta curricular para el aprendizaje de la Geometría Plana basada en el Modelo de Razonamiento de Van Hiele (Gutiérrez Rodríguez), parece inferirse que las unidades de enseñanza diseñadas y experimentadas son un buen material de partida, una buena ejemplificación para que el profesorado de Enseñanza Secundaria que comparta el modelo de razonamiento de Van Hiele diseñe sus propias unidades didácticas en esta área curricular.

— *Metaevaluación*

Las evaluaciones integradas en este grupo no parecen contemplar una metaevaluación, aunque la de Alonso Tapia y colaboradores parte de una evaluación de las evaluaciones contenidas en los programas de entrenamiento incluidos en el Proyecto Inteligencia de la Universidad de Harvard para su posterior adaptación al contexto de educación compensatoria de nuestro país. Finalmente apunta a una metaevaluación que propone la modificación de la aplicación de los distintos programas de entrenamiento de habilidades y enriquecimiento motivacional a la vista de los resultados obtenidos.

b. *Evaluación de programas desarrollados en Educación Superior*

— *Objeto de la evaluación*

Las tres evaluaciones incluidas en este grupo se refieren a programas aplicados en niveles universitarios, por lo que merecen un tratamiento aparte. Se trata de:

- *“Diseño, puesta en práctica y evaluación de un método didáctico con enfoque constructivista para la enseñanza de la*

Física a nivel universitario", dirigida por María Concesa Caballero Sahelices (actualmente en curso).

- "*Evaluación de métodos de simulación informática en Enseñanzas Técnicas Superiores*", dirigida por J. Manuel Sánchez Alciturri (1987).
- "*Evaluación de programas didácticos y de su proceso de realización*", dirigida por Antonio Bartolomé Pina (1987).

La investigación de Caballero Sahelices evalúa un programa específico de intervención educativa que pretende establecer estrategias de instrucción útiles para programar unidades didácticas, cuyo objetivo es el aprendizaje significativo de la Física. La muestra con la que se experimenta el programa son los alumnos de 1.º curso de Ciencias Químicas que cursan la asignatura Física General.

Por su parte, en la investigación dirigida por Sánchez Alciturri, tras comprobarse la necesidad de dedicar más tiempo al análisis de los ejercicios prácticos de la asignatura Geotécnia I impartida en 4.º curso de Ingeniería de Caminos, se elaboraron y evaluaron cuatro programas informáticos específicos con objeto de ahorrar determinados cálculos a los alumnos, y guiar su trabajo personal hacia actividades más profundas de comprensión, comparación y análisis de los resultados obtenidos.

En cuanto a la investigación dirigida por Bartolomé Pina, consiste en la elaboración y evaluación de una serie de programas videográficos para la enseñanza-aprendizaje de algunos contenidos de Estadística Inferencial aplicada a las Ciencias Sociales. Dichos programas se caracterizan por utilizarse en grupo, mantener un planteamiento activo del aprendizaje, incidir en aspectos educativos más que en los meramente instructivos, y aprovechar todos los recursos del medio audiovisual, especialmente sus aspectos formales. La experiencia se realizó con un grupo natural de alumnos de 2.º curso de Ciencias de la Educación que cursaban la asignatura Estadística II.

— *Características definitorias de la evaluación*

Agentes evaluadores

Las tres investigaciones suponen una evaluación interna de los programas, realizada por los respectivos profesores universitarios que

imparten docencia en la asignatura a que hace referencia cada estudio: Física, Geotécnia I y Estadística II, respectivamente.

Objetivos y funciones de la evaluación

El objetivo principal de la investigación “*Diseño, puesta en práctica y evaluación de un método didáctico con enfoque constructivista para la enseñanza de la Física a nivel universitario*”, se expresa claramente en su título. Más concretamente, los objetivos se enuncian como sigue:

- Diseñar un modelo didáctico para la enseñanza de la Física en la Universidad, que promueva el aprendizaje significativo.
- Comparar el grado de aprendizaje y el cambio conceptual producido a corto y largo plazo, en los alumnos que seguían la metodología propuesta y en aquellos con los que se mantenía la enseñanza tradicional expositiva.
- Averiguar si existen interacciones entre el método de enseñanza y las variables sexo, nivel sociocultural, estilo cognitivo, nivel de razonamiento y actitud hacia la metodología.

Este estudio se enmarcaba en una línea de investigación orientada a la elaboración de un cuerpo coherente de conocimientos sobre didáctica de las Ciencias, que tiene por objeto encontrar alternativas, debidamente contrastadas, al modelo predominante de enseñanza universitaria basado en la transmisión/asimilación de conocimientos ya elaborados. Así, en la medida que trata de mejorar el proceso de enseñanza-aprendizaje, la evaluación del programa se puede considerar formativa, aunque, como en todo estudio que trata de determinar la validez o no de un programa para decidir su adopción, cumpliría también una función sumativa.

Por su parte, el objetivo principal del estudio consistente en la “*Evaluación de métodos de simulación informática en Enseñanzas Técnicas Superiores*”, fue introducir técnicas de simulación informática en la enseñanza de la Geotécnia y comprobar si su utilización permitía disminuir el tiempo preciso para trabajar cada tema, en comparación con el que era necesario utilizando el método de enseñanza habitual. En definitiva, el estudio se ocupó en contrastar la siguiente hipótesis: “El alumno obtiene beneficios mayores o iguales en tiempos menores con la utilización de las técnicas docentes propuestas (...)”.

Se trata en este caso de una evaluación con función sumativa, ya que se centra básicamente en los resultados del programa y pretende

informar la toma de decisiones sobre el cambio de una metodología de enseñanza tradicional a otra más innovadora.

La investigación consistente en la “*Evaluación de programas didácticos y de su proceso de realización*”, llevó a cabo una evaluación formativa de los programas, que pretendía desarrollar técnicas fiables para mejorar la eficacia de los programas educativos durante su proceso de aplicación. Asimismo, trataba de averiguar qué programa de vídeo es más adecuado para cada tipo de tarea o para cada tipo de planteamiento didáctico. Por tanto, al iniciar esta investigación se esperaba poder disponer de elementos de referencia fiables para orientar la realización de programas didácticos en vídeo. Sus objetivos se pueden concretar como sigue:

- Diseñar unos sistemas de elaboración de programas educativos en vídeo.
- Evaluar la eficacia y rentabilidad de dichos programas.
- Estudiar los contextos educativos apropiados para los diferentes usos de estos programas didácticos.

Dimensiones evaluadas y momento de la evaluación

En el estudio dirigido por Caballero Sahelices, según el diseño pretest-postest, se realizan medidas antes y después de aplicarse el programa de enseñanza de la Física. En concreto, las medidas se realizan sobre las siguientes variables:

- Aprendizaje de conceptos
- Aprendizaje de procedimientos rutinarios.
- Habilidad en la resolución de problemas.
- Competencia en el uso de los métodos de la Ciencia.
- Rendimiento global.
- Actitud hacia la metodología.
- Conceptualización de conceptos básicos de Física.
- Usos metodológicos.

Del mismo modo, en el estudio dirigido por Sánchez Alciturri se contemplaron variables de entrada y de producto (medidas antes y después del desarrollo de las clases teóricas y prácticas en las que se emplearon técnicas de simulación informática en la materia de Geo-

técnica), que se refieren a los siguientes aspectos del rendimiento de los alumnos de 4.º curso de Ingeniería de Caminos: conocimientos, actitudes y capacidad de abordar problemas.

Por su parte, en la investigación dirigida por Bartolomé Pina, se tomaron medidas de la variable dependiente (información que retiene el alumno inmediatamente después de ver el vídeo) solamente tras la aplicación de cada uno de los programas que se experimentaban. Así, en este caso parece que nos encontramos con una evaluación centrada en los productos, en la que se controlan algunas variables extrañas (inteligencia, dependencia/ independencia de campo, actitud previa ante la asignatura, clase social, lengua materna (castellano/catalán), estudios previos cursados en la rama de Letras o de Ciencias, sexo y sistema de acceso a la Universidad) con el fin de poder apreciar mejor las variaciones en el recuerdo de la información debidas sólo a la variable método.

Patrón empleado para la emisión de juicios de valor

Dos de los estudios de este bloque (los de Caballero Sahélices y Bartolomé Pina) responden a una evaluación de carácter normativo ya que básicamente ambos consisten en un diseño comparativo entre grupos experimental/control: se lleva a cabo una asignación de individuos a ambos grupos, se someten a dos tratamientos diferenciados y se mide la variable objeto de estudio. Por el contrario, la investigación dirigida por Sánchez Alcuturri es una evaluación criterial en la que, a partir de medidas pretest/postest, se analiza el rendimiento de los alumnos en cada una de las metodologías estudiadas.

— Diseño y desarrollo de la evaluación

Los tres estudios que conforman este grupo de evaluaciones de programas en Educación Superior, adoptan un diseño experimental con dos grupos y medidas pretest/postest.

El diseño de la investigación en curso que dirige Caballero Sahélices es, como se ha dicho, experimental y de dos grupos, trabajando con una muestra incidental. Los grupos están constituidos como sigue:

- *Grupo experimental*: la mitad de los alumnos matriculados en la asignatura de Física general de 1.º curso de Ciencias Químicas (60 alumnos aproximadamente).
- *Grupo control*: el resto de los alumnos del curso.

En el diseño no se prevé hacer sólo dos medidas para recoger o controlar la información definida en las variables, sino tres (antes de emplear la metodología constructivista en la enseñanza de la Física, inmediatamente después, y a los tres meses de su aplicación), ya que la persistencia de la formación conceptual y metodológica adquirida por el alumno, es una de las variables dependientes más importantes del diseño. El análisis de los datos se plantea con técnicas estadísticas de carácter inferencial.

Asimismo, la investigación dirigida por Sánchez Alciturri adopta un diseño experimental, de dos grupos y con medidas antes y después de las clases teóricas y de las prácticas de laboratorio en las que se ejecutan los programas informáticos. Sin embargo, las condiciones técnicas del diseño no aparecen claramente explicitadas, ofreciéndose sólo referencias generales. La muestra, por ejemplo, se define como formada por *“todos los alumnos de la asignatura en el curso 1985-1986...que se dividen en dos grupos aleatorios que reciben enseñanza con cada uno de los métodos (clásico o con simulación informática)”*. Se realizó un estudio estadístico descriptivo de los datos resultantes.

Por su parte, la investigación dirigida por Bartolomé Pina dispone de un diseño mixto, en el que hay un tratamiento estadístico de los datos, pero también se recogen una serie de elementos del proceso de experimentación de los programas que permiten un análisis cualitativo del mismo. En este estudio se contrastaba la hipótesis de que al aumentar la información verbal en los programas educativos en vídeo, ésta no es correctamente procesada. Así, se prepararon dos versiones de cada uno de los programas experimentados, en los que la única diferencia era la existencia de una mayor información verbal en una de las versiones. Así, se configuraron dos tratamientos que conformaban la variable independiente: 1) Programas no saturados de información verbal; y 2) programas saturados de información verbal. Siendo variables dependientes: la información retenida después del visionado, la actitud final ante los videos, y el rendimiento en Estadística. Asimismo, se controlan una serie de variables que pueden intervenir en el proceso como son: inteligencia, dependencia-independencia de campo, actitud previa ante la asignatura, clase social, estudios de letras o ciencias, sistema de acceso a la Universidad y el ser o no repetidor.

Los instrumentos de medida utilizados son: Prueba de Actitudes Mentales Primarias, Test de Figuras Enmascaradas, Cuestionarios sobre actitudes y características personales del sujeto, Test de informa-

ción Retenida, Test de rendimiento, habiendo sido elaborados específicamente los tres últimos.

— *Difusión y uso de los resultados de la evaluación*

La investigación sobre Evaluación de Métodos de Simulación Informática en Enseñanzas Técnicas Superiores dirigida por Sánchez Alciturri es difícilmente aprovechable: como señalan sus autores “la serie de programas desarrollados no son adecuados a la finalidad perseguida, y el sistema utilizado y el diseño de actividades han sido erróneos”. Ante esta evidencia no cabe hacer comentario alguno, si bien es cierto que el estudio constituye, en el contexto del centro donde se ha desarrollado, un trabajo novedoso en el que las actividades realizadas pueden resultar útiles para continuar en labores semejantes.

Con respecto al estudio de Bartolomé Pina sobre la Evaluación de Programas Didácticos y sus procesos de realización, hay que destacar que se comprueba que el incremento de explicación verbal en un vídeo didáctico no sólo no ayuda, sino que puede dificultar la captación de información. Por otra parte, el carácter del ítem (verbal o icónico-verbal) no afecta al resultado de las pruebas utilizadas para medir el conocimiento.

La principal aportación de esta investigación se centra en torno a la capacidad informativa del audiovisual, desde un punto de vista integrador; además se sugieren alternativas de posibles diseños sobre el papel concreto de la banda sonora y el mensaje verbal en este medio de comunicación, trabajando desde otras áreas y otros niveles de conocimiento.

En el caso del estudio de Caballero Sahelices, al estar todavía en curso, no puede decirse aún nada sobre el uso y difusión de sus resultados.

— *Metaevaluación*

Ninguna de las investigaciones incluidas en este grupo hacen referencia a la metaevaluación.

4.2.3. Evaluación de Programas de Formación Inicial y Permanente del Profesorado

4.2.3.1. Programas de Formación Inicial

Pertenecen a este grupo de investigaciones centradas en la evaluación de programas de formación inicial del profesorado, las siguientes:

- “Evaluación de un programa experimental de entrenamiento cognitivo social para profesores en formación” de Alvarez Castillo, José Luis (1995).
 - “Propuesta, experimentación y evaluación de un programa de formación del profesorado en la aplicación e investigación de la enseñanza constructivista” de Barrón Ruíz, Angela y Del Río Sánchez, José (en curso).
 - “Diseño, desarrollo y evaluación de un programa de inducción a la práctica profesional para profesores principiantes basado en la reflexión y supervisión” (Marcelo García, Carlos. 1990).
- *Objeto de la evaluación*

La investigación dirigida por Marcelo García consiste en un estudio de necesidades en el que el objeto de evaluación es la práctica diaria de profesores principiantes y las necesidades derivadas de ella, con la intención de, a partir de la información recabada, elaborar el mencionado programa de inducción.

Las otras dos investigaciones de este grupo, dirigidas por Barrón Ruiz y Alvarez Castillo, tienen en común su carácter experimental, siendo dichos programas muy específicos y de alcance restringido; la primera se centra en un programa de enseñanza de tipo constructivista y la segunda en el entrenamiento cognitivo-social de profesores en formación.

- *Características definitorias de la evaluación*

Agentes evaluadores

Los tres estudios que conforman este apartado, que tienen como finalidad la evaluación de un determinado programa de formación del profesorado, representan una evaluación interna porque en todas ellas los responsables del diseño de los programas son las mismas personas que los llevan a la práctica y los evalúan.

Objetivos y funciones de la evaluación

El objetivo central de la investigación de Alvarez Castillo es experimentar un programa que se enmarca en un contexto de mejora de la

calidad de la enseñanza y de adopción de estrategias innovadoras en los planes de formación del profesorado.

Dada la influencia que en el rendimiento de los alumnos tienen las expectativas que sobre ellos mantienen los profesores, se resalta la conveniencia de incluir en los Planes de Formación del Profesorado algún tipo de programa de entrenamiento cognitivo-social que les ayude a conocer sus mecanismos perceptivos y a corregir los sesgos presentes en el procesamiento de la información social.

El programa en cuestión fué diseñado siguiendo las fases siguientes: diseño de objetivos; selección de contenidos referentes a los sesgos del conocimiento social que se pretenden erradicar; metodología didáctica y actividades de aula; criterios de evaluación; secuenciación de los contenidos y temporalización; recursos materiales y personales.

Los investigadores presentan en el siguiente esquema los objetivos perseguidos en su estudio:

Objetivo general: comprobar la eficacia de un programa de entrenamiento cognitivo-social.

Objetivos intermedios:

- a) Conocimiento de mecanismos perceptivos y corrección de sesgos cognitivos:
 - Reducción de sesgos atribucionales de carácter motivacional.
 - Reducción de sesgos atribucionales de carácter cognitivo.
 - Incremento del control de los procesos cognitivo-sociales.
 - Incremento del control de las estructuras cognitivo-sociales.
 - Identificación de variables moduladoras de la percepción social.

- b) Concesión de prominencia a las características positivas de los alumnos/as.
 - Creación de expectativas positivas, partiendo de la modificación de las fases de atención y selección de estímulos.

Se lleva a cabo una evaluación con función sumativa que se propone comprobar si se consigue alcanzar el principal objetivo del programa tras su realización (mejorar las habilidades cognitivo-sociales de los profesores).

El trabajo de Barrón Ruíz para evaluar un programa de formación del profesorado en el ámbito de la enseñanza constructivista, se divide

en dos partes que responden a dos evaluaciones distintas. En la primera fase se pretende el desarrollo en el profesorado de estrategias de enseñanza acordes con el modelo pedagógico constructivista, resaltando sus ventajas frente al modelo tradicional de enseñanza expositiva no significativa. Esta evaluación, aunque trata de evaluar los resultados obtenidos en la aplicación del programa, tiene una función más formativa que sumativa al atender a todo el desarrollo del proceso educativo con el fin de tomar conciencia de los avances realizados, así como de las dificultades encontradas y de los errores cometidos para proceder a su corrección.

En la segunda fase se compara la eficacia de tres metodologías de enseñanza: la expositiva tradicional, y otras dos basadas en planteamientos constructivistas: la metodología expositiva-significativa y la de aprendizaje por descubrimiento, por lo que la evaluación reviste cierto carácter normativo.

El estudio de Marcelo García sobre un programa de inducción a la práctica profesional para profesores principiantes, persigue los siguientes objetivos:

- Identificar problemas de enseñanza referidos a: planificación, evaluación, relaciones con alumnos, padres y otros profesores, dirección y administración, recursos escolares, percepción del entorno escolar, satisfacción profesional y tiempo de dedicación docente.
- Analizar el proceso de socialización de los profesores:
 - Dimensión personal: experiencias pasadas/presentes, creencias personales, cómo se perciben a sí mismos.
 - Dimensión didáctica: el profesor y su percepción del aula y de la tarea docente.
 - Dimensión institucional: centro educativo, ambiente en los centros, infraestructura, curriculum escolar, relaciones con los padres.
- Describir el ambiente de las clases mediante el estudio de:
 - La personalización de la enseñanza: interés y preocupación por interactuar con los alumnos, atención y comprensión a sus demandas y opiniones.
 - La implicación: interés del alumno por las materias impartidas por el profesor.

- La cohesión entre los estudiantes.
- La satisfacción de los estudiantes ante el modo en que el profesor conduce la clase.
- La orientación en la tarea: grado en que los alumnos perciben la clase como correctamente organizada.
- La innovación de las tareas.
- El grado de respeto a las habilidades y destrezas individuales.

Se trata de una evaluación orientada al diagnóstico, la planificación y la toma de decisiones generales acerca de la naturaleza y características que debía reunir el programa que se proponían desarrollar.

Dimensiones evaluadas y momento de la evaluación

La investigación de Barrón Ruiz se centra, en una primera fase, en la evaluación del estilo pedagógico del profesorado antes y después de asistir a un curso teórico-práctico sobre la enseñanza constructivista de las Matemáticas diseñado previamente. Por lo tanto, se llevó a cabo una evaluación tanto inicial como final, que incluía las siguientes variables: exploración del pensamiento de los alumnos; exposiciones; discusiones profesor-alumno y alumno-alumno; trabajo práctico manipulativo; destrezas y rutinas básicas; problemas de la vida cotidiana; y problemas del libro de texto.

En la segunda fase, en la que se compara la eficacia de tres metodologías de enseñanza, las variables consideradas dependientes fueron: el rendimiento a corto y largo plazo en el manejo de estructuras conceptuales, los procedimientos algorítmicos y de resolución de problemas, el rendimiento global y la actitud hacia las Matemáticas. Con respecto al momento de la evaluación, se prevé llevarla a cabo antes de comenzar la experimentación, inmediatamente después de acabarla, y a los dos meses de la misma.

Marcelo García desarrolla una evaluación inicial que trata de describir el primer año de docencia desde la perspectiva de los propios sujetos implicados, es decir, desde las características, necesidades y carencias reales de los destinatarios del programa (profesores de E.G.B., de B.U.P., de F.P. y de Educación Especial).

En el estudio de Alvarez Castillo también hay evaluación inicial y final, pero en este caso las variables dependientes fueron medidas en

dos ocasiones (pretest y postest). Dichas variables son definidas por los autores como: favorabilidad de las heteropercepciones (integrada por seis subvariables: “popular/impopular, trabajador/vago, seguro/inseguro, sociable/insociable, inteligente/torpe, y equilibrado/descontrolado”); favorabilidad de las autopercepciones (las seis subvariables que la integran se refieren a las mismas escalas que en el caso anterior); favorabilidad de las expectativas (integrada por tres subvariables: éxito/fracaso económico, éxito/fracaso afectivo, favorecido/desfavorecido en cuanto a atractivo físico); tipo de atribución (externa/interna); ponderación estímulo-escala; y frecuencia de constructos interpersonales positivos.

Patrones empleados para la emisión de juicios de valor

La primera parte de la investigación dirigida por Barrón Ruiz, con diseño pretest/postest para detectar la virtud de un programa de formación de profesores, se puede considerar una evaluación criterial que utiliza como pautas de comparación las variables anteriormente mencionadas. En la segunda parte, que persigue comparar la eficacia de tres metodologías de enseñanza, el criterio de comparación es un grupo normativo, el grupo de control, al que se le aplica un método tradicional de enseñanza de las Matemáticas frente a dos grupos experimentales a los que se le aplican métodos constructivistas, si bien también se realizan pruebas antes y después de la experimentación.

Semejantes características a las de la segunda parte del estudio de Barrón Ruiz presenta el de Alvarez Castillo, que asimismo se puede considerar una evaluación normativa, en este caso con tres grupos de control y uno experimental y con dos medidas pretest y dos postest.

Por último, la investigación dirigida por Marcelo García es una evaluación de carácter exploratorio o diagnóstico, que no utiliza grupos de comparación y que lleva a cabo un estudio previo de las necesidades y carencias percibidas por profesores principiantes en su práctica diaria a partir de sus propias opiniones para, a partir de su conocimiento, diseñar un programa de inducción a la práctica profesional.

— Diseño y desarrollo de la evaluación

Aunque no resulta fácil alinear los modelos de evaluación que proponen los investigadores a alguno de los modelos clásicos definidos en la literatura especializada, todos parten de un planteamiento que se concreta en un determinado diseño de evaluación.

En este sentido, se distinguen dos enfoques claros: el naturalista-cualitativo y el experimental-cuantitativo. Uno de los estudios, el dirigido por Marcelo García, persigue la complementariedad en el enfoque y metodología adoptados. Sólo dos investigaciones incorporan un diseño experimental: la propuesta, experimentación y evaluación de un programa de formación del profesorado en la aplicación e investigación de la enseñanza constructivista, de Barrón Ruiz y la evaluación de un programa experimental de entrenamiento cognitivo-social para profesores en formación, de Alvarez Castillo.

En la investigación "Diseño, desarrollo y evaluación de un programa de inducción a la práctica profesional para profesores principiantes basado en la reflexión y la supervisión" (Marcelo García), se lleva a cabo un estudio descriptivo mixto con técnicas de análisis tanto cualitativas como cuantitativas. El diseño de la investigación es del tipo "*multisites*", en el que, dependiendo de las necesidades del proceso, los aspectos de reducción de datos, y la estructuración y presentación de conclusiones, se interrelacionan e influyen unos con otros.

Se puede dividir en tres estudios independientes con un diseño propio cada uno. El primero, que coincide con el primer objetivo de la investigación, se centra en identificar los problemas de los profesores principiantes y emplea fuentes personales de información. Concretamente se trata de profesores principiantes clasificados por sexo, especialidad académica ejercida, situación geográfica y nivel de enseñanza (E.G.B., B.U.P., F.P. y Educación Especial). Los procedimientos que se utilizan para recoger información incluyen un *Inventario de Problemas de Enseñanza* (Jordell, 1985), traducido y adaptado por el autor de la investigación, y entrevistas (individuales y estructuradas) con preguntas que corresponden a las siguientes categorías:

- *Experiencias*: los sujetos deben describir sucesos, preocupaciones y opiniones.
- *Preguntas "mini-tours"*: descripción narrativa de algún acontecimiento cotidiano.
- *Preguntas de "lenguaje nativo"*: los profesores responden con sus propias palabras expresando un lenguaje directo.

El segundo estudio (2.^a objetivo) se centra en las creencias de los profesores desde un punto de vista personal (experiencias pasadas/presentes, creencias personales, cómo se perciben a sí mismos),

didáctico (percepción del aula y de la tarea docente) e institucional (centro educativo, infraestructura, curriculum escolar y relaciones con los padres). Para recoger la información sobre estos aspectos se utiliza el *Inventario de Creencias del Profesor* (de Tabachnick y Zeichner) adaptado y validado para este estudio, que consta de seis dimensiones:

- A. *Locus de control* (describir y analizar las diferencias en el "locus" y extensión del control clase). Los ítems se agrupan en cuatro subdimensiones: procedimientos en clase; curriculum; implicación de los padres y profesores en las decisiones de la escuela; conducta del alumno.
- B. Diversidad: cómo manejar las diferencias entre los niños.
- C. Aprendizaje: aprendizaje cooperativo, errores en el aprendizaje y tipo de motivación.
- D. Rol del profesor.
- E. Relación escuela/sociedad: participación del profesor y de la escuela en la transformación de la sociedad y papel de la escuela frente a las desigualdades sociales.
- F. Conocimiento: énfasis del curriculum, énfasis en la separación/integración de las materias.

En el tercer estudio, cuyo objeto es conocer cómo perciben los alumnos y profesores el ambiente de clase, las fuentes de información son, obviamente, los profesores y los alumnos. El instrumento para recabar dicha información es el *"Inventario de Ambiente de Clase en Facultades y Escuelas Universitarias"*, traducido y adaptado por Marcelo García.

El análisis cualitativo de los datos procedentes de las entrevistas tiene diversas fases de desarrollo:

- * Elaboración de un primer sistema de códigos para reducir e interpretar los resultados. Se establecen una serie de categorías para cada una de las siguientes dimensiones: personal, de clase e institucional.
- * Comprobación de la fiabilidad entre codificadores (categorías) y análisis de la pertinencia del sistema de categorías creado.
- * Reelaboración del primer sistema de categorías.

Se consideran problemas importantes aquellos que han sido identificados por, al menos, el 25% de los profesores principiantes. Para facilitar la organización de los resultados, los problemas se agrupan en dimensiones (de enseñanza, planificación, evaluación, medios y recursos, tiempo para llevar a cabo los objetivos educativos marcados, relaciones con alumnos, padres y otros profesores, dirección y administración, dimensión personal y dimensión entorno).

Al analizar los datos se tienen en cuenta los criterios utilizados en la selección de la muestra: sexo, nivel o modalidad educativa (E.G.B., B.U.P., F.P. y Educación Especial), especialidad (Ciencias, Lengua, Ciencias Humanas, Expresión), situación y ciudad (Granada y Sevilla). También se lleva a cabo un análisis cuantitativo del *Inventario de Problemas de Enseñanza*, que proporciona 5 factores (contenidos de enseñanza, problemas en las relaciones con los padres, problemas referidos al tiempo, problemas referidos a la enseñanza y problemas referidos a la escuela y los compañeros).

Con respecto a los datos procedentes del *Inventario de Creencias del Profesor*, se realiza un análisis descriptivo de frecuencias en cada una de las dimensiones que se recogen en el mismo.

Del *Inventario de Ambiente de Clase* se analiza cada dimensión teniendo en cuenta los diferentes criterios que se han utilizado para seleccionar la muestra de sujetos.

Por su parte, el estudio dirigido por Barrón Ruiz, consta, como ya se ha mencionado anteriormente, de dos partes diferenciadas que suponen dos evaluaciones distintas. La primera adopta un diseño experimental del tipo pretest/postest. En este caso la selección de los profesores no es aleatoria, y se realiza en función de los criterios de pertenecer a un mismo centro y la disposición personal hacia metodologías didácticas innovadoras.

Se emplean fuentes directas de información (profesores y alumnos) de las que, en función de los instrumentos que se utilizan, se obtiene información de tipo cualitativo y cuantitativo:

- Observación de los comportamientos desarrollados en clase, que se apoya con grabaciones ocasionales en video;
- Cuestionarios refrendados a través de dinámicas de grupo.

Por otra parte, en la segunda fase de la investigación, donde se compara la eficacia de tres metodologías de enseñanza para el aprendizaje de contenidos estructurantes de difícil adquisición, se utiliza

también un diseño experimental con medidas pretest/postest para identificar relaciones causales entre variables. Se definen tres grupos de comparación, dos experimentales y uno de control:

- * Experimental I: compuesto por 100 alumnos a los que se aplica un método de enseñanza por descubrimiento orientado.
- * Experimental II: formado por 100 alumnos que reciben una enseñanza expositiva significativa.
- * Control: 100 alumnos que siguen la metodología expositiva no significativa o tradicional.

Por su posición en el diseño, las variables consideradas (independientes y dependientes) fueron las siguientes:

- * Variables independientes: los tres métodos de enseñanza (por descubrimiento orientado, expositivo-significativo y expositivo no significativo).
- * Variables dependientes (rendimiento): aprendizaje de estructuras conceptuales a corto y largo plazo, cambio conceptual a corto y largo plazo, aprendizaje de procedimientos algorítmicos, aprendizaje de resolución de problemas, rendimiento global a corto y largo plazo, y actitud hacia las Matemáticas.

En este caso los procedimientos para obtener la información son exclusivamente cuantitativos (aplicación de pruebas estandarizadas):

- Prueba sobre conocimientos y habilidades.
- Prueba de estructuras conceptuales.
- Prueba de procedimientos algorítmicos.
- Prueba de resolución de problemas.
- Test de actitud hacia las matemáticas (TAM).

En la misma línea, por su diseño y enfoque metodológico, se encuentra el trabajo de Álvarez Castillo para evaluar la eficacia de un programa de entrenamiento cognitivo-social. En este caso se optó por tomar cuatro grupos de comparación (uno experimental, uno de control interno y dos de control externo); se emplea un diseño que incluye dos pretests, con el fin de comprobar la estabilidad de la percepción

social previa (medida de fiabilidad) y dos postests para verificar la efectividad en el aprendizaje de las habilidades previamente adquiridas; y se presentan 16 hipótesis operativas de partida. El proceso de obtención de datos antes y después de la intervención se basa en el método de observación. Los sujetos con los que se ha trabajado en este estudio no fueron seleccionados aleatoriamente, el procedimiento de muestreo no fue probabilístico, sino que los cuatro grupos ya estaban formados de manera natural y eran alumnos de centros universitarios que cursaban 1.º de Magisterio de la especialidad de Educación Infantil y Primaria.

En cuanto a los procedimientos de recogida de información antes y después de la aplicación del programa se emplearon técnicas cualitativas y cuantitativas. Se planificó un procedimiento general de obtención de datos basado en la observación de secuencias de interacción infantil en escenarios escolares, en cuya percepción podían operar sesgos cognitivos y motivacionales, y en la grabación en video de secuencias conductuales en los diferentes escenarios escolares (aulas, talleres, patio y comedor). Asimismo, se ha utilizado un cuestionario que consta de seis tipos de preguntas con diferentes formatos de respuesta (escalas, opción múltiple, no estructurada) y que pretendían analizar el estatus de los cuatro grupos estudiados en las habilidades cognitivo-sociales que se querían mejorar con el programa de entrenamiento.

Posteriormente, y para contrastar las hipótesis de partida, se llevó a cabo un análisis estadístico de los datos obtenidos. Finalmente, se realiza una descripción de los resultados obtenidos y se resumen los resultados globales de la investigación en relación a la consecución de los objetivos.

— *Difusión y uso de los resultados de la evaluación*

Los resultados del estudio de Alvarez Castillo se pueden sintetizar en torno a los dos tipos de hipótesis formuladas: las que tratan de verificar la existencia de un sesgo inicial y las que predicen la eficacia del programa de entrenamiento cognitivo-social. Según esta clasificación se pueden establecer los siguientes resultados globales:

1. La percepción social de los sujetos está caracterizada por diversos tipos de sesgos y por la ausencia total de control sobre sus estructuras y procesos cognitivos.

2. El programa de entrenamiento al que es sometido el grupo experimental no logra modificar sustancialmente el funcionamiento cognitivo-social de los sujetos.

En definitiva, las hipótesis centrales para la investigación, aquellas que anticipan la eficacia del programa cognitivo, no se verifican, si bien se obtienen confirmaciones parciales de las hipótesis alternativas en tres de las nueve expectativas formuladas. La relevancia de este estudio en el campo del perfeccionamiento del profesorado se basa en que llama la atención sobre la importancia de la percepción que los profesores tienen de sus alumnos para un eficaz proceso de enseñanza y aprendizaje y, por consiguiente, de los programas de entrenamiento docente.

Por su parte, los resultados del estudio dirigido por Marcelo García ponen de manifiesto la necesidad sentida por un elevado número de profesores principiantes de que se les proporcione asesoramiento pedagógico. Para el profesorado esto representa, según los investigadores, disponer de información relevante para afrontar con eficacia su primer año de docencia, en este caso por medio de los Módulos Reflexivos, cuyo contenido se corresponde con los problemas y necesidades formativas identificadas por los mismos profesores.

Los resultados se dirigen a profesores que se encuentren en alguna de las siguientes situaciones:

- Profesores en sus primeros años de enseñanza, inmediatamente después de su graduación (su experiencia se limita a prácticas realizadas durante la carrera).
 - Profesores con alguna experiencia docente, que habiendo abandonado la docencia durante un considerable período de tiempo, deciden retornar a la enseñanza.
 - Profesores que asumen diferentes roles dentro del sistema educativo: que cambian de escuela, de nivel de enseñanza o de materia a enseñar.
- *Metaevaluación*

En ninguna de las investigaciones de este grupo se lleva a cabo una metaevaluación.

4.2.3.2. Programas de Formación Permanente

— Objeto de la evaluación

Las investigaciones de este grupo emprenden la evaluación de dos tipos de programas:

- * *Institucionales*: programas que generalmente tienen un ámbito de aplicación amplio, y que han sido propuestos desde la Administración educativa central o autonómica.
- * *Experimentales*: programas de formación propuestos por el equipo investigador y de aplicación restringida, aunque en ocasiones se aspire a su generalización.

Las investigaciones que evalúan programas institucionales tienen en común el hecho de contribuir al desarrollo del Plan Marco de Formación del Profesorado que, dentro del contexto general de la reforma del sistema educativo, diseña el Ministerio de Educación y Ciencia y que supone el punto de referencia para la planificación anual de todas las actividades a desarrollar en el ámbito de la formación permanente del profesorado. Generalmente son programas de formación en centros que promueven el cambio y la innovación educativa en la medida en que inciden directamente en todos los procesos que se generan en las aulas y en los propios centros docentes.

El Plan Anual de Formación Permanente del Profesorado para el curso 1994 define esta modalidad de formación como *una actuación formativa que tiene como eje el centro docente. Está basada en las necesidades de formación surgidas en un Claustro o equipo de profesores que imparten la docencia en un centro escolar y cuyo objetivo primordial es mejorar la calidad de la enseñanza en el mismo, tanto en lo referente a la actuación en el aula como en lo relativo a la gestión y coordinación pedagógica del centro.*

Desde esta perspectiva, las cinco investigaciones de este grupo -dirigidas por San Fabián Maroto (1994), Porlán Ariza (en curso), Gairín Sallán (1994), Nicolás Prieto (en curso) y Marcelo García (1994)-, están referidas a la formación permanente de profesores, y evalúan programas de carácter institucional siendo, por tanto, de amplio alcance. A continuación se comentan sus respectivos objetos de estudio.

San Fabián Maroto en su investigación "*Condiciones y posibilidades de la formación del profesorado en centros*" (1994), se propone describir la organización de los centros que desarrollan iniciativas de formación del profesorado en Asturias. Por ello, se revisan la mayoría de los Proyectos de Formación realizados en esta región desde las primeras convocatorias, y se lleva a cabo un análisis exhaustivo y sistemático de los seleccionados en la convocatoria de 1992, cuyo período de realización correspondía al curso 1992-93.

El trabajo de Porlán Ariza titulado "*Evaluación formativa de un proyecto institucional de formación permanente del profesorado: El Programa de Actualización Científica y Didáctica (modalidad A)*", todavía en curso, recoge una serie de actividades formativas cuya finalidad es hacer posible una constante actualización del profesorado no universitario en los distintos campos disciplinares que componen las áreas curriculares. El estudio se aplica a dos tipos de muestras:

- * La totalidad de los cursos que se desarrollan en el seno del Programa A.C.D. Modalidad-A (en concreto, 50 cursos en esta región) de los que se realiza un seguimiento indirecto y extensivo.
- * Una muestra representativa del conjunto de cursos impartidos (un 10% del total: aproximadamente 5 cursos). En este caso se lleva a cabo un seguimiento intensivo y directo.

El marco general en el que se inscribe el conjunto de actuaciones que describe Gairín Sallán en su "*Estudio de las necesidades de formación de los equipos directivos de los centros educativos*" (1994) es el modelo español de dirección escolar, y por tanto se trata de un programa de amplio alcance. La naturaleza de esta investigación y la diversificación de sus fuentes de información delimitan como población de estudio a los equipos directivos de niveles no universitarios, en concreto se lleva a cabo:

- a) Una evaluación demorada de los siguientes cursos:
 - 3 cursos del Ministerio de Educación y Ciencia.
 - 1 curso de la Generalitat de Catalunya.
 - 1 curso del Gobierno Vasco.
- b) Una evaluación de necesidades, tal y como son expresadas por 400 directivos del Estado con diferentes responsabilidades en los centros educativos.

c) Una evaluación comparada de necesidades entre:

- 10 expertos nacionales.
- 10 expertos europeos no españoles.

Por su parte, Nicolás Prieto en el estudio *“Investigación operativa de los proyectos de formación del profesorado en centros de escuelas rurales para el desarrollo curricular en el marco de la reforma del sistema educativo”*, también en curso, se propone analizar proyectos de formación para el desarrollo curricular. Dichos proyectos están dirigidos a profesores de educación Infantil y Primaria de escuelas rurales de la provincia de Valladolid, y la muestra de estudio la componen 71 centros y 133 profesores de esta provincia.

Marcelo García en su estudio *“Facilitar el cambio en la escuela: análisis y evaluación de dimensiones personales, didácticas, organizativas y de asesoramiento en el proceso de innovación y formación en la escuela”* (1994), evalúa proyectos de innovación educativa aprobados por la Consejería de Educación de la Junta de Andalucía que se están llevando a cabo en 50 centros de Educación Infantil, Primaria y Secundaria de las provincias de Huelva y Sevilla. La población objeto de estudio, la constituyen una muestra representativa de 50 centros que llevan a cabo proyectos de innovación, de la totalidad de centros de Educación Infantil, Primaria y Secundaria de las citadas provincias, que tienen aprobada por la Consejería de Educación de la Junta de Andalucía la realización de proyectos de innovación educativa. A partir de esta muestra se seleccionan diez profesores para realizar un estudio de casos.

— *Características definatorias de la evaluación*

Agentes evaluadores

El estudio de Gairín Sallán sobre las necesidades de formación de los equipos directivos de centros educativos es una heteroevaluación realizada por expertos en investigación sobre formación de directivos, que han colaborado con diferentes Administraciones en la planificación, desarrollo y evaluación de programas y cursos en esta área de conocimiento. De hecho, el diseño contempla explícitamente la realización de una evaluación externa que actúa como elemento de contraste del proceso de investigación, con la cual se pretende tener una referen-

cia exterior que valide las decisiones que el equipo de investigación va tomando. Sin embargo, hay que resaltar que es también una evaluación participativa, en cierta medida interna, puesto que prevé incorporar aportaciones de los participantes y se negocian con ellos determinados aspectos de la evaluación: tipo de entrevistas, modelo de información, etc.

Por otra parte, Porlán Ariza en su análisis del Programa de Actualización Científica y Didáctica (Modalidad A), desarrolla un tipo de evaluación mixta: se opta por un análisis de contrastes que haga posible la evaluación, la negociación y la toma de decisiones fundamentadas en dos ámbitos:

- * *Interno*: implica a los profesores participantes en la actividad de formación.
- * *Externo*: evaluadores-investigadores del Plan de Formación.

Se trata pues, de una coevaluación a través de la triangulación de perspectivas y del contraste de puntos de vista entre responsables de las actividades de formación, participantes y evaluadores.

Por otra parte, la investigación de Nicolás Prieto, consiste en una evaluación externa, ya que el equipo investigador, formado por responsables en formación del profesorado de escuelas rurales preocupados por mejorar las ofertas formativas en este ámbito, evalúa proyectos de formación propuestos desde el Ministerio de Educación y Ciencia.

Por el mismo motivo que en el caso anterior, se puede considerar que son evaluaciones externas, aunque ofrecen un mayor o menor grado de participación a los implicados en los programas, las investigaciones de Marcelo García y San Fabián Maroto. El primero de ellos en su investigación para facilitar el cambio en la escuela a través del análisis y evaluación de dimensiones personales, didácticas, organizativas y de asesoramiento en el proceso de innovación y formación en la escuela, incluye, además, una autoevaluación de los participantes en el programa. El estudio de San Fabián Maroto, por su parte, resalta la importancia de la participación activa de los sujetos implicados.

Objetivos y funciones de la evaluación

El objetivo general del estudio de Gairín Sallán es identificar situaciones insatisfactorias individuales, grupales o generales, en el ám-

bito de la formación. En consecuencia, se consideran necesidades derivadas de la respuesta a la normativa vigente, necesidades expresadas, percibidas o comparadas según se refieran al sistema educativo, a la demanda, a las percepciones de los sujetos, o a los resultados de comparar distintas situaciones o grupos respectivamente.

Se realiza una evaluación con doble función —diagnóstica y formativa—, orientada a la toma de decisiones para fundamentar las actuaciones que se vienen realizando en el campo de la formación para mejorar la calidad de las prácticas educativas. En este sentido, se pretende que la información obtenida sea la base para la elaboración de propuestas formativas que tengan en cuenta las siguientes cuestiones: diferentes tipos de directivos (director, jefe de estudios, secretario y coordinadores), su forma de selección, su situación (noveles o experimentados), niveles educativos, distintas modalidades formativas (cursos, seminarios, formación en centros), y diferencias entre Comunidades Autónomas. Además deben relacionarse con los marcos en que se desarrolla actualmente la formación de directivos y con las posibilidades reales que tiene el sistema educativo para asumirla.

En este contexto, la evaluación de necesidades constituye el punto de partida para identificar situaciones deficitarias o insatisfactorias, que reclaman con urgencia la atención y asignación de recursos a la formación de directivos escolares.

Por su parte, Porlán Ariza en su análisis del Programa de Actualización Científica y Didáctica (Modalidad A), pretende alcanzar los siguientes objetivos:

- Caracterizar el modelo de formación, las estrategias y los recursos propuestos en el Programa de ACD-A.
- Contrastar las concepciones de los responsables directos de los cursos con el desarrollo práctico de los mismos.
- Asesorar a las diferentes instancias implicadas en función de los datos obtenidos.

En el proyecto se plantea realizar una evaluación con función formativa en la idea de sistematizar y ampliar el trabajo de asesoramiento y seguimiento realizado por parte del equipo investigador durante cursos anteriores del Programa de Actualización Científica y Didáctica. Es una evaluación orientada a la toma de decisiones y reguladora del proceso de aprendizaje profesional con la finalidad de hacer evolucionar

nar la actividad profesional del profesorado hacia perspectivas más próximas a la reforma.

Nicolás Prieto propone los siguientes objetivos:

- Obtener una caracterización socio-profesional del profesorado destinado en escuelas rurales.
- Elaborar estrategias que permitan describir y evaluar la funcionalidad de la formación recibida para la mejora del desarrollo curricular.
- Conocer la opinión de los profesores que participan en esta modalidad formativa.
- Aportar sugerencias para futuras actividades de formación con profesores de escuelas rurales incompletas.
- Analizar los materiales curriculares generados dentro del proceso formativo y su operatividad para articular la práctica docente.

Esta investigación, al recoger las necesidades sentidas por los profesores de escuelas rurales tras su participación en Proyectos de Formación en Centros, lleva a cabo una evaluación formativa que trata de describir la funcionalidad de la formación que han recibido los profesores tras asistir a uno de estos programas. Está orientada al perfeccionamiento de la propia situación en que se llevarán a cabo los Proyectos de Formación del Profesorado en Centros, mediante su adecuación a las características de los profesores de escuelas rurales, con objeto de que puedan hacer frente en mejores condiciones a las funciones que se les asignan dentro del nuevo marco curricular. Asimismo, se aportan sugerencias para futuras actividades de formación del profesorado.

San Fabián Maroto realiza una evaluación formativa orientada a informar la toma de decisiones, cuya finalidad última es proporcionar a la Administración educativa y a los propios usuarios, orientaciones útiles que posibiliten la mejora de los programas (elaboración de pautas para diseñar programas de formación). Más concretamente se persiguen los siguientes objetivos:

- a) *Describir el contenido y el contexto de los proyectos de formación del profesorado que se han desarrollado en el ámbito de Asturias surgidos y desarrollados al amparo de la convocatoria oficial.*

- b) Analizar desde una doble perspectiva, externa e interna, los diferentes factores que condicionan el desarrollo de las experiencias de formación en centros.
- c) Hacer propuestas para mejorar la formación en centros y trazar las líneas de un modelo de evaluación comprensivo que abarque dimensiones contextuales, procesuales y de producto, asociadas a la dinámica formativa en centros.

Se elaboran criterios y procedimientos que, presumiblemente, permitirán a los profesores evaluar sus propios proyectos y propiciar su reflexión, superando modelos tradicionales de formación que se centran en el profesor como un elemento aislado del resto de la escuela. Asimismo, se quiere proporcionar a la Administración información y orientación para mejorar la calidad de los Programas.

Por último, la investigación dirigida por Marcelo García se puede considerar una evaluación formativa que contribuye al cambio que tienen que experimentar los profesores en cuanto a estilo y modo de enseñanza de acuerdo con los nuevos planteamientos curriculares surgidos tras la reforma educativa. Por otra parte, se pretende proporcionar información teórica y metodológica sobre los procesos de innovación y formación en la escuela. Se destacan los siguientes objetivos generales:

- a) Describir, comprender y analizar los procesos de innovación educativa desde una perspectiva personal, didáctica e institucional.
- b) Analizar el perfil del asesor de formación: descripción de las condiciones profesionales de su actividad y estudio de sus necesidades formativas.

Patrón para la emisión de juicios de valor

La investigación de San Fabián Maroto es una evaluación que, en sentido estricto, no se puede considerar criterial, porque no establece criterios específicos a priori para juzgar la bondad de los Proyectos de Formación en Centros; tampoco puede considerarse normativa, porque no se establecen grupos de comparación; consiste en una valoración de expertos que llevan a cabo el análisis de los Proyectos a partir de la opinión de los asesores de los mismos y de los agentes directamente implicados en su realización (los coordinadores y los propios

profesores participantes) sobre una serie de aspectos, y a partir de su experiencia profesional.

Desde esta perspectiva, las investigaciones dirigidas por Marcelo García y por Gairín Sallán tampoco utilizan grupos de comparación, ni criterios de comparación establecidos previamente. En el caso de Marcelo García, para analizar el proceso de innovación educativa se utilizaron una serie de descriptores (contexto interno, inicio, diseño y desarrollo de la innovación, resultados) sobre los que se recopiló información a través de unos cuestionarios a los que respondieron la mayoría de los participantes en los proyectos de innovación. Del mismo modo, Gairín Sallán en su estudio sobre las necesidades de formación de los Equipos Directivos de los Centros Educativos, se basa en las concepciones y opiniones de los sujetos implicados sobre los objetos centrales del estudio: formación y función directiva.

Asimismo, el estudio de Porlán Ariza evalúa el Programa de Actualización Científica y Didáctica (Modalidad A) mediante la caracterización general de cursos de formación que desarrollan Programas de este tipo y teniendo en cuenta las dimensiones ya mencionadas en el apartado anterior.

Por su parte, la evaluación de Nicolás Prieto utiliza el juicio de expertos y la autorreferencia como instancias de comparación e interpretación. Se analizan los Proyectos de Formación en Centros a partir de las opiniones aportadas por los participantes en los mismos, los asesores responsables de su seguimiento y los coordinadores.

Dimensiones evaluadas y momento de la evaluación

El estudio de Gairín Sallán realiza una evaluación global, en la que se contemplan variables de entrada (formación previa), de contexto (datos personales y profesionales tanto docentes como directivos, formación que se está recibiendo —modalidad, contenido, características—, problemáticas que afectan a la función directiva —formación, tareas, funciones—) y de producto (sondeo y contraste de la opinión de los participantes sobre el programa, y comparaciones entre las respuestas).

Este diseño de evaluación de necesidades incluye una evaluación demorada de cursos de formación ya realizados previamente y propuestos desde la Subdirección General de Formación del Profesorado (M.E.C.) y desde algunas Comunidades Autónomas.

La investigación que dirige Porlan Ariza prevé, en función de los objetivos, una evaluación global con la consideración de variables de entrada resultantes del diagnóstico inicial de los cursos y de cómo son percibidos por sus responsables (modelo de formación que desarrollan, objetivos, estrategias y recursos, y opiniones de los responsables sobre sus experiencias); de proceso (caracterización y modificación de los diseños y seguimiento intensivo del desarrollo de cinco de los cursos analizados previamente); y de producto (*caracterización final* de los cursos).

En la investigación que dirige Marcelo García se analizan los efectos de los programas una vez realizados, tratándose, por tanto, de una evaluación final.

Marcelo García se plantea evaluar el impacto de los procesos de innovación curricular en los profesores, los alumnos y los centros donde éstos se desarrollan. En primer lugar, se estudian las preocupaciones y problemas que se les plantean a los profesores cuando pretenden llevar a cabo en sus aulas un modelo de enseñanza que responda a los principios y *estrategias didácticas propuestas* en los proyectos de formación, a través del estudio de las siguientes variables: personales (acomodación, inseguridad), gestión (dedicación, tiempo), consecuencia (repercusión en el aprendizaje de los alumnos), colaboración entre compañeros, y reajuste (capacidad de modificar el propio proceso de innovación). En segundo lugar, se pretende abordar la dimensión didáctica de los procesos de cambio, es decir, conocer qué uso se da en las clases respecto a las propuestas de innovación realizadas por los profesores. Y por último, se indaga sobre una nueva figura: los asesores de formación adscritos a los Centros de Profesores. Se trata de averiguar en este caso, qué estilos de asesoramiento asumen y practican en relación a los procesos de formación en centros. En este aspecto se tienen en cuenta las siguientes variables: relaciones con la comunidad educativa, *desarrollo del profesorado*, planificación y cambio, comunicación, curriculum, programa instructivo, servicio a los profesores, resolución de problemas y toma de decisiones, evaluación de programas, motivación y organización y desarrollo personal.

En la investigación dirigida por San Fabian Maroto se consideran tres dimensiones del programa: el contexto en el que se inserta (organización escolar), el proceso de desarrollo (incluye los diversos aspectos que conforman el programa así como su puesta en marcha y funcionamiento) y los resultados obtenidos. Para ello se evalúan diversas

experiencias de formación del profesorado desarrolladas en Asturias. En definitiva, se pretende conocer las condiciones en que se desarrollan los Proyectos de Formación en Centros, los factores que facilitan o dificultan la implantación de dichos Proyectos y cómo se han desarrollado y qué efectos han producido.

Nicolas Prieto en su investigación sobre Proyectos de Formación del Profesorado en centros rurales pretende describir y evaluar la funcionalidad de la formación recibida (cómo inciden los Programas de Formación en el curriculum escolar) a través de la valoración de dichos programas por parte de los propios profesores (procedimientos organizativos, estrategias, estructura, recursos). Se analizan, asimismo, los materiales curriculares elaborados dentro del proceso formativo y su operatividad para articular la práctica docente.

— *Diseño y desarrollo de la evaluación*

La mayoría de las investigaciones de este grupo de evaluación de programas de formación permanente del profesorado intentan compaginar las posturas naturalista y experimental. De hecho, las seis que lo conforman —dirigidas por Porlán Ariza, Gairín Sallán, Nicolás Prieto, San Fabián Maroto y Marcelo García— presentan dicha complementariedad.

San Fabián Maroto sitúa su investigación sobre las posibilidades de la formación del profesorado en centros en el paradigma naturalista-interpretativo, orientado por planteamientos filosóficos de carácter fenomenológico, por lo que se adopta un diseño etnográfico centrado en el estudio de casos. Esto posibilita el conocimiento y la comprensión del fenómeno estudiado mediante el examen sistemático y en profundidad del caso concreto dentro del propio contexto donde se produce.

Este estudio emplea un modelo integrador de evaluación desde una perspectiva holística, que contempla la interdependencia del conjunto de actuaciones que tienen lugar en la escuela. Se prioriza, pues, el enfoque cualitativo aunque, sin embargo, se utilizan también técnicas de carácter cuantitativo para el análisis de algunos datos.

Se realiza un análisis de los Proyectos de Formación en Centros desde una doble perspectiva, externa e interna. La primera incluye fuentes de información directas e indirectas, documentales y personales, que se refieren, respectivamente:

- Al contenido de la mayoría de los Programas de Formación en Centros desarrollados en Asturias durante el curso 1992-1993. Para ello se confecciona una Ficha de registro del desarrollo del proyecto de formación que tiene la siguiente estructura: 1) Naturaleza, características y recursos del centro; 2) Naturaleza y características del proyecto; 3) Aspectos organizativos del proyecto; 4) Evaluación.
- A la opinión de los asesores de los Proyectos sobre el modelo de formación en centros, las bases de la convocatoria oficial, el papel que ellos desempeñan, y el desarrollo de los Proyectos.

Desde la perspectiva interna se recaba la opinión de los agentes directamente implicados en la realización de los Proyectos: los coordinadores y los propios profesores participantes. En este caso los procedimientos empleados son:

- Entrevistas a los coordinadores para recoger su opinión sobre el modelo de formación, las bases de la convocatoria, el contexto organizativo, el apoyo externo, su desarrollo y los resultados.
- Cuestionario dirigido a los profesores participantes, que incluye los siguientes aspectos: temática y modalidad de trabajo, análisis de la realidad del centro, objetivos, contenidos y metodología, organización de los recursos, criterios y procedimientos de evaluación, problemas encontrados, y evolución del grupo.

Dada la naturaleza de los instrumentos empleados para recoger información, se realiza un análisis de contenido de carácter descriptivo. A su vez, la interpretación supone, por una parte, una descripción interpretativa de los resultados; es decir, una descripción de las conclusiones esenciales que se derivan del análisis de los datos recogidos y, también, una valoración de sus implicaciones prácticas para el perfeccionamiento de la formación en centros referida, tanto a su diseño y desarrollo, como a su evaluación. De esta manera se resaltan las limitaciones de este tipo de formación y se presentan propuestas de mejora.

En esta misma línea, aunque sin explicitar el marco de referencia conceptual, se encuentra el trabajo de Nicolás Prieto sobre la formación del profesorado en centros escolares rurales. Presenta un diseño

principalmente cualitativo y se prevé la posibilidad de llevar a cabo un tratamiento cuantitativo de los datos obtenidos.

Se adopta un modelo de evaluación que trata de describir la situación concreta generada tras la aplicación de ocho Proyectos de Formación del Profesorado de escuelas rurales que se han acogido a esta modalidad de formación voluntariamente. Fundamentalmente, la investigación se basa en un método descriptivo de estudio de casos, para examinar y analizar la interacción de los factores que producen el cambio de las prácticas formativas basadas en los principios psicopedagógicos de la teoría de la formación del profesorado en el marco institucional de los P.F.P.C. a través del seguimiento longitudinal de las categorías significativas del proceso formativo.

De acuerdo con los objetivos propuestos se formulan cuatro hipótesis de carácter descriptivo que dirigen el desarrollo de todo el proceso de evaluación. Asimismo, se trata de un estudio longitudinal a realizar durante dos años académicos consecutivos.

Las fuentes de información son personales (profesores) y documentales (programas de formación), y la información que se recoge es de tipo cuantitativo y cualitativo. Se elaboraron cuestionarios y encuestas que comprenden preguntas de respuesta cerrada y abierta sobre hechos y opiniones, y preguntas directas e indirectas sobre intenciones de actuación. Posteriormente, y para profundizar en las respuestas, se llevaron a cabo entrevistas, así como variados procedimientos de observación (rejillas de observación sistemática, listas de control, etc.). Para el análisis de documentos, los investigadores emplean los métodos siguientes: intensivo y extensivo, interno y externo y análisis de contenido. En esta investigación la interpretación cualitativa de los datos tiene un peso mayor que el análisis cuantitativo.

También desde esta perspectiva de investigación descriptiva-interpretativa, Porlán Ariza lleva a cabo la evaluación del Programa de Actualización Científica y Didáctica (Modalidad A). Se trata de una evaluación que presenta un enfoque a medio camino entre la evaluación y el asesoramiento, es decir, se desarrollan actividades centradas en la recogida y análisis de datos, junto a otras centradas en la presentación, discusión y negociación del resultado de dichos análisis con las diferentes personas e instituciones implicadas. En definitiva, esto resalta el carácter formativo y procesual de la evaluación.

Se trata de un estudio de casos que combina un enfoque extensivo de análisis de carácter cuantitativo (referido a la totalidad de cursos a

desarrollar en el seno del Programa ACD-A), con un enfoque intensivo, más cualitativo, que focaliza la atención de los investigadores en los casos más representativos.

El equipo de investigación ofrece un esquema que resume con precisión y claridad el desarrollo de su investigación (sus fuentes de información directas e indirectas, los procedimientos de recogida de la misma y las técnicas de análisis):

a. Caracterización inicial de los cursos:

- * Análisis de los materiales institucionales (análisis del contenido semántico): determinación de intenciones y objetivos explícitos; modelo de formación que desarrolla; estrategias, medios y recursos.
- * Análisis de los puntos de vista de los responsables de los cursos a través de cuestionarios: concepciones explícitas sobre la formación permanente; experiencia en formación; primeras impresiones sobre el curso.
- * Elaboración y distribución del Informe de evaluación número 1: caracterización del contexto de partida de los cursos ACD-A, y recomendaciones fundamentales para el diseño.

b. Caracterización y modificación de los pre-diseños:

- * Análisis de los pre-diseños de los cursos (análisis semántico del contenido): modelos de formación que desarrollan; objetivos que persiguen; actividades que proponen; coherencia interna y externa.
- * Elaboración y distribución del Informe de evaluación número 2: caracterización de cada diseño y recomendaciones para su mejora.
- * Asesoramiento presencial para la elaboración del diseño definitivo: poner en evidencia los aspectos positivos y las incongruencias presentes en los diseños; presentar problemas a resolver; contrastar las opiniones de los responsables de los cursos con otras informaciones prácticas y teóricas adecuadas a los problemas detectados; elaboración de la propuesta final de diseño.
- * Asesoramiento no presencial relativo a la preparación de los cursos.

- c. Seguimiento extensivo de los cursos: consiste en un análisis del punto de vista de los responsables de los cursos sobre el desarrollo de los mismos; análisis del desarrollo de los cursos contrastando información procedente de los responsables del curso, los participantes y los evaluadores; elaboración del Informe de evaluación n.º 3.
- d. Caracterización final de los cursos: análisis del punto de vista final de los responsables; análisis comparativo entre el diagnóstico inicial, el seguimiento y el diagnóstico final; elaboración del Informe final de la evaluación.

En cuanto a los procedimientos para recoger información, se prevé la combinación de técnicas cuantitativas (aplicación de cuestionarios) y cualitativas (observación, entrevistas, diarios de curso).

Por su parte, Gairín Sallán también asume una postura ecléctica en su investigación sobre el estudio de las necesidades de formación de los Equipos Directivos de los centros educativos, utilizando métodos y técnicas tanto cualitativas como cuantitativas. Desde este enfoque, y de acuerdo a los objetivos marcados, combina dos diseños para el análisis:

- a. El estudio de casos, por medio del cual se contrastan diferentes modelos de formación de directivos, su rol y sus funciones;
- b. El diseño de investigación *ex-post-facto*, en el que se comprueba la eficacia de las variables consideradas independientes sin disponer de información sobre la situación de partida.

La naturaleza de la investigación propuesta y la diversificación de las fuentes de información, delimitan como población de estudio a los equipos directivos de niveles no universitarios. Sin embargo, no se establece el criterio empleado para seleccionar la muestra, aunque posteriormente ésta se caracteriza en función de las siguientes variables: sexo, edad, titulación, cargo directivo, forma de acceso al cargo, antigüedad, nivel educativo, tamaño de los centros donde ejercen su función directiva, Comunidad Autónoma, localidad y horas de dedicación.

Las fuentes de información son directas e indirectas (personales, documentales e información de campo) dependiendo del tipo de necesidades consideradas. Asimismo, se incorpora como un elemento bási-

co la evaluación demorada de los cursos de formación que se han venido realizando.

Las ofertas de formación en el ámbito de la dirección y gestión de centros se analizan y comparan a partir de las siguientes elementos:

- Opinión sobre temas formativos y de dirección:
 - Expectativas previas a los cursos de formación.
 - Percepción posterior a la asistencia a los cursos de formación.
 - Momento en que se debe realizar la formación.
 - Proceso de formación.
 - Modalidad formativa preferida.
 - Tareas que el directivo debe desempeñar.
 - Problemática del actual modelo de formación.
 - Problemas que afectan a los directivos en su actuación.
 - Nivel de preparación en distintos ámbitos.
 - Importancia de los contenidos de formación.
 - Formación previa al acceso al cargo.
 - Tipos de formación, selección y promoción que se ven como deseables.

- Opinión sobre la formación recibida: sobre el diseño del curso; el apoyo logístico a los curso de formación; los contenidos; problemas y sugerencias sobre la formación recibida.

En la configuración de los instrumentos se consideran los tres modelos de evaluación que sirven de marco de referencia (discrepancia, marketing y toma de decisiones) que dan lugar a los siguientes:

- * Cuestionarios de tres tipos según se dirijan a directivos formados, en formación y sin formación. Se pretende recopilar información de carácter general que permita centrar y orientar las entrevistas posteriores. Se estructuran en tres apartados: datos generales (personales y profesionales); opiniones sobre temas formativos y de dirección; opiniones sobre la formación recibida en otras cuestiones del diseño formativo.
- * Entrevistas estandarizadas y realizadas con el soporte de los cuestionarios ya elaborados.
- * Análisis de documentación:

- Producción científica en relación a la detección de necesidades de formación de directivos: textos, artículos e investigaciones que inciden en el tema.
- Planes y programas de formación de directivos ya aplicados.
- Legislación en relación a la dirección escolar, sus funciones y atribuciones.

Además de la instrumentalización citada, a lo largo de la investigación se desarrollan otros instrumentos que pueden identificarse como:

- Pautas para el análisis de procesos historiográficos y de observación: autoinformes que sobre sí mismo o sobre situaciones observadas realizaron los miembros del equipo de investigación, expertos en el tema de estudio.
- Guía de análisis para expertos, con objeto de ordenar la perspectiva que diferentes expertos podían dar a partir de entrevistas o producciones textuales publicadas.

El tratamiento de la información combina análisis estadísticos y cualitativos, e incluye la posibilidad de diferenciar las aportaciones por Comunidades Autónomas y por tipología de cursos realizados.

Se realiza un análisis descriptivo de las variables por cada colectivo estudiado (directivos sin formación, en formación y formados; profesores e inspectores), procediendo posteriormente a relacionarlas valorando y contrastando las opiniones de directivos, profesores e inspectores para obtener conclusiones.

Desde el punto de vista metodológico, el estudio de Marcelo García, *“Facilitar el cambio en la escuela: análisis y evaluación de dimensiones personales, didácticas, organizativas y de asesoramiento en el proceso de innovación y formación en la escuela”*, adopta, igualmente, un modelo de evaluación mixto, si bien en este caso no existe un diseño de tipo experimental. Se realiza un análisis estadístico descriptivo y un estudio de casos. El muestreo es estratificado en función de los siguientes criterios:

- Provincia: Sevilla/Huelva
- Nivel educativo (Infantil, Primaria y Secundaria).
- Área del currículum.

- Representación de las comarcas de las dos provincias.
- Antigüedad del proyecto.

Se utilizan fuentes de información tanto personales (profesores, asesores de formación y alumnos), como documentales (proyectos de innovación educativa en centros).

Para recopilar información sobre los problemas y preocupaciones del profesorado se utiliza el “Cuestionario sobre Etapas de Preocupaciones”, y se realizan entrevistas estructuradas para recabar las opiniones personales de los profesores sobre los siguientes aspectos de los proyectos de innovación en los que han participado:

- * Datos descriptivos.
- * Experiencia previa en otros proyectos.
- * Motivaciones para desarrollar el proyecto.
- * Expectativas y preocupaciones respecto al proyecto.
- * Necesidades formativas respecto al proyecto.
- * Características del Proyecto: metas y objetivos; metodología de trabajo; recursos; temporalización; ambiente o relaciones entre el grupo de profesores; evaluación del proyecto.
- * Apoyo externo recibido.

Por otra parte, se evalúan las fases en que se desarrollan los procesos de innovación a través de un estudio de casos con diez profesores seleccionados de la muestra inicial. Para ello se emplean los siguientes procedimientos:

- Entrevista para conocer el nivel de uso de la innovación (orientación, preparación, uso mecánico, rutina, integración y renovación).
- Observación en clase para contrastar la información obtenida en las entrevistas.
- Inventario para medir el ambiente de clase.

Con respecto a los asesores de formación se utiliza el denominado “Inventario sobre Necesidades Formativas de los Asesores” construido al efecto.

Los procedimientos de análisis vienen determinados por el tipo de información recogida, tanto cuantitativa como cualitativa. Los datos

cualitativos de las entrevistas y observaciones se analizan con el programa AQUAD 3.0., que permite la codificación, extracción de texto codificado, contraste de hipótesis y comparación de casos a través de la minimalización.

— *Difusión y uso de los resultados de la evaluación*

De las tres investigaciones de este grupo que han sido concluidas en 1994 (Marcelo García, Gairín Sallán y San Fabián Maroto), se dispone de las respectivas memorias finales donde se presentan las principales conclusiones a las que han llegado tras el estudio realizado, seleccionando las ideas esenciales derivadas de los datos recogidos así como sus implicaciones prácticas para la mejora de la formación en centros.

Con respecto a la investigación de Gairín Sallán se destacan las siguientes consideraciones:

- Necesidad de validar con nuevas aplicaciones el modelo utilizado y estudiar su viabilidad para la elaboración de planes de formación en centros.
- La diversificación de ofertas y la concreción de un Plan Marco exige el conocimiento exhaustivo de las propuestas de formación que se están realizando y que no han sido evaluadas y el apoyo decidido a la experimentación de nuevas alternativas. Sólo desde esta información se podrán establecer políticas coherentes con la realidad. Algunos de los aspectos a considerar en ese proceso serían:
 - a) Capacidad para cambiar conductas y actitudes.
 - b) Conexión con las finalidades de la organización en que se aplica.
 - c) Coherencia con el contexto donde la organización se sitúa.
 - d) Uso de técnicas de formación activas.
- La eficacia de los proyectos de innovación en centros queda muchas veces condicionada por la falta de previsión sobre la incidencia que en los aspectos organizativos de los centros tiene la innovación que se pretende. Incidir sobre la realidad escolar supone considerar todas las perspectivas que la configuran y, entre ellas, las relacionadas con el papel de los directivos.
- La aparición y desarrollo de normativa referida a los reglamentos orgánicos y la aplicación de las medidas de calidad propuestas por

el Ministerio de Educación y Ciencia, puede conllevar nuevas necesidades formativas que cabrá incorporar a la propuesta realizada en esta investigación.

Por otra parte, una de las principales aportaciones de la investigación de Marcelo García ha sido contrastar de forma empírica un modelo de análisis para el estudio y evaluación de procesos de innovación educativa; por lo demás, se presenta un análisis meramente descriptivo de los datos recogidos en el estudio.

Por último, San Fabián Maroto en la Memoria Final de la investigación que dirige, recoge las principales conclusiones mediante un análisis interpretativo de los resultados y que incluye una serie de propuestas para la mejora del modelo de Formación en Centros. Así, se derivan diversas consecuencias y posibilidades para mejorar los Proyectos de Formación en Centros, ya que todos los centros deberían contar con programas de formación interna vinculados a su Proyecto Curricular y a su Programación Anual.

— *Metaevaluación*

De todas las investigaciones incluidas en este grupo, tan sólo en la de Gairín Sallán, "*Estudio de las necesidades de formación de los equipos directivos de los centros educativos*" se lleva a cabo una auténtica metaevaluación.

Se utilizan cuatro criterios: valor de verdad (credibilidad), aplicabilidad (transferibilidad), consistencia (dependencia) y neutralidad (confirmabilidad), que se definen en el estudio como sigue:

- Valor de verdad: interroga sobre la confianza que puede ser depositada en los resultados de la investigación y en los procedimientos empleados. Se intenta establecer contrastando la credibilidad de las creencias e interpretaciones del investigador con las diferentes fuentes de las que se han obtenido los datos.
- Aplicabilidad: validez externa en el enfoque empírico-analítico. Intenta determinar la relevancia y las posibilidades de aplicación de las explicaciones/interpretaciones producto de esta investigación concreta, en otros sujetos, otros contextos, otros problemas.
- Consistencia: pretende establecer el grado de relación entre las conclusiones de la investigación y los procedimientos de recogida y análisis de la información.

- Neutralidad: se propone asegurar que los resultados de la investigación están en función del problema estudiado y no de los intereses y juicios de los investigadores.

En cuanto al valor de verdad o credibilidad, se han contrastado diversas fuentes de información, situadas, a su vez, en diferentes perspectivas y posiciones en relación al problema de estudio. Así, sobre las necesidades de formación, han informado tanto directivos (en formación, sin formación, con formación) como inspectores, profesores y responsables de programas. También se han usado diferentes instrumentos respecto a una misma información y se han aplicado en momentos distintos.

Respecto a la aplicabilidad, ofrece una cierta garantía el partir del análisis de situaciones reales y contextualizadas y el permitir en su desarrollo la implicación de usuarios y de responsables de las acciones formativas.

La consistencia o dependencia de los datos respecto a los procedimientos de análisis se consigue por la propia minuciosidad seguida en la confección de los instrumentos y por la diversidad de dimensiones contempladas.

La neutralidad del proceso no se puede garantizar de manera absoluta, si bien el hecho de actuar el equipo de investigación como elemento de confirmabilidad, de mantener un evaluador externo y de aplicar procesos de triangulación (combinando metodologías de estudio y procesos de contrastación de la información), permite mejorar las expectativas.

Además de la atención a estos aspectos, también está garantizado el rigor de la investigación en lo que afecta a la dimensión más cuantitativa, a través del control constante del proceso y de la exhaustividad en la aplicación de los diferentes requisitos que el proceso de construcción de instrumentos ha conllevado.

4.2.4. Evaluación de Programas Educativos para la prevención de drogodependencias en el ámbito escolar

- *Objeto de la evaluación*

A pesar de no ser muy numerosas en este período, las evaluaciones de programas de prevención de drogodependencias en el ámbito escolar financiadas por el C.I.D.E. se han constituido, por la especificidad de su objeto de estudio, en un grupo independiente dentro del apartado de evaluación de programas.

Son evaluaciones que se han llevado a cabo sobre programas realizados en colaboración con distintas instituciones de carácter educativo, con objeto de estimar el grado de adecuación y eficacia de algunos de los programas diseñados para su aplicación en el aula, y cuya finalidad primordial es determinar la situación y nivel de drogodependencia entre los estudiantes de Educación Secundaria, por considerarse, en razón de su edad, una población especialmente vulnerable en lo que respecta al consumo de drogas.

En este grupo se incluyen dos evaluaciones de programas que responden a estos objetivos, una de las cuales todavía se encuentra en curso:

- *Evaluación del programa de prevención escolar "Tú decides"* (Amengal Munar, 1990).
- *Evaluación de programas de prevención de drogas en la escuela* (March Cerdá, en curso).

Para examinar con detenimiento el ámbito y alcance de estos programas, se precisa una aproximación al fenómeno de las drogodependencias en los estudiantes, así como a sus consecuencias, dado que de otro modo no se entendería cuál ha sido el móvil que ha llevado a diseñar determinados programas, y a someterlos, además, a una evaluación sistematizada.

Se puede constatar cómo distintos organismos o instancias públicas y privadas se han comprometido en aunar esfuerzos para la prevención de las drogodependencias, aunque parece no existir todavía una determinación de cuáles son las pautas de actuación más eficaces en este sentido.

Todos los programas de prevención del consumo de drogas participen de un mismo propósito común enmarcado en la educación para la salud, que ha sido definida desde las instancias gubernamentales competentes como "*un proceso intelectual, psicológico y social que comporta actividades útiles para desarrollar la aptitud de los individuos para tomar, con conocimiento de causa, decisiones importantes para su bienestar personal, familiar y social*" (1985)¹³. Sin embargo, esta coincidencia en los objetivos generales permite que se generen una diversidad de programas que responden a planteamientos teóricos y a estrategias distintos, y que persiguen múltiples objetivos, en ocasiones contradictorios entre sí.

¹³ Consejería de Salud (1985): *Propuestas de educación para la salud*. Comunidad de Madrid, Madrid, pág. 14.

El estudio de Amengal Munar (1990) se propone evaluar el programa de prevención de drogodependencias denominado "Tú decides" que estaba siendo llevado a cabo en una muestra de 192 alumnos de segundo ciclo de Educación Secundaria Obligatoria, de distintos centros escolares de Mallorca; además, en dicho programa se implicaron una quincena de profesores de ese nivel. El propósito de la evaluación era estimar el grado de eficacia de dicho programa para decidir su posterior generalización.

Por su parte, en el proyecto de investigación dirigido por March Cerdá, que se encuentra en curso en el momento de efectuar esta revisión, se propone evaluar todos los programas nacionales orientados a la prevención de drogodependencias en los jóvenes, y que, desde una perspectiva educativa sean aplicables en el medio escolar. Su propósito es delimitar las estrategias comunes a todos ellos que se reconocen como de mayor eficacia, e identificar los elementos específicos de cada programa que potencian el logro de sus objetivos. Para ello, el equipo de investigación se propone comenzar haciendo un inventario de los programas preventivos a nivel estatal.

— *Características definitorias de las evaluaciones*

Agentes evaluadores

En ambas evaluaciones se encuentran implicados equipos de investigadores con una trayectoria y experiencia dilatada en la prevención de drogodependencias en el ámbito escolar. Como ya se ha apuntado, en dichas iniciativas se hallan implicadas también distintas instituciones, por entender que esta tarea no es competencia exclusiva de la escuela, dado que en esta problemática confluyen numerosos elementos de orden sociocultural y económico. De este modo, en el programa "Tú decides" han participado organismos como el Servicio de Información y Prevención de Drogodependencias de la Comisión de Sanidad del Consejo Insular de Mallorca, Ayuntamientos, Asociaciones de consumidores, medios de comunicación (periódicos locales) etc., colaborando en la financiación de gastos y apoyando la promoción de dicho programa.

En concreto, la evaluación dirigida por Amengal Munar (1990) es una evaluación interna puesto que ha sido realizada por los propios integrantes del equipo de diseño del Programa "Tú decides". Además, dado que el equipo evaluador y los diseñadores del programa evaluado

son las mismas personas, nos encontramos ante un proceso de autoevaluación del programa.

Del desarrollo de la propuesta de March Cerdá se desprende que se trata de una evaluación externa, dado que se realiza sobre programas diseñados y aplicados por otros.

Objetivos y finalidad de la evaluación

La investigación dirigida por *Amengal Munar* (1990) se centra en la evaluación del programa denominado "Tú decides", que fue concebido con unos objetivos específicos referidos al ámbito de las actitudes. Así pues, su propuesta consistía en dotar a los alumnos/as de habilidades y destrezas para la toma de decisiones y para la identificación de situaciones que requieren tomar una decisión tras valorar unos factores interrelacionados, y que implican determinadas actitudes respecto al consumo de drogas.

Los indicadores que se han empleado en la medición de la eficacia de dicho programa de prevención de drogodependencias han sido:

- Resultados de la conducta medible (consumo de drogas, gasto en drogas).
- Intenciones de conducta y valores manifestados por los encuestados.

En función de estos objetivos se puede afirmar que la finalidad de esta evaluación es formativa, porque no se pretende la mera estimación del valor del programa "Tú decides" para alcanzar el objetivo preventivo que persigue, sino orientar la toma de decisiones para ajustar y mejorar el programa, con el fin de poder generalizarlo si procede.

La otra investigación, coordinada por *March Cerdá* (en curso) se centra en el análisis y la evaluación del impacto de los diferentes programas de prevención de drogodependencias llevados a cabo en el territorio español desde ámbitos educativos, con el fin de conocer las consecuencias y los resultados de los mismos, y de replantear las estrategias y las actividades a desarrollar que sean consideradas más eficaces para la prevención del consumo de drogas en la población juvenil de mayor riesgo. En definitiva, los objetivos que se proponen alcanzar con el desarrollo de la presente investigación son:

1. Conocer el estado de la prevención de las drogodependencias en la escuela.

2. Identificar los programas que existen en esta línea dentro del Estado español, y enunciar sus actividades.
3. Confeccionar un instrumento de evaluación que permita conocer el grado de eficacia de los programas.
4. Realizar una propuesta metodológica para el diseño y la elaboración de programas escolares de prevención de drogodependencias con el fin de optimizar su idoneidad, su eficacia y su eficiencia.

En función de estos objetivos puede decirse que el carácter de la evaluación es doble: por un lado, determinar aquellos programas que son más eficaces (evaluación sumativa); por otro, proponer una propuesta que permita diseñar programas preventivos eficaces a la luz de la experiencia obtenida con otros programas (evaluación formativa).

Por otra parte, al tratarse de una investigación en curso sólo se puede exponer la intención de los autores recogida en el proyecto inicial de dicha investigación, en la que la obtención de una muestra de los programas escolares de prevención de drogodependencias que se están llevando a cabo en la actualidad en España, se atiene a los siguientes criterios:

- Ser programas representativos de lo que se realiza en este campo en las distintas Comunidades Autónomas.
- Estar dirigidos a la población escolar.
- Tener una antigüedad mínima de dos años.

Dimensiones evaluadas y momento de la evaluación

La evaluación llevada a cabo por el equipo de Amengal es de tipo global, dado que ha considerado las cuatro dimensiones clásicas atendiendo a las variables de entrada (que definen a los sujetos que son sometidos al programa); de contexto (rasgos del entorno en que se aplica el programa); de proceso (que han surgido o han afectado a la marcha del programa); y de resultados (situación del consumo y actitudes hacia el mismo detectadas al finalizar la aplicación del programa).

No sucede lo mismo con el estudio evaluativo propuesto por March Cerdá, que posee un carácter más parcial, centrado en la evaluación de productos, es decir, en los resultados ya obtenidos por los distintos programas objeto de estudio.

Atendiendo al momento en que se ha efectuado la evaluación, e independientemente de las dimensiones consideradas o de las variables incluidas en el diseño, ambas evaluaciones son finales, por haberse efectuado al término de un período de aplicación de los programas objeto de estudio.

Patrón empleado para la emisión de juicios de valor

La investigación de Amengal puede definirse como una evaluación criterial, dado que los resultados y los logros del programa "Tú decides" se contrastan con los objetivos previamente establecidos tanto en el programa, como por parte de los propios evaluadores, que en este caso son también los diseñadores del Programa. Puesto que la valoración de la información recogida se ha realizado tomando como referencia el propio programa, la evaluación puede decirse que emplea una instancia de autorreferencia.

Por su parte, el proyecto de evaluación dirigido actualmente por March Cerdá hay que encuadrarlo dentro de un modelo de evaluación mixto en relación a la instancia de referencia empleada. Por un lado es normativa, puesto que compara la eficacia relativa de unos programas de prevención del consumo de drogas con otros. Por otro lado puede considerarse también criterial, puesto que los resultados y características de los programas preventivos se contrastan también con unos criterios independientes, previamente fijados por el propio equipo de investigación para valorar la idoneidad y eficacia de los programas.

— Diseño y desarrollo de la evaluación

Se puede afirmar que la evaluación del programa diseñado por el equipo de Miguel Amengal (1990) se ajusta a un modelo cuyo diseño es de carácter cuasi-experimental.

Las estrategias descritas por dicho equipo en el programa de prevención del consumo de drogas, fueron evaluadas por primera vez durante los cursos 1980-81, 1981-82 y 1982-83. De aquella primera evaluación se concluyó que:

- Tras la aplicación del programa no se constataban efectos contrapreventivos.
- La información de los alumnos/as con respecto al consumo de drogas había mejorado considerablemente.

- El consumo de drogas tales como el tabaco y el alcohol se había mantenido o incluso había aumentado en el grupo de control, es decir, en los sujetos que no participaron en el programa “Tú decides”.
- Los mejores resultados se correspondían con una mayor identificación con los objetivos, es decir, con una mayor toma de conciencia de las repercusiones de sus propias decisiones en materia de consumo, y de un mayor respeto y tolerancia con quienes no adoptan actitudes idénticas a las propias.

Con el fin de mejorar su eficacia, se diseñó de nuevo el material inicial consistente en fichas explicativas que describían situaciones ante las que hay que pronunciarse y tomar decisiones. Este material se hizo más atractivo, despojándolo de contenidos informativos e incluyendo aquellos elementos que durante la experimentación habían permitido dinamizar las discusiones con los alumnos/as. Este nuevo material gráfico, que utiliza un esquema semejante al adoptado por los comics, consta de una parte específica para el profesor y otra para el alumnado.

Esta evaluación se llevó a cabo en centros educativos de Mallorca y en los niveles educativos correspondientes a los últimos cursos de la Educación Secundaria Obligatoria.

Las muestras utilizadas en la experiencia estaban constituidas por 12 centros escolares mallorquines, cuya selección se efectuó en función de distintas variables: titularidad del centro, tamaño del mismo, niveles de enseñanza y estatus socio-económico de los alumnos/as, tras lo cual fueron asignados aleatoriamente a los grupos experimental y de control:

- Grupo experimental, constituido por 192 alumnos de 8.º de E.G.B., de 1.º de B.U.P. y de 1.º de Formación Profesional.
- Grupo de control, constituido por 184 alumnos de los mismos niveles y centros educativos.

Las fuentes de información utilizadas por los evaluadores del Programa “Tú decides” fueron sido de tipo personal y documental, y los procedimientos de obtención fueron la aplicación de pruebas (test y postest) diseñadas para detectar las actitudes y conductas más comunes manifestadas por los alumnos participantes en el programa respecto al consumo de drogas.

La presentación de los resultados de la evaluación se ha realizado en términos descriptivos y explicativos, destacando la modificación sustancial producida en los alumnos tras participar en el programa experimental evaluado.

Mientras que en el grupo de control no se realizó ninguna actividad relacionada con las drogas, el grupo experimental siguió el programa preventivo que consistía en lo siguiente:

- a) Preparación de los profesores: durante el primer trimestre del curso 1987-88 se realizaron unos cursillos para la formación del profesorado en prevención de las drogodependencias. A estos cursillos asistieron, entre otros, los profesores que debían encargarse de la aplicación del programa en los centros del grupo experimental. Los cursillos fueron de tres sesiones de hora y media cada uno, realizadas una vez por semana después del horario laboral, y se centraban en el entrenamiento para la utilización del programa "Tú decides".
- b) Aplicación del programa: cada profesor pudo organizar la aplicación libremente. Lo único que se mantuvo invariable fue la utilización del material de clase. Se pretendía evaluar el programa en un contexto real, lo que implicaba una variedad de situaciones y de factores entre los cuales estaban todos los que conciernen al profesor. Partiendo del hecho de que cada profesor tiene su personalidad, su estilo de trabajo, su formación, y otro conjunto de características que iban a suponer una aplicación particular del método, se trataba de incluir esta realidad en la experimentación del programa, ya que si algo interesaba especialmente era su eficacia en situaciones reales, y no el posible éxito en condiciones difícilmente reproducibles.

A los alumnos se les entregaba una carpeta conteniendo los fascículos correspondientes a las cuatro lecciones en que se divide el programa. Cada fascículo consta de una historieta representativa del tema, una serie de preguntas para orientar la reflexión, una propuesta de situaciones conflictivas sobre las que había que pronunciarse, y una breve exposición de informaciones útiles para la reflexión.

El 1.º fascículo consistía en una introducción general al tema e incluía algunos conceptos básicos que hicieran posible una discusión coherente durante las sesiones posteriores. En el 2.º fascículo se plan-

teaban los distintos procesos de iniciación al consumo de drogas, y en el 3.º se planteaba la posibilidad de ayudar a otro y, llegado el caso, cómo hacerlo.

Se pretendía que el material del alumno fuese de apoyo para el trabajo en clase. La utilización que se hizo del mismo fue, por tanto, variable, y dependía de la forma en que cada profesor organizaba sus clases. También la duración fue variable, oscilando entre 4 y 15 horas de clase, según el profesor. Como se puede observar, el equipo de evaluación fue extremadamente respetuoso con la idiosincrasia de cada profesor, comprobándose, eso sí, si la reflexión con los alumnos se centraba en los puntos que se indicaban en el programa.

Cada profesor dispuso de un manual en el que se ofrecía la información necesaria para aplicar el programa. Además, durante la aplicación del mismo los profesores fueron asistidos por el equipo de investigación, por si se les presentaba algún problema o dificultad en la tarea.

- c) Reunión con los padres de los alumnos del grupo experimental, por cursos. Durante la misma, además de explicar el programa preventivo que iban a seguir sus hijos, se procuraba centrar la discusión sobre algunos factores de riesgo en el consumo de cualquier tipo de droga ligados a la problemática de la estructuración de la personalidad del sujeto, especialmente con la carencia de modelos de conducta adecuados con los que los jóvenes pudieran identificarse.

La recogida de datos se llevó a cabo en dos momentos, mediante la aplicación de un pretest y un postest, y empleando instrumentos de medida confeccionados a tal efecto, que consistían en un cuestionario que constataba las diferencias existentes entre ambos grupos, tras someter al grupo experimental al programa preventivo "Tú decides".

Los resultados que se han recogido con respecto a la efectividad del programa se refieren a tres ámbitos de consumo (alcohol, tabaco y "otras drogas"), y fueron objeto de un análisis de tipo estadístico.:

- a) Consumo de alcohol: antes del tratamiento no había diferencias significativas entre los grupos, aunque el grupo experimental tenía una proporción mayor de bebedores en todas las modalidades, excepto la de fin de semana. Después de aplicar el programa el grupo experimental no muestra cambios. En el grupo de control,

sin embargo, se observa un aumento significativo de la frecuencia en el consumo de alcohol. Se encontró, además, una correlación positiva, de forma que el nivel más alto de consumo identificado en el postest, correspondía a los sujetos que en el pretest también mostraban un consumo elevado de alcohol.

- b) Tabaco: la frecuencia de consumo es mayor en el grupo de control, manteniéndose estable en el grupo experimental.
- c) Otras drogas: para identificar el grado de consumo de otras drogas (heroína, hachís, etc.), los investigadores dicen no haber dispuesto de un número de casos suficiente para poder estimar la significación estadística de las diferencias. En todo caso, ambos grupos parecen haber mantenido sus cifras de consumo.

Las mayores diferencias encontradas entre los grupos se refieren al proceso de toma de decisiones y a las actitudes hacia el consumo: los alumnos del grupo experimental son más conscientes de los riesgos inherentes al consumo de tabaco. Sobre la posibilidad de comprar drogas, en ambos grupos predomina una postura activa de rechazo, aunque en el grupo control parece que aumentan las respuestas emocionales en detrimento de otras más racionalizadas. Estos resultados concuerdan, por tanto, con lo esperado según los objetivos del programa, que eran, según se ha visto, enseñar a los alumnos a tomar decisiones conscientes.

Otra de las conclusiones importantes de la evaluación del programa es que su aplicación no había producido efectos contraproducentes en los alumnos respecto al consumo de drogas.

Este programa puede considerarse un programa específico, en tanto que entre sus objetivos está enseñar a tomar decisiones sobre el consumo de drogas, especialmente de drogas legales (alcohol y tabaco). Sin embargo, va más allá de este propósito, en la medida en que el tema se trata en el contexto de las otras dificultades con las que se encuentran los adolescentes. Este programa es perfectamente integrable, como lo demuestra esta práctica, en programas educativos más amplios como los de promoción de la salud, sirviendo como una posible base para el ulterior desarrollo de programas más amplios.

Dado que la investigación dirigida por March Cerdá se encuentra en curso, respecto a su desarrollo sólo puede referirse lo que se dispone en el proyecto. Los investigadores se proponen emplear fuentes documentales para la recogida de información sobre los programas de

prevención del consumo de drogas existentes en España, limitados a aquellos dirigidos a la población escolar y de aplicación reciente (con no más de dos años de antigüedad). Todo parece indicar que el análisis de los datos será de tipo cualitativo.

— *Difusión y uso de los resultados de las evaluaciones*

El programa "Tú decides" ha tenido una gran resonancia tras su puesta en marcha y evaluación. Se han llevado a cabo aplicaciones del mismo más o menos fieles al original, aunque se han efectuado algunas modificaciones que no afectan de modo sustancial al programa concebido por el equipo de Amengal.

El convencimiento de la eficacia de este programa de prevención de consumo de drogas, especialmente del tabaco y el alcohol, en la población juvenil, y en el medio escolar, ha sido el móvil que ha llevado tanto a los investigadores, como a los centros escolares, a divulgar estos materiales (fichas y cuadernillos) para que profesores y alumnos puedan trabajar con ellos en el contexto de los temas transversales del currículo (educación para la salud, para la tolerancia, etc.).

Los autores, por su parte, han contribuido a su difusión a través de artículos en revistas científicas, tanto sanitarias como educativas.

De la investigación de March Cerdá y colaboradores parece desprenderse la intención de identificar y divulgar los elementos más relevantes que determinan la eficacia de los diversos programas preventivos escolares que existen a nivel nacional, aunque todavía habrá que esperar a conocer los resultados finales del estudio y la proyección de los mismos.

— *Metaevaluación*

En el estudio sobre el programa "Tú decides" no prevé de modo explícito realizar una metaevaluación. Sin embargo, la propuesta de March Cerdá, consiste precisamente en una evaluación de las evaluaciones realizadas con anterioridad de los programas de prevención de drogodependencias.

4.3. Evaluación de Centros

— *Objeto de la evaluación*

En este grupo se incluyen trabajos que, desde distintas perspectivas y con distinto alcance, se centran en el análisis y evaluación de

instituciones y centros educativos. De este modo, se revisan estudios que abordan la evaluación de centros desde una perspectiva global e investigaciones que, por el contrario, se ocupan de aspectos determinados o de dimensiones concretas de la institución objeto de estudio. Además, y por lo que se refiere al tipo de centro que es objeto del análisis, se incluyen investigaciones relativas tanto a centros de enseñanza (de niveles no universitarios y universitarios) como a instituciones de apoyo a los mismos.

Entre éstos últimos cabe incluir el estudio de Sara Morgenstern Pitcovsky y Eustaquio Martín Rodríguez (1993): *La evaluación de los Centros de Profesores. Análisis de su contribución a la Reforma educativa*. Este trabajo constituye una investigación evaluativa sobre la red de Centros de Profesores (C.E.P.), en cuanto modelo institucional que, a partir del año 1985, se responsabilizaba de la formación permanente del profesorado no universitario, constituyendo plataformas estables apoyadas por la Administración para el trabajo en equipo de los profesores.

Concretamente, el ámbito de estudio de esta evaluación estuvo definido por aquellos C.E.P. creados por el Ministerio de Educación y Ciencia durante el primer año de implantación (en total 56) y que, por lo mismo, contaban con una cierta biografía institucional en el momento de la recogida de información.

Asumiendo como objeto de evaluación estos mismos centros, pero con un ámbito de estudio claramente distinto, se revisa la investigación dirigida por Isidro Moreno Herrero con el título *No llorarse las mentiras sino cantarse las verdades: el cuestionamiento de nuestro quehacer institucional* (en curso). El trabajo se centra, igualmente desde una perspectiva global, en la evaluación del C.E.P. de Getafe, y se conceptualiza, en consecuencia, como una evaluación institucional.

Situado ya en la perspectiva del análisis de centros de enseñanza, el estudio de M.^a del Carmen Palomares Aguirre, *Evaluación del Programa de Colegios Rurales Agrupados de Cuenca*, actualmente en curso, analiza el modelo organizativo que constituyen los Colegios Rurales Agrupados (CRA), implantados a partir del año 1986 por la Administración educativa. El ámbito de la evaluación estuvo constituido, concretamente, por un total de cinco Colegios Rurales Agrupados de la provincia de Cuenca.

Dentro de las evaluaciones globales de centros educativos, se revisan, por último, dos trabajos llevados a cabo en el ámbito de la Educación Superior.

De este modo, la investigación dirigida por Fco. Javier Tejedor Tejedor, *Evaluación de las condiciones personales, materiales y funcionales en las que se desarrolla la docencia en la Universidad de Salamanca* (en curso), asume y parte de los estudios recientes sobre la evaluación del profesorado universitario, como estimación del nivel de calidad de la enseñanza universitaria. A partir de esta dimensión se amplía el marco de referencia al conjunto de elementos que intervienen en la docencia universitaria. A juicio de los autores, esta perspectiva conlleva necesariamente la integración del tratamiento del tema en el marco más amplio de la evaluación de centros. En este estudio no se ha considerado otra gran dimensión de la actividad universitaria, la investigación, centrándose la evaluación exclusivamente sobre la actividad docente. El ámbito de la investigación comprende, en concreto, la totalidad de los centros, departamentos, alumnos y profesores de la Universidad de Salamanca.

Por su parte, el estudio dirigido por Ricardo Marín Ibañez *La evaluación de la enseñanza superior a distancia* (1986), se compone de dos evaluaciones institucionales correspondientes a otros tantos Centros Asociados de la Universidad Nacional de Educación a Distancia (U.N.E.D.). Los Centros Asociados de Palma de Mallorca y de Mérida, objeto de evaluación, son considerados representativos de la U.N.E.D. en el ámbito geográfico insular y peninsular, y se caracterizan por responder a fuertes demandas sociales en Educación Superior.

Finalmente, se incluyen dos investigaciones que se centran en aspectos o dimensiones concretas de un centro escolar. Así, el estudio encargado por el C.I.D.E. a Carmen González Enríquez y Ana María Delgado Solís, titulado *Estudio sobre la utilización del equipamiento didáctico en centros públicos de Enseñanza Básica y Media* (1986), puede considerarse un trabajo específico de diagnóstico-evaluación de centros educativos. Se preocupa de un elemento concreto de las instituciones escolares: los recursos didácticos, su disponibilidad y uso por parte de los docentes. El ámbito del estudio estuvo constituido por los centros públicos de Enseñanza General Básica, B.U.P. y Formación Profesional del territorio gestionado por el Ministerio de Educación y Ciencia.

Igualmente, la investigación dirigida por José Antonio Del Barrio del Campo bajo el título *Evaluación de criterios ecológicos sobre percepción de centros escolares y estrategias para un plan de interven-*

ción (1986), se centra en un aspecto concreto de los centros, habitualmente incluido en los modelos generales: la evaluación del ambiente o clima institucional. Concretamente esta evaluación se lleva a cabo a partir de las percepciones de los sujetos implicados directa o indirectamente en la estructura y funcionamiento de 15 centros de Enseñanza General Básica y B.U.P. de la Comunidad de Cantabria.

— *Características definitorias de la evaluación*

Agentes evaluadores

La práctica totalidad de los estudios incluidos en este grupo reúnen los rasgos que definen a los procesos de evaluación externa, dirigidos y llevados a cabo por técnicos y que no contemplan entre sus procedimientos metodológicos la autoevaluación. Sólo el estudio evaluativo del C.E.P. de Getafe, dirigido por Moreno Herrero, destaca en este sentido por sus características netamente distintas.

Efectivamente, se prevé y diseña en este caso un proyecto de autoevaluación institucional de carácter interno. La investigación se articula a partir del trabajo de un "equipo de evaluación", compuesto por agentes con distintos grados de implicación. Este equipo está compuesto por todos los miembros del Equipo Pedagógico del propio C.E.P., al que se une un número reducido de agentes externos (miembros del Consejo del C.E.P., de los centros escolares de la zona, de los Equipos de Sector y expertos universitarios).

Por otra parte, el estudio dirigido por Tejedor Tejedor sobre la Universidad de Salamanca se puede considerar una evaluación externa, a pesar de que el propio director y la totalidad de los componentes del equipo investigador mantienen relaciones laborales con la institución evaluada. Asumir el carácter interno del proceso evaluativo exigiría considerar como unidad estructural y funcional del análisis la Universidad de Salamanca en su conjunto. Si consideramos como unidades de referencia sus distintos centros, o los departamentos, perspectiva posiblemente más adecuada, la evaluación habría de calificarse más propiamente como externa (los agentes evaluadores pertenecen a un Centro particular, y aún a un Departamento universitario concreto). Igualmente cabría apuntar que, si bien este estudio consiste básicamente en una heteroevaluación, incorpora en su diseño y desarrollo ciertos procesos particulares de autoevaluación (por ejemplo en relación al profesorado).

Objetivos y funciones de la evaluación

Aunque no se formulan explícitamente los objetivos concretos de la evaluación sobre la red de Centros de Formación de Profesores llevada a cabo por Morgenstern Pitcovsky y Martín Rodríguez, éstos se apuntan, de modo más o menos formalizado, a lo largo del estudio. Concretamente, puede decirse que el estudio tiene como objetivo general, la valoración de la trayectoria y los logros de los C.E.P., su quehacer institucional en tanto que responsables del perfeccionamiento del profesorado, desde la perspectiva de su contribución a la reforma del sistema educativo. Junto a este objetivo general, se define el estudio como una primera aproximación centrada en las tendencias generales, asumiendo la existencia de una problemática, compartida por muchos C.E.P., que tiene características estructurales y no institucionales.

El estudio adopta una perspectiva calificada por los autores como evaluativa, afirmándose que “los resultados y conclusiones de este estudio no pretenden otra cosa que contribuir al conocimiento de la dinámica y trayectoria institucional de los C.E.P.”. Asumido el carácter independiente del estudio, se pretende en definitiva “contribuir de una manera problematizadora a la construcción de un modelo innovador y efectivo de perfeccionamiento profesional”.

Esta orientación formativa es la que quiere articularse, adoptando un enfoque de evaluación que comprende el diagnóstico de las situaciones, su valoración, y la propuesta de las líneas de acción necesarias. Junto a lo anterior, y como parte sustantiva de las conclusiones del estudio, se incluyen una serie de sugerencias para la reflexión, el reajuste y la mejora, dirigidas a la propia Administración, responsable de la red de C.E.P.

Esta orientación a la toma de decisiones para la reconducción y la mejora, es igualmente clara en el trabajo diseñado para la autoevaluación institucional del C.E.P. de Getafe. Los objetivos a que se ordena tal evaluación, entendida como valoración crítica del quehacer educativo del C.E.P. en y desde el propio C.E.P., son, de acuerdo con los autores, los siguientes:

- Clarificar elementos y criterios de actuación que posibiliten la mejora del funcionamiento interno del C.E.P., la elaboración del Proyecto del centro, su concreción en los planes anuales y el desarrollo de las distintas actuaciones de formación.

- Desarrollar procesos de autoformación del equipo pedagógico del C.E.P. en coherencia con el modelo de formación por el que se ha optado.
- Introducir dinámicas de autoevaluación permanente en todos los ámbitos de la institución, experimentando en el equipo pedagógico lo problemático y enriquecedor de estos procesos.
- Generar un modelo útil de autoevaluación institucional, que experimentado y vivido en la institución, pueda ser ejemplificador, proyectarse y contrastarse en instituciones similares.

Se prevé y diseña, por lo tanto, una autoevaluación incardinada en las estructuras y en los procesos institucionales mismos, con un carácter netamente formativo, en cuanto orientada a la mejora de los distintos elementos institucionales y a la reconducción de los procesos y factores del centro como fruto de la información y reflexión generada a partir de la misma.

Desde un enfoque considerablemente distinto, la evaluación diseñada para la Evaluación de los Centros Rurales Agrupados (C.R.A.) comparte con los dos estudios de evaluación de Centros Asociados de la U.N.E.D. una característica: formulan hipótesis que tratan de ser contrastadas en la investigación. De este modo, los objetivos de tales estudios consisten en última instancia en obtener conclusiones a partir del contraste de las hipótesis, como es habitual en la investigación aplicada.

El objetivo general de la evaluación de los C.R.A., llevada a cabo por Palomares Aguirre, es comprobar si, efectivamente, éstos centros cubren los objetivos para los que fueron creados, y dan respuesta a las necesidades que se trataba de satisfacer: contribuir a un mayor desarrollo educativo de los alumnos de zonas de poblamiento disperso, y favorecer las condiciones de igualdad de las zonas rurales en el terreno educativo. A partir de este objetivo general, la evaluación se diseña concretamente para contrastar las siguientes hipótesis:

- Los C.R.A. son organizaciones de centros para zonas rurales que favorecen la práctica docente y mejoran la calidad de la enseñanza así como el rendimiento educativo de los alumnos.
- En los centros incompletos de menos de 3 unidades que *imparten Ciclo Superior*, el rendimiento educativo de los alumnos es menor que el conseguido en los C.R.A.

- Los C.R.A. favorecen las condiciones de igualdad educativa en zonas rurales desfavorecidas, paliando las desigualdades socioeducativas que sirven de freno en el acceso a estudios superiores.

Se entiende que el análisis de los datos de la evaluación hará posible el replanteamiento de los C.R.A., es decir, realizar nuevas modificaciones o mantenerlo como modelo organizativo para la implantación de la nueva etapa de Educación Primaria que entonces comenzaba a producirse. Esto es, se pretende obtener resultados válidos y fiables que permitan confirmar o replantear la conveniencia de este tipo de organización. Parece deducirse, en consecuencia, una orientación clara hacia la toma de decisión, y una finalidad de naturaleza formativa, orientada a la formulación de propuestas de reconducción y de mejora.

Por su parte, la evaluación llevada a cabo en el Centro Asociado de la U.N.E.D. en Baleares, tiene por objetivo general analizar el rendimiento de la Enseñanza Superior a distancia en Baleares. Las hipótesis que guían tal análisis son las siguientes:

- El perfil del alumno de la U.N.E.D. en Baleares responde a las características sociodemográficas y culturales asignadas tradicionalmente al alumnado de la U.N.E.D. en general.
- Las Unidades Didácticas de la U.N.E.D., por su calidad, pueden sustituir la acción personal del profesor, y por lo tanto el alumno puede lograr por su cuenta los objetivos programados.
- Existen relación entre la opinión de alumnos y de tutores acerca de los materiales didácticos.
- Los materiales de apoyo utilizados en la U.N.E.D. (cassetes, emisiones radiofónicas, guías, etc.) son necesarios y desempeñan un función importante en la Enseñanza Superior a distancia.
- Las opiniones de alumnos y tutores del Centro Asociado acerca del sistema de evaluación son positivas, y no difieren entre sí de manera notable.
- El rendimiento académico del Centro Asociado de la U.N.E.D. en Baleares no difiere del rendimiento académico de la U.N.E.D. a nivel nacional, y está relacionado con el sexo de los alumnos, siendo superior en su totalidad el de los alumnos varones.

- El perfil del licenciado por la U.N.E.D. en el Centro Asociado de Baleares responde al del alumnado general de la U.N.E.D., y como grupo, los licenciados comparten los rasgos más definitorios de la población nacional.

La evaluación institucional del Centro Asociado de la U.N.E.D. en Mérida se define como *“un análisis de la eficacia de la U.N.E.D. en Extremadura, o del grado de consecución de los logros que se pretendían alcanzar”*. Concretamente el estudio se propone contrastar la siguiente hipótesis central: *“la Educación Superior a distancia en Extremadura supone una alternativa eficaz y de resultados al menos similares a los de la enseñanza presencial”*.

A partir de esta hipótesis, se formulan hipótesis secundarias en torno a dos presupuestos:

- *“El centro Asociado de la U.N.E.D. en Mérida tiene justificada su creación y función actual, dada la población atendida, las actividades desarrolladas y el escaso coste por alumno”*.
- *“Existen diferencias significativas entre el perfil y opiniones del alumno actual y los de los primeros que obtuvieron la licenciatura”*.

Las finalidades atribuidas a los procesos de evaluación llevados a cabo en los Centros Asociados de Mérida y Baleares no se presentan de modo explícito. En todo caso, y del análisis del material disponible, no parece derivarse una orientación prioritaria a la toma de decisiones. Los diseños y planteamientos del estudio apuntan más bien al establecimiento y justificación valorativa del estado actual de las instituciones evaluadas, desde una perspectiva asimilable a la filosofía de “rendición de cuentas”.

En el trabajo evaluativo sobre la Universidad de Salamanca, los autores pretenden analizar los distintos factores y dimensiones identificadas como relevantes para una evaluación global y comprensiva del componente “docencia” en la institución universitaria, así como las relaciones existentes entre los mismos. Se proponen como intenciones fundamentales del estudio, el conjunto de objetivos que suelen ser propios de las evaluaciones de contexto-proceso-producto. Los objetivos específicos de cada fase, son los siguientes:

- Evaluación de contexto: identificar los elementos que caracterizan a cada centro, señalar carencias o necesidades, posibilitar la evaluación de procesos y de productos, y cuestionar la validez de los actuales elementos organizativos.
- Evaluación de proceso: reorientar estrategias y actividades que se llevan a cabo para la consecución de los objetivos planteados, proceder al seguimiento de la actividad organizativa en centros y departamentos, y evaluar objetivamente la actividad docente realizada.
- Evaluación de producto: cuantificar y analizar logros, relacionar variables implicadas en el modelo, diagnosticar las causas de éxitos y fracasos, y formular propuestas de corrección.

Junto a estos objetivos analíticos, que se refieren básicamente a las dimensiones consideradas en la evaluación, se apuntan también los dos siguientes:

- Determinar objetivamente el cuadro de necesidades materiales, personales y organizativas que condicionan la calidad de la docencia universitaria.
- Establecer prioridades y sugerir estrategias para la mejora de la docencia universitaria.

Estos dos objetivos finales, junto a aquéllos propios de la evaluación integradoras del contexto, los procesos y los productos, dan a la evaluación diseñada una orientación hacia la toma de decisiones de mejora, por lo tanto con función formativa. Así, entre los elementos de la planificación del proceso, y concretamente en su fase final, se prevé la determinación de necesidades y la formulación de sugerencias de intervención (prioridades, dotación de recursos, actividades formativas y proyección al ámbito socio-profesional).

Por su parte, el objetivo del estudio de González Enríquez y Delgado Solís sobre la utilización del equipamiento didáctico en centros de enseñanza, fue indagar sobre el grado de aprovechamiento real del material didáctico, con la hipótesis de partida (implícita) de la existencia de una cierta infrautilización del equipamiento, por motivos relacionados con su forma de distribución, de conservación y de integración en las aulas y en los programas de estudio. La investigación trataba, por tanto, de aportar información descriptiva, y también expli-

cativa, que permitiera una racionalización del gasto público en esta materia. Por lo mismo, entendemos que constituye un estudio orientado a la toma de decisiones de mejora de la eficiencia.

En el trabajo dirigido por Del Barrio del Campo sobre percepción de centros escolares, no se lleva a cabo una exposición sistemática de los objetivos, más o menos operativos, del estudio. En todo caso, el estudio parece orientarse al establecimiento de conclusiones generalizables sobre la percepción que los distintos agentes de la comunidad escolar tienen del funcionamiento y de la dinámica de los centros escolares. Así, se apunta en el apartado de conclusiones: "La validez externa de los resultados, (...) es desigual en función del número de participantes y de su representatividad (...)". A continuación se afirma: "Con todo esto, creemos que podemos presentar una buena panorámica de la dinámica existente en los centros escolares (...)” (p. 434).

Asimismo, y como ya se apunta en el propio título del trabajo, puede atribuirse al estudio una orientación hacia la sugerencia de algunos factores que deberían ser considerados a la hora de intervenir en contextos escolares. En definitiva, se pretendió llevar a cabo una evaluación inicial del clima institucional con el objetivo de establecer conclusiones y sugerencias de intervención que pudieran ser generalizables a otros centros, situándose en un modelo de estudio más propio de la investigación aplicada que de la estrictamente evaluativa.

*Dimensiones evaluadas y momento de la evaluación*¹⁴

El estudio evaluativo de la red de Centros de Profesores llevado a cabo por Morgenstern Pitcovsky y Martín Rodríguez, parece comprender, adoptando la terminología clásica, dimensiones del contexto y de los productos o logros de la institución.

De este modo, y tratando de aportar informaciones relevantes que permitan ponderar todo el estudio, se lleva a cabo una evaluación de los contextos institucional y personal de los C.E.P.

Como características definitorias del contexto institucional se incluyen: la localización de los C.E.P. (contexto geográfico); los Movi-

¹⁴ Se entiende que, salvo excepciones, el momento en que opera la evaluación no resulta una información relevante para este tipo de estudios, al versar sobre instituciones supuestamente consolidadas desde el punto de vista de su estructura y funcionamiento. Los datos relevantes en este sentido se han introducido, cuando era oportuno, bien en la presentación de la evaluación, bien en el desarrollo de la misma.

mientos de Renovación Pedagógica en el entorno del C.E.P. (contexto cultural); la infraestructura (características del edificio, dotaciones, libros, revistas y videos); y la organización y funcionamiento de los C.E.P. (sus relaciones con el exterior, la planificación y gestión institucional y las tareas desempeñadas por sus responsables). Además, a partir del análisis de las características del personal del centro, se trata de inferir tendencias de reclutamiento, motivaciones, actitudes ante el cambio profesional y ante la propia formación; asimismo se persigue una aproximación más objetiva hacia posibles desequilibrios en la composición del personal, en función de características como la titulación o el nivel educativo del que provienen.

Junto a esta valoración de las condiciones contextuales de la institución, se lleva a cabo una evaluación de logros, o del avance en la consecución de lo que pueden considerarse objetivos institucionales de los C.E.P.

De este modo, se analiza y valora la influencia del C.E.P., o grado de "penetración" en su entorno. Se consideran a tal efecto como indicadores, la participación del profesorado en los C.E.P. y los motivos y razones de la misma. Además, y con el fin de evitar un enfoque meramente eficientista, se incluye de nuevo una dimensión contextual, de naturaleza, no obstante, más subjetiva que la que se consideró inicialmente. Concretamente, el contexto de actuación de los C.E.P. se valora a partir de la actitud de los docentes ante la reforma educativa de la L.O.G.S.E., la receptividad de los docentes ante las actividades de formación, y la imagen del C.E.P. y sus agentes entre el profesorado.

Junto a lo anterior, y considerando que una parte muy importante de las tareas asignadas a un C.E.P. giran en torno a la concreción de su oferta de formación permanente, se evalúan las actividades llevadas a cabo por éstos en este campo. Concretamente, se valoran los criterios de organización y planificación de las actividades de formación, y las características de éstas (temática, duración, acreditación y evaluación, etc.).

La consideración de estas dimensiones básicas permitió igualmente una valoración más específica de la contribución de los C.E.P. a los procesos de reforma, tratando de determinar si efectivamente se habían producido los cambios suficientes en la formación permanente del profesorado que hicieran de este modelo un esquema coherente.

Para ello, y junto con la valoración de la opinión del personal del C.E.P. con respecto a la reforma, se retoma desde esta perspectiva la planificación de actividades, y se valora la relación de los C.E.P. con los centros experimentales.

Por su parte, la evaluación del C.E.P. de Getafe se articula sobre tres “macrodimensiones”, de acuerdo con la terminología empleada por los autores:

- Se prevé llevar a cabo, en primer lugar, una evaluación del diseño del *Plan del C.E.P.*, a partir de la consideración de los planteamientos y mecanismos relativos al diagnóstico de las necesidades de formación, al modelo de formación, a la definición de actividades prioritarias, a la tipología de actividades, a los modos de selección y acreditación de los participantes, a las opciones organizativas y metodológicas, a la evaluación del desarrollo y a la ejecución del Plan, etc.
- La *organización y funcionamiento internos* serán valorados a partir de las siguientes dimensiones entre otras: organigrama formal y funcional del C.E.P., distribución de cometidos y tareas, ámbitos de responsabilidad en la gestión, formas de participación y toma de decisiones en los procesos de planificación, coordinación y distribución de la información, dinámicas de relación, etc.
- Por último, la consideración de las *actuaciones de formación* se entiende que permitirá valorar el desarrollo, ejecución y resultados de una gran parte del Plan del C.E.P., proporcionando información acerca de cómo se llevan a cabo efectivamente las acciones proyectadas, y de cómo inciden éstas en la mejora de la realidad educativa.

Se trata, por lo tanto, de una evaluación que incorpora el análisis de dimensiones de entrada, de los procesos y los productos de la institución.

En el trabajo de evaluación de los Centros Rurales Agrupados en la provincia de Cuenca, y de acuerdo con las dos subhipótesis que se pretendían contrastar, se consideran principalmente dos ámbitos para la evaluación: el contexto en el que se inscriben los C.R.A., y los rendimientos educativos. Más concretamente, las dimensiones que se analizan, de acuerdo siempre con la autora, son:

- Contexto: zona geográfica y características
- Nivel jurídico: plantilla de profesores y distribución en secciones, material didáctico y espacios, distribución del alumnado.
- Nivel didáctico: órganos unipersonales y colegiados, reuniones de ciclo y periodicidad, constitución de Asociaciones de Padres, rendimiento académico (en los ciclos Inicial, Medio y Superior, y obtención del Graduado Escolar o del Certificado de Escolaridad), y acceso a programas de innovación y reforma (Orientación en centros, Educación Física por especialista, oferta de idioma en Ciclo Medio, Integración y Educación Infantil).

Posteriormente, en el informe de progreso de la investigación se informa de la selección de cinco ámbitos para la evaluación de las condiciones de igualdad entre los centros, y que deberán incluirse en todos los instrumentos de recogida de información que se elaboren como referentes de la calidad educativa de los centros: recursos humanos y materiales, organización del centro, rendimiento educativo de los alumnos, acceso a programas de innovación y satisfacción docente.

Por su parte, y como ya se adelantó, el estudio dirigido por Tejedor Tejedor sobre la Universidad de Salamanca incluye como dimensiones evaluativas: el contexto, los procesos y los productos de la institución universitaria evaluada.

De este modo, se apuntan como categorías generales que incluyen las variables de contexto: características del alumnado (sociológicas, psicológicas y pedagógicas), recursos del centro (materiales, personales y servicios), características del profesorado (ratios, titulación y formación, etc.) e indicadores relativos a la relación Universidad-sociedad.

Las informaciones relativas a los procesos se agrupan en las siguientes categorías: organización departamental, desarrollo de la docencia en el aula, funcionamiento del centro y clima institucional. Finalmente, como productos institucionales se consideran el rendimiento y la satisfacción de los alumnos y del profesorado.

También en el estudio evaluativo de la Educación Superior a Distancia podría considerarse que se incluye como dimensiones el contexto, los procesos y los productos de la institución, si bien no son con-

ceptualizadas en estos términos por los autores, ni aparecen claramente definidas, sistematizadas o sistemáticamente tratadas.

Concretamente, para la evaluación del Centro Asociado de Baleares, y como ya se apuntó en parte cuando se presentaron las hipótesis, se consideraron los siguientes factores:

- Características sociodemográficas y sociolaborales del alumnado para establecer el perfil del estudiante adscrito al Centro Asociado de la U.N.E.D. en Baleares.
- Evolución de las tasas de matriculación de alumnos por curso y carreras del Centro Asociado de Baleares desde su creación.
- Calidad del sistema didáctico multimedia de la U.N.E.D. (Unidades Didácticas de carácter obligatorio, cassettes, emisiones de radio, video-cassettes, guías del curso, programas, etc.).
- La eficacia del sistema de evaluación utilizado por la U.N.E.D., así como los resultados académicos obtenidos por los alumnos mediante dicho sistema.
- Resultados o productos de la institución:
 - Cuantificación y análisis de los licenciados en el Centro Asociado durante los ocho primeros años, incluyendo su comparación con los datos nacionales de la U.N.E.D. y con el conjunto de alumnos del centro.
 - Tasa de éxitos y porcentajes de abandono.

En este caso, se estudian los nueve cursos de funcionamiento del centro, desde el comienzo de su actividad académica en el curso 1977/78.

Por lo que se refiere al estudio sobre el Centro Asociado de Mérida, los aspectos generales evaluados para el contraste de la hipótesis principal, relativa a la eficacia de la institución, fueron los siguientes:

- Necesidades de creación del centro y ubicación geográfica.
- Labor académica.
- Educación permanente, promoción cultural y perfeccionamiento profesional.
- Investigación y publicaciones.

- Aceptación del centro.
- Función social.
- Costes.

En el estudio de Del Barrio del Campo centrado sobre la evaluación del clima institucional, las principales variables consideradas son las siguientes:

- Marco social y participación de los padres (existencia y eficacia de las Asociaciones de Padres).
- Condiciones materiales del centro (existencia de instalaciones y de servicios complementarios en el centro escolar).
- Programación (tipo de programación considerado más adecuado y llevado a cabo por el profesor, tipo de evaluación seguida y material didáctico utilizado).
- Percepción de la tutorías y de la dirección (existencia, tiempo dedicado y eficacia de la labor tutorial, y buen funcionamiento de la dirección).
- Interacción alumno-profesor (percepción de profesores y de alumnos).
- Percepción profesorado (por padres y alumnos).

Finalmente, en orden a conocer y analizar las dotaciones de los centros y el uso que de ellas se hace, el trabajo de González Enríquez y Delgado Solís recoge información en los centros públicos de los siguientes aspectos:

- Instalaciones complementarias a las aulas y material didáctico (inventario general del centro).
- Material disponible para el profesorado, suficiencia/insuficiencia del mismo, estado de conservación, mecanismos para el mantenimiento de aparatos, formas de adquisición del material, etc.
- Política estatal de distribución del material a los centros.
- Uso concreto de los elementos didácticos (segunda fase del estudio):

- Elementos didácticos más y menos apreciados por los profesores.
- Materiales no usados y motivos de esa falta de uso.
- Observaciones relativas a los elementos considerados más relevantes (uso, problemas y rendimiento de cada elemento utilizado): el retroproyector, el proyector de diapositivas, el proyector de cine, el video, el magnetofón, los equipos de laboratorio (materiales específicos de Física, Química y Ciencias Naturales), y el ordenador.

Se analiza también la opinión sobre la formación que sería necesaria para manejar los aparatos, y algunos aspectos relativos a la producción propia de material de paso.

Patrón empleado para la emisión de juicios de valor

La mayor parte de los estudios revisados adoptan una perspectiva comparativa o normativa en la elaboración de conclusiones y juicios de valor. Sólo en el trabajo de Morgenstern Pitcovsky y Martín Rodríguez sobre la red de C.E.P. la evaluación adopta una perspectiva próxima a la criterial, a partir de la definición del marco valorativo que preside el análisis.

En este trabajo se formularon y adoptaron desde el comienzo un conjunto de criterios de evaluación, que expresaban los objetivos de la reforma que han privilegiado los investigadores por su relevancia para reconducir la actualización del profesorado. Estos criterios son los siguientes:

- Implantación del C.E.P. en la zona.
- Capacidad de promover la innovación educativa, a través de actividades que contemplen:
 - Prioridades curriculares tanto en lo relativo al contenido como a su didáctica.
 - Proyectos de formación en los propios centros.
 - Proyectos inter-centros, interdisciplinares e inter-nivelaes.
 - Importancia asignada a la enseñanza Técnico-Profesional por ser la rama con menor tradición en la formación docente.

- Introducción de temáticas relevantes, aunque no tradicionalmente “ortodoxas”.
- Apertura a la reflexión:
- Capacidad de los miembros del C.E.P. para revisar y corregir los planteamientos institucionales y para autoevaluar su papel como gestores y promotores de la innovación.
 - Trascender una formación particularista para lograr un visión comprensiva de lo educativo.
 - Planificación de la autoformación del equipo.
 - Conciencia de la necesidad de superar el pragmatismo introduciendo de forma creciente la fundamentación teórica.
- Creación de la necesidad de la formación permanente.
- Racionalidad global de la planificación:
- Equilibrio en la programación de actividades por ciclos y áreas.
 - Complementariedad entre los contenidos disciplinares y la fundamentación psicopedagógica.
 - Traducción de las prioridades formativas en la elaboración de los presupuestos.
 - Seguimiento y evaluación de las actividades individuales.
 - Relación costo-beneficio de las actividades realizadas.

No obstante, y a lo largo del estudio, se encontraron condicionantes que relativizaron la aplicación de dichos criterios. Así, el análisis de la racionalidad global de la planificación institucional fué muy restringido porque, de acuerdo con los investigadores, no se dispuso de toda la información pertinente. En todo caso, las conclusiones y valoraciones, tanto generales como particulares, se derivan de la interpretación de los resultados obtenidos a la luz de los criterios de evaluación apuntados.

Por su parte, la evaluación de los Centros Rurales Agrupados adopta una perspectiva fundamentalmente comparativa. En este sentido los autores apuntan: *“las conclusiones se establecerán al tomar*

como referencia la comparación entre dos tipos de agrupamientos y su funcionamiento como centros educativos en zonas rurales". Efectivamente, el diseño de la evaluación, como se verá más adelante, se basa en el contraste de la información obtenida en los C.R.A. con la obtenida en las escuelas incompletas de tres unidades.

También los dos estudios evaluativos llevados a cabo sobre Centros Asociados de la U.N.E.D. adoptan una perspectiva básicamente normativa a la hora de valorar dichas instituciones. Concretamente, son los datos nacionales correspondientes al conjunto de Centros Asociados de la U.N.E.D., el patrón de comparación para el establecimiento de conclusiones. Junto a este enfoque, y como una consecuencia natural de dos estudios de corte cuantitativo-estadístico orientados desde la formulación de hipótesis iniciales, los criterios de significación estadística habituales constituyen el otro gran patrón de referencia a la hora de concluir e interpretar los resultados obtenidos.

En otros casos, no puede hablarse propiamente de formulación de juicios de valor. Así, el trabajo de Del Barrio del Campo constituye una investigación aplicada que no formula en sus conclusiones juicios de valor, y que, básicamente, interpreta los resultados conforme a los criterios habituales de significación estadística a partir de datos recabados de instrumentos de evaluación que no parecen haberse construido con referencia criterial.

Esta perspectiva fundamentalmente descriptiva caracteriza igualmente el trabajo de González Enríquez y Delgado Solís sobre recursos didácticos en centros escolares.

Finalmente cabe apuntar que, en el caso de los estudios evaluativos aún en curso, no se puede identificar aún la perspectiva adoptada en relación con este punto. Sólo cabe decir que, puesto que no se ofrece una definición operativa y sistemática de criterios previos de evaluación en los diseños iniciales, posiblemente la óptica a adoptar sea fundamentalmente normativa o comparativa, aún cuando también se puedan aplicar criterios para el contraste, asumidos de modo implícito, o generados a lo largo de la investigación.

— *Diseño y desarrollo de la evaluación*

El estudio de Morgenstern Pitcovsky y Martín Rodríguez se define como una primera aproximación al estudio de la evolución de los Centros de Profesores. En consecuencia, y considerando las razones

derivadas de la diversidad de situaciones contextuales y funcionales de los distintos Centros, así como la ausencia de estudios previos, la investigación se diseña a partir de una metodología que recogiera prioritariamente las tendencias generales con que han venido evolucionando los C.E.P. desde su creación.

Concretamente, si bien el modelo adoptado no es estrictamente descrito, el diseño evaluativo se define a partir de las siguientes características:

- Metodológicamente plural, en cuanto a la utilización de diversas técnicas de investigación, asumiendo complementariamente enfoques cuantitativos y cualitativos.
- Construido sobre el análisis de los problemas desde una perspectiva contrastada. Los procesos en este sentido se articularon en tres niveles distintos:
 - a. Información proporcionada por los sujetos a través de entrevistas y encuestas, y perspectiva aportada por la información documental disponible en los C.E.P. donde se hicieron los estudios de casos.
 - b. Las respuestas de los directores por una parte y de los asesores por otra.
 - c. El grado de congruencia y verosimilitud de las respuestas de un mismo sujeto.
- Preocupado por considerar en el mayor grado posible, en una evaluación externa como ésta, la participación de los propios sujetos implicados en las instituciones evaluadas, aunque preservando siempre el tratamiento anónimo y la confidencialidad de la información recogida -salvo excepciones expresamente autorizadas.

De los 106 Centros de Profesores existentes en el momento de diseño del estudio (abril 1989), y como ya se apuntó, se eligieron los 56 centros creados durante el primer año de implantación (1985). Por distintos motivos, dos centros fueron finalmente excluidos del estudio, por lo que la muestra total sobre la que se lleva a cabo el trabajo es de 54 centros, todos ellos pertenecientes al territorio gestionado por el Ministerio de Educación y Ciencia.

De las fuentes de información empleadas cabe decir que, si bien fueron variadas, el mayor volumen de información procede de fuentes

personales (directores y asesores de los C.E.P.). Como se deriva de lo anterior, el tipo de informaciones recogidas es de carácter indirecto y de naturaleza subjetiva (percepción y valoración de asesores y directores de los C.E.P.).

No obstante, en algunas ocasiones hay información sobre las instituciones o su contexto. Esta procede de fuentes documentales empleadas ocasionalmente o de la observación asistemática de los evaluadores.

La evaluación incluye también como fuente documental las memorias anuales de los C.E.P. existentes en la Subdirección General de Formación del Profesorado. Esta última fuente, sin embargo, es descartada a lo largo del estudio por no recoger de modo completo y sistemático información básica relativa a todos los C.E.P., y es empleada sólo en alguna ocasión muy concreta.

El desarrollo del proceso evaluativo fue el siguiente. Se diseñaron tres estudios de casos, correspondiendo a los C.E.P. de tipo A, B y C.¹⁵ Las características de la modalidad Centros-D se presuponen comunes a la de tipo C. Paralelamente se realizaron entrevistas en profundidad con personas que estuvieron o estaban aún vinculadas a la gestión de los C.E.P. y con asesores que tuvieran por lo menos dos años de antigüedad en el cargo. Finalmente, a partir de estas informaciones se construyeron los cuestionarios, base de las encuestas realizadas y procedimiento fundamental de recogida de información en el estudio. Concretamente se construyeron tres cuestionarios:

- Cuestionario de información general del C.E.P.
- Encuesta a los asesores de formación permanente de los C.E.P.
- Encuesta a los directores de los C.E.P.

Los cuestionarios incluían preguntas de respuesta cerrada y abierta. Éstas últimas fueron categorizadas y tabuladas a partir de la construcción de categorías previas.

Hay que considerar que, aunque las encuestas a los asesores y directores registraron una alta tasa de contestación, la encuesta general

¹⁵ Se tomaron los módulos adoptados por el Ministerio de Educación y Ciencia para la adjudicación de plantillas en función del profesorado adscrito. Módulo A: de más de 3.000 profesores; Módulo B: de 1.501 a 3.000 profesores; Módulo C: de 601 a 1.500; Módulo D: hasta 600 profesores.

del C.E.P. arrojó resultados muy precarios, por lo que se hizo un uso muy limitado de dicha información.

Con los datos obtenidos, y en la mayor parte de los casos, se optó por un análisis estadístico a través de tablas de contingencia bivariadas con porcentajes verticales, aunque para algunas variables se ofrecen únicamente los estadísticos principales.

Los resultados son interpretados desde una perspectiva múltiple que va desde la descripción, a la emisión efectiva de juicios de valor.

De acuerdo con unos objetivos netamente distintos a los planteados en el estudio anterior, los autores del proyecto de autoevaluación del C.E.P. de Getafe definen el modelo de evaluación adoptado como un elemento esencial de la planificación educativa, llevado a cabo por los propios agentes de la institución en un permanente proceso crítico-reflexivo de investigación-acción. El curso de las acciones diseñadas, por lo tanto, debía permitir la experimentación y la autoformación en su desarrollo, ser potencialmente transformador de la propia práctica, y participativo. En consecuencia, el enfoque metodológico adoptado es el ideográfico.

A pesar de no contar con información completa al tratarse de un estudio aún en curso, puede señalarse que en el desarrollo de la evaluación prevista parecen tener un lugar importante la dinámica de trabajo por equipos, y los procedimientos de cooperación e intercambio entre las distintas comisiones de trabajo que se formarán en el Equipo de Evaluación, incluyendo una Comisión Coordinadora. Por lo tanto, juegan un importante papel las técnicas de debate y de puesta en común para la toma de decisiones. Se prevé en este sentido la elaboración de documentos-síntesis de las distintas sesiones de trabajo, y de un diario que describa todo el proceso, así como la grabación en video del desarrollo de las sesiones para permitir la revisión y el análisis posteriores.

Las fuentes de información previstas son de naturaleza documental (Memoria del C.E.P., memorias de los asesores, etc.), personal, y comportamental. Como instrumentos para su recogida se citan: plantillas y registros para el análisis documental, observaciones sistemáticas (de carácter etnográfico), entrevistas semiestructuradas, cuestionarios e informes descriptivos para el análisis de la detección y diagnóstico de necesidades, de las actuaciones de formación, y de la organización y funcionamiento.

Como instrumentos para el análisis de la información recabada se prevén: cuadros de doble entrada, tablas de invención, parrillas y cues-

tionarios que faciliten una visión sintética de algunos datos, con el fin de establecer el análisis y discusión en pequeños grupos o en gran grupo.

En la evaluación de los Centros Rurales Agrupados, dirigido por Palomares Aguirre, no se explicita estrictamente el modelo evaluativo adoptado. Sin embargo, y por lo que se refiere a su diseño, éste podría calificarse de ecléctico, asumiendo algunos componentes de los modelos clásicos de evaluación que incluyen grupo de comparación, junto a un enfoque principalmente cualitativo en el empleo de fuentes de información y en los procedimientos para su recogida.

La muestra del estudio está constituida por un grupo de referencia o control del que forman parte todos los centros incompletos de menos de tres unidades de la provincia de Cuenca, y otro grupo, calificado de experimental, que lo forman los cinco Centros Rurales Agrupados seleccionados en la misma provincia. No obstante, debe tenerse presente que no se trata en ningún caso de un estudio experimental.

Se prevé el empleo de fuentes documentales, personales y comportamentales para recabar informaciones de carácter directo e indirecto. Así, como procedimientos para la recogida de información se citan:

- Visitas a los centros y reuniones con el personal (observación directa de instalaciones y recursos de los centros).
- Análisis de documentos e información administrativa (actas, horarios, plantillas, etc.).
- Cuestionario para profesores, a partir del que se pretende disponer de la opinión de los docentes de escuelas unitarias, de los C.R.A. y de centros incompletos de menos de tres unidades sobre los cinco factores o indicadores considerados nucleares.
- Entrevistas con los Directores y Claustros de centros.

Tanto para las entrevistas, como para las reuniones, la valoración de instalaciones y recursos y el análisis documental, se construyen guías que pautan los contenidos e informaciones a recabar. En la última fase de análisis y elaboración de conclusiones, se prevé la realización de esquemas y cuadros comparativos que aclaren los resultados obtenidos.

El proyecto formulado inicialmente para la evaluación de la Universidad de Salamanca, actualmente en curso y dirigida por Tejedor

Tejedor, explicita el modelo adoptado para llevar a cabo la evaluación. Concretamente éste se define como trifásico (de contexto, proceso, producto), y basado, por tanto, en la propuesta global ya clásica de Stufflebeam. El enfoque metodológico es fundamentalmente cuantitativo.

Como ya se apuntó, la población de análisis queda definida por el ámbito de la Universidad de Salamanca, en la totalidad de sus centros, departamentos, alumnos y profesores. Los investigadores igualmente prevén, en su caso, la definición de muestras representativas y suficientes, con estructuras de conglomerados y selección aleatoria para la obtención de informaciones relevantes para la consecución de los objetivos formulados.

Las técnicas de obtención de datos, y de acuerdo con la naturaleza de las variables que se consideran y con el tipo de información requerida, son notablemente variadas. De este modo se incluyen fuentes documentales (registros personales de Secretaría, partidas presupuestarias, registros del Instituto Nacional de Empleo, etc.), entrevistas personales a miembros de la comunidad universitaria (Consejo Social), encuestas y cuestionarios a alumnos y profesores, etc. Los tipos de información sobre los que se lleva a cabo la evaluación parecen ser, por tanto, directas e indirectas.

Concretamente, la documentación de avance aportada por los autores, informa del empleo de los siguientes tipos de instrumentos:

- Cuestionario sobre recursos del centro (espacios y servicios): Elaborado por el equipo investigador y aplicado a un responsable de cada centro universitario (Decano, Vicedecanos, etc.).
- Pruebas estandarizadas y/o adaptadas a los objetivos del estudio para la medida de las variables relativas a alumnos (pruebas comercializadas de aptitud intelectual, de personalidad, etc., ; cuestionario de hábitos de estudio, de satisfacción, de la participación en el ámbito universitario, elaboradas ad hoc por el equipo...).
- Cuestionario de evaluación del profesorado por los alumnos, autoinforme del profesorado y cuestionario de profesores, elaborados por los investigadores.

Además, se emplea cuando es pertinente, la información disponible del Centro de Proceso de Datos de la Universidad de Salamanca.

Las técnicas de análisis, dado el carácter del estudio, se prevén de naturaleza exploratoria y descriptiva, sin renunciar al posible nivel explicativo e intentando establecer entre las variables relaciones de tipo causal. Se emplearán, de acuerdo con los autores, técnicas descriptivas, inferenciales (para comparar distintos datos en centros y departamentos), y con especial atención a los modelos de regresión y modelos logarítmico-lineales, dada la naturaleza categórica de alguna de las variables consideradas. La documentación disponible hasta el momento incluye básicamente análisis de tipo descriptivo (análisis de frecuencias, estadísticos de posición y dispersión y estudios correlacionales).

Asimismo, en los dos estudios evaluativos incluidos en el trabajo sobre la Educación Superior a distancia, dirigido por Marín Ibañez, se adopta un enfoque cuantitativo-estadístico, en esta ocasión guiado, como se dijo, por la formulación inicial de hipótesis generales y parciales. Además cabe apuntar cómo ambas evaluaciones adoptan una perspectiva longitudinal en el contraste de algunas de las hipótesis formuladas, (relativas al Centro Asociado de Baleares), o en algunos de los análisis (relativos al Centro Asociado de Mérida).

En la evaluación institucional del Centro Asociado de Baleares, las muestras empleadas para el contraste de las hipótesis formuladas varían de acuerdo con la naturaleza específica de las mismas, y con los procedimientos empleados para la recogida de la información. En unos casos se trabaja con el universo total de alumnos matriculados en el curso 1984/85 (N=1152) y, en otros, con las muestras incidentales extraídas del mismo de acuerdo con la disponibilidad de los datos. Para los análisis longitudinales se considera como universo muestral a los alumnos matriculados en el Centro Asociado desde el curso 1977/78.

El tipo de información recabada es de naturaleza tanto directa (variables sociodemográficas, rendimiento, tasas de abandono) como indirecta (valoración y sugerencias relativas a materiales, a procedimientos de evaluación empleados, y datos referidos a otros aspectos que se consideran asociados al rendimiento de los alumnos).

Los fuentes de información empleadas son en consecuencia diversas, e incluyen fuentes tanto documentales como personales. Concretamente, como procedimientos de recogida de datos se emplean: la ficha de datos sociológicos incluida en el impreso de matrícula del curso 1984/85, registros de la Secretaria del Centro Asociado (período 1977-1985), y cuestionarios diseñados *ad hoc* para alumnos y profesores-tutores.

En la evaluación del Centro Asociado de Mérida, el diseño y desarrollo de la evaluación es en muchos aspectos similar. Se emplean igualmente fuentes de información documental (registros administrativos, tanto de la Secretaría del Centro Asociado de Mérida como de la de los Centros Asociados de la Sede Central; la memoria económica y de actividades, etc.); así como datos personales (alumnos licenciados y alumnos matriculados en el momento de la investigación). También en este caso se emplea como procedimiento de recogida de información el cuestionario.

Los universos muestrales de referencia igualmente estuvieron constituidos por licenciados en cursos anteriores, y por alumnos matriculados en el momento de la recogida de la información.

La medición de las variables se realizó a través de un cuestionario remitido en el curso 1982-83 a los licenciados por la U.N.E.D. en el centro de Mérida¹⁶; y un cuestionario enviado a los alumnos de matriculados en el Centro Asociado de Mérida en el curso 1984-85. Las muestras finales estuvieron constituidas por 84 licenciados y 644 alumnos.

Los resultados de los análisis se interpretan en términos de descripción, explicación y valoración, siempre teniendo como telón de fondo el contraste estadístico de las hipótesis y la comparación con los datos nacionales de la U.N.E.D. (del conjunto de Centros Asociados).

Por su parte, la investigación llevada a cabo sobre la utilización del equipamiento didáctico en centros públicos de enseñanza, constituye un estudio de los denominados "tipo survey", e incluye en su diseño dos fases, con una naturaleza distinta en algunos aspectos, pero con características básicas comunes.

De este modo, y en ambas fases, se recoge información indirecta acerca de las dimensiones de análisis anteriormente apuntadas, se emplean fuentes de información personales, y se utiliza como procedimiento de recogida de información el cuestionario aplicado mediante entrevista personal.

La primera fase del estudio operó sobre una muestra de 152 centros educativos públicos, de E.G.B., B.U.P. y Formación Profesional, seleccionados aleatoriamente de las Comunidades Autónomas si-

¹⁶ Aplicado con motivo de una investigación anterior de uno de los componentes del equipo actual: GARCIA AREITO, J. (1985): *Licenciados extremeños de la UNED*. UNED-Centro Regional de Extremadura, Badajoz.

guientes: Asturias, Aragón, Baleares, Cantabria, Castilla-León, Extremadura, Rioja, Murcia y Navarra.

En cada uno de estos centros se entrevistó, a partir de un cuestionario estructurado, al Director y a dos profesores. En total se efectuaron 457 entrevistas: 152 a directores y 305 a profesores. En todos los casos se entrevistó a profesores que impartían docencia en 8º de E.G.B., 1.º de B.U.P. o 1.º de F.P., y pertenecientes a algunas de las siguientes cuatro áreas: Ciencias Sociales, Ciencias, Idiomas y Gimnasia.

En la segunda fase, se adoptó un enfoque más cualitativo y abierto en la recogida de datos. Con un cuestionario-guión de entrevista, apenas formalizado, se entrevistó a tres profesores de cada centro. La muestra final estuvo compuesta por 192 profesores pertenecientes a 96 centros de Castilla-La Mancha y de Madrid.

Los datos recogidos en las entrevistas mantenidas en la primera fase fueron tratados estadísticamente. Los de la segunda fase fueron objeto de tratamiento cualitativo mediante examen de todas las respuestas recogidas y extracción de conclusiones por áreas temáticas.

Finalmente, el estudio dirigido por Del Barrio Del Campo se enmarca en la ya clásica línea de investigación sobre ambientes institucionales y educativos. Concretamente se alinea con los enfoques que emplean como método de evaluación del clima medidas perceptivas de los distintos sectores implicados, tomando como unidad de análisis el centro educativo.

Como consecuencia del modelo adoptado para la evaluación del clima institucional, las informaciones o medidas tomadas son en todo caso indirectas, siendo recabadas de alumnos, padres y profesores. Se adopta, además, una perspectiva cuantitativa tanto en los procedimientos de recogida de los datos como en su análisis y tratamiento.

Los instrumentos empleados en esta investigación, constituidos por tres cuestionarios, tratan de evaluar el ambiente escolar percibido por los alumnos (EVECE-alumnos), los profesores (CECE-profesores) y los padres (CECE-padres)¹⁷.

¹⁷ De acuerdo con los autores, el antecedente más próximo de tales instrumentos es el cuestionario de evaluación de ambientes escolares de V. Pelechano, a partir de la University Residence Environment Scale (URES) de Moos y Gert (1974) y de la Ward Atmosphere Scale (WAS), de Moos (1974). La escala primitiva, denominada CEAE, trató de adecuar el instrumento a la realidad de nuestro sistema educativo y a los problemas y situaciones planteados en la E.G.B. A partir de ella, se elabora en primer lugar la prueba para profesores, de la que surgen las otras dos para alumnos y padres.

Las muestras empleadas en el estudio estuvieron constituidas por 80 profesores, 277 padres y 1383 alumnos del Ciclo Superior de E.G.B. de quince centros escolares de Cantabria¹⁸. Nada se apunta acerca del procedimiento de muestreo, aparte de que se hayan establecido ciertos "controles en la extracción muestral", por lo que cabe imaginar que éste fue incidental.

La interpretación que se realiza de los resultados obtenidos de los análisis estadísticos adopta un enfoque esencialmente descriptivo. A partir de la misma, se formulan las conclusiones correspondientes en términos de pautas a considerar en posibles procesos de intervención posteriores. No se emiten, por lo tanto, juicios de valor explícitos.

— *Difusión y uso de los resultados de la evaluación*

En general, para la mayor parte de estudios revisados en este apartado, poco se puede apuntar en relación con los procedimientos de devolución de resultados a las audiencias interesadas y con su papel en la toma de decisiones, si procedía. La mayor parte de las memorias de investigación no informan de los procedimientos previstos o empleados para la difusión de los resultados obtenidos, ni para la devolución de datos a los propios centros implicados en el estudio. En la mayor parte de los proyectos iniciales e informes de progreso analizados en el caso de las investigaciones en curso, tampoco se prevé nada en este sentido.

Sin embargo, cabría mencionar dos excepciones. En el trabajo de Morgenstern Pitcovsky y Martín Rodríguez sobre la red de Centros de Profesores, y a juicio de los autores, la modalidad y alcance del estudio no permitieron la participación plena de los sujetos evaluados, aunque se intentó en la medida de lo posible contar con los distintos responsables de los C.E.P. a la hora de concebir y obtener la información necesaria para llevar a cabo la evaluación. De este modo, los investigadores ponen de manifiesto su compromiso de devolver la información a los participantes para que procedan a su validación. Esta labor se considera necesaria para la elaboración definitiva del informe

¹⁸ Debe destacarse la alta indefinición de las muestras por lo que se refiere a los centros de donde fueron extraídas. Sólo en el caso de los alumnos se informa de que proceden de 15 centros y que son de ciclo superior de EGB. En las muestras de profesores y padres, aunque se caracterizan a nivel de sujetos (titularidad del centro, niveles que imparte, etc.), no se puede determinar el número de centros empleados, que en este caso parecen ser de EGB y BUP. Se desconoce por tanto si se trata en todos los casos (para las tres muestras) de los mismos centros.

e, incluso, los autores defienden que no puede quedar definitivamente cerrado el estudio hasta que no se cumpla este requisito.

Por otro lado, en la investigación dirigida por Tejedor Tejedor sobre la Universidad de Salamanca, se prevé la elaboración de un Informe de evaluación global, así como de informes particulares para cada centro.

Por lo que se refiere al uso de los resultados obtenidos en las evaluaciones ya finalizadas, de nuevo debe apuntarse cómo sólo pueden recogerse aquí las áreas de utilidad identificadas por los autores, en la medida en que la práctica totalidad de los evaluadores no poseen competencias reales de intervención sobre las instituciones y centros evaluados.

En este sentido, la memoria de investigación presentada tras la evaluación de la red de Centros de Profesores, incluye un conjunto de sugerencias y conclusiones destinadas a la Administración educativa, responsable de la misma. En el trabajo de Del Barrio Del Campo sobre criterios ecológicos sobre la percepción de centros escolares, se señala lo siguiente: *“Los resultados alcanzados en esta investigación constituyen un conjunto de guías útiles a la hora de diseñar y poner en marcha programas de intervención que afecten a los centros escolares, tanto globalmente, como a algunos de sus estamentos componentes”* (p. 434). Finalmente, en el trabajo dirigido por Marín Ibañez sobre la Educación Superior a distancia, nada se apunta en este sentido.

— *Metaevaluación*

Aunque no pueda hablarse de una metaevaluación formal en ninguno de los estudios (entendida como actividad intencional, planificada y sistemática), en las dos evaluaciones llevadas a cabo sobre Centros de Profesores (C.E.P.) se pueden identificar elementos cercanos a esta actividad.

El trabajo sobre la red de Centros de Profesores, incluye en el capítulo de conclusiones un apartado específico donde los autores ponen de manifiesto las limitaciones de la evaluación llevada a cabo, sus aspectos mejorables y aún sus errores. Por su parte, en el proyecto de evaluación del C.E.P. de Getafe, se prevén procedimientos que permitan una “metaevaluación” posterior. Concretamente, se procederá a grabar en video las sesiones de trabajo, como registro que permitirá la

valoración posterior del desarrollo de la autoevaluación institucional, y que se entiende contribuirá a su enriquecimiento.

4.4. Investigaciones centradas en la elaboración de modelos para la evaluación de programas o centros educativos

Ha parecido interesante dedicar un espacio propio a este grupo de investigaciones por la aportación adicional que suponen a este campo de estudio. Así, aunque estos seis estudios se centran en la evaluación de programas de formación de profesorado y en la evaluación institucional, en vez de incorporarlos a sus correspondientes grupos en función de su objeto de estudio, se ha preferido destacar su carácter instrumental, su valor como modelos de evaluación.

4.4.1. Modelos de evaluación de Programas de Formación del Profesorado

En este grupo se incluyen las siguientes investigaciones:

- *Análisis de los resultados de los programas de formación permanente del profesorado del Ministerio de Educación y Ciencia mediante la evaluación demorada de sus efectos. Determinación de criterios y elaboración de instrumentos.* (S. Antúnez Marcos, en curso).
 - *Marco conceptual para el diseño de sistemas integrados de evaluación de programas de desarrollo profesional del docente en centros educativos.* (L.M. Villar Angulo, 1992).
 - *Virtualidad de un modelo de evaluación para cursos de Actualización Científico-Didáctica de los C.E.P.* (M.^a J. Sáez Brezmes, en curso).
- *Objeto de la evaluación*

La investigación dirigida por Antúnez trata de enfrentar de forma novedosa la evaluación de programas de formación permanente promovidos por la Subdirección General de Formación del Profesorado. Como se afirma en el propio proyecto de investigación, hasta ahora

“Los responsables de cada actividad han diseñado frecuentemente instrumentos para la evaluación de la formación. Su finalidad ha consistido, en general, en recoger información para evaluar el proceso y, la mayoría de las veces, para verificar la satisfacción de las personas participantes... A menudo se han utilizado los cuestionarios y, esporádicamente, las técnicas de grupo y la observación... Pocas veces se ha puesto el énfasis en comprobar los resultados efectivos de la formación a partir de las aplicaciones que, de los aprendizajes que la formación ha promovido, pudieran hacer las personas receptoras en su lugar de trabajo...” (pág. 3). Esto último es precisamente el objeto de esta evaluación: los efectos que ha tenido la formación recibida sobre el propio trabajo docente. En concreto, los efectos de tres cursos de distintas modalidades de formación: para responsables de formación de personas, para asesores de formación tipo “A”, y para asesores de formación tipo “B”.

Por su parte, la investigación dirigida por Villar Angulo se propone diseñar sistemas integrados de evaluación de programas de desarrollo profesional del docente en centros educativos. En concreto, tiene por objeto la creación y aplicación de una forma de evaluación de programas de perfeccionamiento docente desde un marco conceptual con unos determinados principios teóricos, entre los que se pueden mencionar los siguientes: a) la escuela es la unidad básica del cambio; b) las instituciones políticas tienen la primera responsabilidad en el perfeccionamiento profesional de los docentes; c) los directores de los centros impulsan o frenan iniciativas de innovación, promoción de seminarios permanentes y actividades de autoformación en los centros; d) los programas de perfeccionamiento combinan la investigación con la práctica; e) la evaluación del perfeccionamiento requiere tiempo y una cuidadosa planificación que incluya la determinación de un sistema de indicadores.

El trabajo incluye el estudio y contrastación de modelos, métodos e instrumentos de evaluación de los 15 programas del Plan de Formación Permanente del Ministerio de Educación y Ciencia y de los 12 programas del Plan Anual de la Comunidad Autónoma de Andalucía.

La investigación dirigida por Sáez Brezmes se centra en la mejora de la utilidad y credibilidad de los procesos de autoevaluación que se llevan a cabo de las actividades de formación permanente del profesorado que, de acuerdo a la normativa, ha de realizar el Equipo Pedagógico de los Centros de Profesores (responsable de evaluar el Plan

Anual de Formación, y de realizar la Memoria final de cada curso). En concreto, el trabajo consiste en la contrastación de un modelo de autoevaluación diseñado para facilitar a los coordinadores de los cursos y a los asesores de los C.E.P. evaluar sus propios cursos de formación.

— *Características definitorias de la evaluación*

Agentes evaluadores

La investigación dirigida por Antúnez tiene carácter externo, puesto que es llevada a cabo por profesores universitarios que no participaron directamente en los cursos cuya efectividad se trata de evaluar.

La investigación dirigida por Villar Angulo, cuyo propósito fundamental es elaborar un sistema integrado de evaluación de la formación, incluye también la evaluación externa de un programa de formación permanente desarrollado en Andalucía. Además, los autores seleccionaron una red de asesores de C.E.P., de miembros de las Comisiones Técnicas Provinciales y de profesores no universitarios, para que colaboraran en la evaluación externa del "Sistema Integrado de Evaluación" elaborado durante la investigación.

La evaluación del "modelo de autoevaluación" de cursos de formación elaborado y experimentado por el equipo de Sáez Brezmes sigue dos modalidades: la que los propios autores realizan de su modelo (evaluación interna) y la que realizan del mismo los miembros del CEP en que dicho modelo es experimentado.

Objetivos y funciones de la evaluación

La evaluación de cursos de formación permanente del profesorado dirigida por Antúnez, persigue los siguientes objetivos:

- Identificar las variables más significativas para determinar la eficacia de la formación mediante el análisis de la práctica posterior de los sujetos en su lugar de trabajo.
- Identificar los criterios para diseñar los instrumentos de evaluación más precisos para detectar la validez de la formación en función de sus efectos a medio plazo.
- Diseñar y experimentar estrategias e instrumentos de evaluación demorada para que sean aplicados en los planes y progra-

mas de la Subdirección General de Formación del Profesorado. De hecho, se proponen diseñar un Plan Experimental de evaluación general demorada de la formación permanente del profesorado, así como directrices para su aplicación.

En relación a esto, se puede afirmar que la función que persigue la evaluación es formativa, ya que los propios autores reconocen que, a partir de ella, la propia Subdirección General de Formación del Profesorado puede perfeccionar el diseño y desarrollo de sus planes de formación permanente, así como los propios instrumentos que se vienen utilizando para su evaluación, desde una perspectiva innovadora: la evaluación demorada de la eficacia de los cursos que se llevan a cabo.

Por su parte, los objetivos de la evaluación que dirige Villar Angulo, son los siguientes:

- Sugerir marcos conceptuales para el diseño de sistemas integrados de evaluación (S.I.E.) de los Planes Anuales de Formación Permanente del Profesorado, con el fin de mejorar la calidad de estos programas de formación permanente.
- Describir los factores clave de los programas de formación permanente de los C.E.P.
- Ofrecer datos a las Comisiones Técnicas Provinciales (CTP) y a los C.E.P. sobre variables eficaces de entrenamiento y sobre materiales que se deben incluir en las actividades de formación. En definitiva, proporcionarles información sobre la aplicación del S.I.E. a un programa concreto en Andalucía.
- Proponer orientaciones para la evaluación de actividades de formación.
- Construir y validar un sistema de indicadores para la evaluación de actividades de los programas de formación permanente.

A partir de los objetivos, y del desarrollo y conclusiones del estudio, se puede afirmar que la función, tanto de la evaluación del programa concreto de formación como del Sistema Integrado de Evaluación elaborado por los autores, tienen como objetivo prioritario orientar la toma de decisiones de mejora de dos elementos centrales: los programas de formación y los modelos utilizados para evaluarla.

En definitiva, su función es conocer el grado en que las iniciativas oficiales (andaluzas y de la Administración central) de formación de

profesores, contribuyen a mejorar la calidad de la enseñanza, así como optimizar los procesos de gestión del perfeccionamiento y de los centros educativos.

Los objetivos de la investigación de Sáez Brezmes y colaboradores son los siguientes:

- Experimentar el modelo de autoevaluación elaborado en los contextos para los que fue diseñado, es decir, cursos de Actualización Científico-Didáctica que serán evaluados por dos asesores del C.E.P.
- Analizar y valorar los aspectos más relevantes del modelo:
 - Si la estructura del modelo propicia la realización de evaluaciones técnicamente correctas.
 - La utilidad del modelo como guía para recoger y suministrar la información necesaria sobre los procesos de enseñanza aprendizaje, cumpliendo así su carácter formativo e incidiendo en los procesos de mejora.
 - La capacidad para identificar problemas.
 - Los efectos sobre la mejora de la comunicación entre los profesores y asistentes a los cursos.
 - La función formadora que se le supone al modelo para aquellos que han de evaluar los cursos.
- Revisar y perfeccionar el modelo.
- Evaluar la utilidad del modelo definitivo en distintos tipos de cursos y con distintos tipos de asesores.

A la vista de estos objetivos parece claro que la finalidad de esta evaluación del modelo es formativa, ya que uno de sus pasos es precisamente la revisión, ajuste y mejora del mismo tras su aplicación en contextos reales.

Dimensiones evaluadas y momento de la evaluación

En la evaluación de cursos de formación permanente del profesorado que dirige Serafín Antúnez, y como los propios autores explicitan, la evaluación se centra en la eficacia, es decir, en los resultados de la formación del profesorado (atendiendo a sus efectos en tres tipos de

práctica profesional: docente, formadora o directiva), considerando que esto ha de ayudar a perfeccionar los diseños de los planes de formación permanente, así como a verificar hasta qué punto fueron útiles los programas o planes que se desarrollaron. Por otra parte, al tratarse de una evaluación de producto es también final, pues, se realizará tras la finalización de los programas de formación.

Las dimensiones que incluye el diseño evaluativo dirigido por Villar Angulo son las clásicas en toda evaluación global: de contexto, entrada, proceso y producto. De hecho, el núcleo de la investigación consistió en la elaboración y contrastación de un modelo de evaluación de la formación del profesorado en activo (el S.I.E.), transcurriendo la evaluación del mismo en paralelo a su propia elaboración.

Por presentar la dimensionalidad del modelo de acuerdo a como ésta se describe en el estudio, habría que decir que en esta investigación se propone un modelo de evaluación de los programas de formación que parte de la interrelación de tres conceptos: evaluación, desarrollo profesional y gestión educativa de centros. El primero se refiere a la intención o función de la evaluación; el segundo a aquello que se debe observar/evaluar; y el tercero a los resultados/productos esperados. Cada concepto lleva asociadas seis categorías que pueden ser objeto de diversas combinaciones susceptibles de ser formalizadas en un diseño concreto de evaluación integrada.

En cuanto a la investigación dirigida por Sáez Brezmes, habría que distinguir las dimensiones del modelo de autoevaluación de los cursos que han elaborado, y la dimensionalidad que tiene la evaluación que los investigadores realizan de dicho modelo.

En cuanto al primer caso, tal y como el propio equipo explicita, el modelo de autoevaluación de la formación del profesorado es un modelo global que atiende a cuatro dimensiones:

El contexto: que se propone caracterizar el curso en el medio en que se realiza, identificar y examinar los objetivos y prioridades de éste, y estudiar la población concreta a la que se dirige, con el fin de ajustar los objetivos y prioridades del curso a esa población.

El input: estudio de la organización y enfoque del curso, así como de los materiales curriculares que se ponen a disposición de los asistentes, siendo valioso conocer las estrategias alternativas que distintos profesores y asistentes al curso harían una vez planteado el curso globalmente.

El proceso: se trata de analizar el impacto de la aplicación de la estrategia de enseñanza-aprendizaje diseñada, del uso eficaz de los recursos y de la comunicación que se establece entre profesores y alumnos. Los informes parciales de evaluación se proponen producir un *feed-back* que mejore algunos aspectos en esta fase.

El output: dirigido a analizar e interpretar los logros del curso, no sólo en cuanto a lo que han aprendido los asistentes, sino también los efectos demorados de la formación: hábitos adquiridos, modificaciones que los profesores aplican en las aulas tras la formación, etc. En definitiva, el impacto del curso en los asistentes y en su contexto.

En cuanto a la dimensionalidad de la evaluación del modelo que se somete a prueba, aunque no se explicita con el mismo detalle en el proyecto, se infiere que el diseño cualitativo de Estudio de Caso que se aplica también recoge información de las cuatro dimensiones: entrada, contexto, proceso e impacto.

Patrón empleado para la emisión de juicios de valor

Del proyecto de investigación diseñado por el equipo que dirige Antúñez, se deduce que una de sus tareas es precisamente la determinación de los criterios para valorar la eficacia de los cursos de formación objeto de estudio. Por consiguiente, puede afirmarse que tratan de llevar a cabo una evaluación referida a unos criterios predeterminados (evaluación criterial).

La evaluación de programas de formación dirigida por Villar Angulo también puede considerarse criterial, al utilizar como instancia de referencia los objetivos de los propios programas oficiales de formación permanente de profesores; de hecho, al hacer una evaluación de los resultados de los programas por sus efectos en la propia labor docente del profesorado, los criterios que se utilizaron fueron los objetivos de los propios cursos en los que éstos habían tomado parte. Pero al mismo tiempo esta evaluación es de tipo normativo, puesto que compara los resultados de unos cursos con los obtenidos por cursos de otros Centros de Formación de Profesores (el grupo normativo de referencia en este caso).

En la evaluación del "modelo de autoevaluación" de cursos que realizan Sáez Brezmes y su equipo se emplea una instancia de referencia criterial, puesto que la información recogida se va a interpretar en

relación a los objetivos que persigue el propio modelo. De hecho, se aleja explícitamente de la instancia normativa, puesto que se excluye la comparación del éxito de la aplicación del modelo entre unos cursos de formación y otros.

— *Diseño y desarrollo de la evaluación*

El análisis de los resultados de los programas de formación permanente del profesorado del Ministerio de Educación y Ciencia mediante la evaluación demorada de sus efectos, dirigido por Antúnez, tiene como propósito fundamental llenar el vacío existente en nuestro país de este tipo de estudios. Los autores pretenden efectivamente construir un Plan o Modelo para realizar este tipo de evaluación. Para ello, les sirve de base un modelo de evaluación ya construido en cuya elaboración participó parte del actual equipo de investigación. A partir de ahí se proponen diseñar y experimentar estrategias e instrumentos específicos. Las fases de la investigación son las siguientes:

- 1.^a) Desarrollo de un trabajo piloto que consistirá en la elaboración y experimentación de un conjunto de instrumentos de evaluación, que se aplicarán a una muestra inicial de personas asistentes a tres cursos de formación auspiciados por la Subdirección General de Formación del Profesorado:
 - Curso de Responsables de Formación de Personas Adultas (desarrollado en 1991).
 - Curso de Asesores de Formación "A" (curso 1990-91).
 - Curso de Asesores de Formación "B" (curso 1991-92).

Asimismo, se realizarán entrevistas a los directores de los Centros de Profesores de la zona, y a profesores y profesoras participantes en los cursos. A partir de ahí se evaluarán los resultados.

- 2.^a) Establecer los criterios, fuentes y pautas hipotéticamente adecuados para evaluar la eficacia demorada de la formación, elaborar un conjunto de instrumentos a partir de los mismos, y diseñar un *Plan Experimental de Evaluación General Demorada* de este tipo de formación.

Por lo demás, en el proyecto de investigación no se especifica el tipo concreto de diseño que se empleará para experimentar el modelo de evaluación, aunque sí se anticipa que será una investigación de tipo cualitativo. En este sentido cabe destacar los siguientes elementos del diseño:

- Los procedimientos para recoger la información serán: entrevistas semiestructuradas, listas de control, escalas, cuestionarios, etc., elaborados al efecto.
- Las fuentes de información: personas individuales (profesores y directores de C.E.P.); colectivos o equipos de responsables de formación en C.E.P., Unidades de Programas, Subdirección General de Formación del Profesorado, etc.; huellas de trabajo.
- El tipo de información a recoger es documental (normativa, Memorias de los C.E.P., programas de formación, etc.) y personal o de opinión (de responsables y destinatarios de la formación).
- El análisis de la información será cualitativo, aunque no se especifica cómo, y se hará sobre la práctica docente de los profesores en su lugar de trabajo tras recibir la formación.

La evaluación dirigida por Villar Angulo se diseña en función de una matriz de trabajo articulada por tres ideas-eje: la naturaleza y funciones de la evaluación, el desarrollo profesional y la gestión educativa en centros. El sistema de evaluación que pretenden diseñar es una actividad de comprobación de la teoría y funciona como un diseño de evaluación mezclado. En este sentido los autores se apoyan en el marco teórico ofrecido por otros autores (Greene, Caracelli y Graham, 1989).

El estudio tiene tres fases:

- 1.^a) Estrategia de planificación del sistema:
 - a. Evaluación de necesidades y prioridades formativas.
 - b. Análisis de factibilidad del sistema de evaluación integrado.
 - c. Planificación del sistema de evaluación.
- 2.^a) Estrategias de desarrollo del producto:
 - a. Diseño y tecnología del sistema de evaluación.

- b. Comprobación en el campo.
- c. Comprobación operativa.

3.ª) Estrategias de difusión.

Con respecto a otros elementos del diseño se puede observar lo siguiente:

- Empleo de la triangulación de fuentes de información (formadores, profesores, alumnos, miembros de Comisiones Técnicas Provinciales), así como de métodos, técnicas e instrumentos de evaluación (entrevistas, sondeos de opinión, observaciones, cuestionarios, etc.).
- Se recogen distintos tipos de información: documental, de opinión y algunos hechos observados. Especialmente se recoge información sobre cómo evalúan los Centros de Profesores sus actividades. Con la siguiente matriz de preguntas se estudian y presentan las evaluaciones realizadas por los C.E.P.;
 - a. ¿Qué preguntas se han formulado los investigadores/ formadores para hacer la evaluación del plan, programa, curso, seminario, actividad o proyecto de innovación?
 - b. ¿Qué información han requerido para evaluarlos?
 - c. ¿Qué fuentes de datos se eligieron?
 - d. ¿Cuáles fueron los métodos para recoger información?
 - e. ¿Quién, cuándo y en qué condiciones se recogió la información?
 - f. ¿Qué procedimientos se han usado para el análisis de datos?
 - g. ¿Qué criterios se han seguido para hacer las interpretaciones?
- Son diversos también los procedimientos utilizados para recoger la información: análisis de documentos (para la evaluación interna de las C.T.P., y de los programas de formación), entrevistas a formadores y profesores participantes, sondeos de opinión, observaciones de sesiones formativas, cuestionarios a profesores, etc.
- En el análisis de la información se han empleado fundamentalmente técnicas cualitativas. Esto lo justifican los autores afirmando que han tratado de evitar las limitaciones de la investi-

gación cuantitativa que, por la dificultad de disponer de muestras aleatorias del profesorado, casi siempre se basa en diseños cuasi-experimentales, con los consiguientes problemas de validez.

Por otra parte, la investigación dirigida por Sáez Brezmes, por las funciones y objetivos que plantea, se enfoca desde una metodología cualitativa y naturalista. Así, el estudio se centra en el elemento más significativo: la capacidad del modelo de autoevaluación para producir los efectos deseados o, dicho de otro modo, su virtualidad.

El modelo de autoevaluación se experimenta en dos cursos de actualización científico-didáctica, uno de tipo B y otro de tipo C. Cada uno de los cursos en los que se aplica el modelo se consideran como unidades de estudio con entidad en sí misma, por lo que se tratan metodológicamente como auténticos Estudios de Caso. A partir de las conclusiones obtenidas en dichos estudios comienza una segunda fase en la que se realiza una metaevaluación (análisis y valoración de la fase de experimentación) que dará como consecuencia la revisión y perfeccionamiento del modelo. Una tercera y última fase consiste en la aplicación del modelo definitivo en diversos cursos de formación.

Como parece exigir este diseño, las fuentes de información son sobre todo directas, recabando la opinión de asistentes, coordinadores y asesores participantes en el proceso de autoevaluación. De hecho, el impacto del modelo se analiza recogiendo opiniones de una pluralidad de fuentes: evaluadores, coordinadores, asistentes, asesores del C.E.P. y de todas las personas implicadas en la fase de experimentación del modelo-guía. El tipo de información que se recoge es documental y personal (opiniones/actitudes).

Se pretende huir de la utilización de instrumentos estandarizados de recogida de la información (cuestionarios, etc.) aplicados en todos los cursos donde se experimenta el modelo. Se opta más bien por emplear aquellas técnicas e instrumentos que el Estudio de Caso vaya demandando, siempre de carácter personalizado y contextualizadas. Así, los procedimientos utilizados para recoger la información son diversos: observación directa e indirecta y desde diferentes puntos de vista, realizar cuadernos de campo, entrevistas, etc.

— *Difusión y uso de los resultados de la evaluación*

En el proyecto dirigido por Antúnez no se especifica el uso que se hará de las conclusiones de la evaluación, y aunque se halla todavía en

curso, se identifica con claridad la audiencia a la que pretenden dirigirlas: la propia Subdirección General de Formación del Profesorado.

La audiencia a la que se dirigen las conclusiones de la investigación de Villar Angulo y colaboradores es similar al caso anterior: los organismos o centros responsables de la formación. De hecho, el propio equipo establece como uno de los principales objetivos de la evaluación, "ofrecer datos a las Comisiones Técnicas Provinciales y a los Centros de Profesores de la aplicación del Sistema Integrado de Evaluación a un programa de formación permanente en la Comunidad Autónoma de Andalucía". Asimismo, el sistema de indicadores que se construye y se valida en el estudio, se pone a disposición de los usuarios de los programas de formación permanente (formadores, profesores y alumnos), para que puedan evaluar sus propias actividades.

Por su parte, el uso que el equipo de Sáez Brezmes pretende hacer de los resultados de la evaluación del modelo que llevan a cabo es el propio perfeccionamiento del modelo. Por otra parte, dado que el modelo de autoevaluación que se experimenta fue un encargo de un Centro de Profesores concreto, parece claro que la difusión que se hará del modelo resultante tiene como clientela prioritaria a éste y otros centros de profesores.

Asimismo, se cuida especialmente la presentación de resultados. Se prevé la presentación de informes parciales (para dar feedback en el proceso de aplicación del modelo-guía) y de un Informe final, persiguiéndose especialmente la credibilidad interna y externa de la autoevaluación.

— *Metaevaluación*

En la investigación proyectada por Antúnez y colaboradores no se lleva a cabo una metaevaluación. Sin embargo, en la investigación dirigida por Villar Angulo, se selecciona y establece una red de asesores de C.E.P., de miembros de las Comisiones Técnicas Provinciales y de profesores no universitarios para colaborar en la evaluación externa del "Sistema Integrado de Evaluación" elaborado y aplicado en la Comunidad andaluza, lo cual podría considerarse una forma de metaevaluación.

En el caso de la investigación dirigida por Sáez Brezmes, como ya se ha reflejado anteriormente, la metaevaluación forma parte del pro-

pio proyecto de investigación. En realidad, éste consiste en una evaluación del modelo de autoevaluación diseñado por los autores.

4.4.2. Modelos de Evaluación de Centros

En este epígrafe se agrupan investigaciones dirigidas a la elaboración y/o experimentación de modelos de evaluación de centros educativos de distintas características (ordinarios, de educación especial, públicos y privados). Este es el caso de las siguientes investigaciones:

- *Modelo de estudio de costes de la educación (M.E.C.E.)* (M. Ibar Albiñana, 1987).
- *Evaluación del C.E.P. de Getafe-Parla a través del estudio de sus ofertas y de las demandas de los profesores y centros educativos de su competencia.* (E. Sánchez Martín, 1992).
- *Contexto y retraso mental: sistema de evaluación e centros para sujetos con retraso mental* (M.J. Espinosa, 1990).

A continuación se describen algunos de los rasgos más definitorios de cada una de ellas.

— *Objeto de la evaluación*

El estudio realizado por el equipo que dirige Ibar Albiñana es definido como una “investigación básica, destinada a la construcción de un modelo matemático-académico-administrativo que sirva para la evaluación económica y organizativa de los centros docentes de nivel no universitario, y con vistas a una aplicación continua para la obtención y publicación de series cronológicas de tendencias medias e índices de gestión”.

Este modelo administrativo (que deja para otros modelos complementarios la valoración de los aspectos cualitativos de la educación que se imparte en un centro), se compone de seis submodelos en los que se analizan y cuantifican los recursos humanos y materiales como factores educativos o de producción del centro, para pasar al estudio de su repercusión en las plantillas, nóminas y resultados contables o de coste.

El modelo se experimentó en centros privados de la Comunidad de Madrid y la muestra estudiada estuvo compuesta por los centros

(22) que voluntariamente respondieron a una demanda de colaboración ofrecida a la totalidad del colectivo de 342 centros de la F.E.R.E. en la provincia de Madrid. El tamaño de los centros participantes oscilaba entre 400 y 1200 alumnos y tenían una oferta escolar diversa (Preescolar, E.G.B., B.U.P., C.O.U. y/o Formación Profesional).

El modelo se adecuaba a los centros privados fundamentalmente, que eran los que en la fecha en que se realizó el estudio podían asignar con autonomía los recursos económicos disponibles a una u otra partida del presupuesto. Sin embargo, los autores señalan también la utilidad del modelo para los centros concertados, pues se puede determinar la cuantía del módulo de las subvenciones de los conciertos (de hecho es uno de los objetivos del modelo). Además, aunque no se ha aplicado a centros públicos, el modelo cuenta con la versión adaptada al estudio de costes de este tipo de centros.

Este nuevo enfoque microeconómico del análisis estructural "coste-eficacia" que proponen los autores, a su juicio permite responder a una serie de cuestiones de capital interés para Directores y Administradores de los centros docentes, con vistas a la buena gestión de los mismos:

- ¿En qué grado afectan las distintas variables pedagógicas a los resultados académicos y a su coste?
- ¿Cómo comparar la estructura pedagógica y económica del propio centro con los demás centros del sector, con el fin de tener una idea de la eficacia relativa propia?
- ¿Se pueden conseguir unos índices, ratios y tendencias centrales del colectivo al que pertenece el centro, que sirvan de guía y orientación tanto para el control de los distintos niveles o colectivos educativos de referencia, como para la gestión del propio centro?
- ¿Cuáles son los aspectos educativos y estructurales que más afectan a la excelencia y en la economía de los esfuerzos pedagógicos?

Para la realización y experimentación de este modelo de evaluación de costes fue muy útil la experiencia de un modelo anterior realizado por el propio M. Ibar: el Modelo de Gestión de Centros Docentes (MGDC), programado en Fortran IV para ordenadores de gran capacidad. Esta vez se trata de adecuarlo a micro-ordenadores disponibles en

los centros docentes, por lo que se programa en hoja de cálculo, a partir de un programa Open Access para PC.

Por las características de la investigación realizada por el equipo que dirige Sánchez Martín, parecería difícil discernir a primera vista si se trata de la evaluación institucional (del Centro de Profesores de Getafe-Parla) o de la evaluación de un programa (de formación de profesores en este caso). Sin embargo, tanto por la denominación del trabajo, como por su contenido, puede concluirse que se trata de una evaluación institucional de un C.E.P., llevada a cabo a través de la evaluación de su oferta de formación para profesores (de su adecuación a la demanda, de su utilidad demorada, etc.).

La investigación sobre contexto y retraso mental (Espinosa, 1990), se ocupa de la construcción de un sistema de evaluación del contexto interno y externo de centros no ordinarios que escolarizan alumnos con retraso mental. Así, la investigación consiste en el diseño y elaboración de una serie de instrumentos para evaluar el contexto de dichos centros, cuya validación se deja para un estudio posterior.

El marco teórico fundamental en que se sitúan los autores gira en torno a la definición del concepto de contexto. De acuerdo con las directrices teóricas, "el prototipo de sistema de evaluación de centros para sujetos con retraso mental" que elaboran los investigadores consta de tres partes:

- 1) Entorno del centro: los aspectos o elementos considerados son localización, accesibilidad, congruencia, estética, conservación, equipamiento, riesgos y clima de relación.
- 2) Centro: se consideran los elementos anteriores, pero se añaden otros como la organización del centro, del personal, de los alumnos, las actividades recreativas, la implicación familiar, la socio-demografía de los alumnos y del profesorado, los elementos psicosociales, el impacto de las diferencias y el clima social.
- 3) Escenarios del centro: servicio de comedor, aulas, talleres, aseos, internado, pasillos y escaleras y salones comunitarios.

A partir de esta estructura se construye un amplísimo instrumento de evaluación que, dependiendo de la dimensión o del aspecto del contexto del centro que se pretenda evaluar, adopta una forma de escala tipo Likert o bien la de una lista de control o ficha de seguimiento con clave dicotómica de respuesta o bien de alternativa múltiple.

— *Características definitorias de la evaluación*

Agentes evaluadores

El modelo de estudio de costes de la educación diseñado por el equipo de Ibar Albiñana (modelo M.E.C.E.) está pensado para que sean los propios centros, sus administradores, los que lo apliquen, por lo que supone una aportación a la evaluación interno-económico-administrativa de los centros escolares. De otro lado, la experimentación del modelo en 22 centros responde más a una evaluación externa, puesto que son los propios investigadores quienes valoran la experiencia, aunque a partir de la información proporcionada por los centros que lo experimentaron.

En el caso de la investigación dirigida por Sánchez Martín, los agentes de la evaluación son los investigadores, aunque parten de dos fuentes más: la propia autoevaluación del Centros de Profesores (tiene un modelo por actividades realizadas ya muy consolidado) y la evaluación que de su oferta hacen los usuarios (profesores y equipos directivos). Por lo tanto, se trata de una evaluación externa que tiene muy en cuenta la opinión de los sectores implicados y las autoevaluaciones previamente realizadas por el propio C.E.P.

La investigación de Espinosa y colaboradores, no consiste en realidad en una evaluación de un programa o un centro, sino exclusivamente, como ya se dijo, en la elaboración de un instrumento de evaluación de los contextos de los centros de educación especial, por lo que no cabría hablar en este caso de evaluación interna o externa. Sin embargo, el propio sistema de evaluación construido puede aplicarse de distintas formas: evaluación externa o interna, bien realizada por expertos o por la dirección del centro.

Objetivos y finalidad de la evaluación

Ibar Albiñana y colaboradores señalan como objetivos prioritarios de su investigación:

- Atender a la necesidad sentida por los administradores de muchos centros escolares (sobre todo privados), de disponer de un programa estadístico de datos que pudiera aplicarse en los propios centros docentes para conseguir unos estudios com-

parativos de costes que sirvieran de orientación y referencia para la gestión.

- La necesidad de adaptar y reprogramar un modelo para ordenadores de gran capacidad elaborado anteriormente por el director de la investigación, a un estudio de costes más sencillo, manejable y concebido para microordenadores.
- Experimentar el modelo M.E.C.E. en una muestra de centros con el fin de detectar posibles problemas y ajustarlo para su mayor utilidad para los centros.

A la vista de estos objetivos, se puede concluir que tanto el modelo M.E.C.E., como la experimentación de que es objeto, tienen una clara finalidad formativa, por orientar la toma de decisiones de mejora, bien de la gestión y administración de los recursos de un centro, bien de las características del propio modelo.

Por su parte, la investigación de Sánchez Martín y sus colaboradores se propone dos objetivos fundamentales:

- Evaluar un Centro de Profesores (el de Getafe-Parla de Madrid) a través de su programación de formación (se evalúa toda la oferta).
- Diseñar un sistema de evaluación de los Centros de Profesores.

En esta línea, se hace evaluación con finalidad formativa, al tratar de mejorarse la calidad de la formación permanente que reciben los profesores a través de la oferta de ese C.E.P. y de otros, así como la calidad de los propios modelos de evaluación de la oferta formativa de los Centros de Profesores.

Los objetivos de la investigación dirigida por M.J. Espinosa están muy condicionados por el convencimiento —ampliamente apoyado por la evidencia empírica— de que para mejorar la calidad de vida de las personas con retraso mental y facilitar su integración en la vida de la comunidad, los ambientes o entornos de estas personas deben cumplir una serie de características.

En este sentido, y tras fundamentar ampliamente el tipo de características que deben perseguirse, este equipo de investigación se propuso crear un sistema (conjunto de instrumentos) de evaluación, que permitiera reconocer en qué medida los centros de educación especial

disponen de un contexto interno y externo favorable a la calidad de vida, al mayor control sobre su entorno, y a la integración en la comunidad (normalización) de los sujetos con retraso mental.

En este sentido, aunque el propio sistema elaborado no es objeto de evaluación en este trabajo, si podemos decir que la finalidad del sistema construido es formativa, puesto que su propósito es facilitar la toma de decisiones de mejora de los contextos de este tipo de centros, por parte de las autoridades educativas competentes. Se piensa que con esto se mejora enormemente la normalización conductual de los sujetos con retraso mental.

En definitiva, se propone una evaluación orientada a la toma de decisiones relacionadas con la aplicación de los principios de:

- Normalización contextual: desarrollo de un contexto que permita integrarse al deficiente.
- Normalización personal: adquisición por el sujeto de las habilidades necesarias para asumir roles y responsabilidades sociales culturalmente normativas.

Dimensiones evaluadas y momento de la evaluación

La investigación dirigida por el profesor Ibar, produce un modelo de estudio de costes de la educación que se puede alinear con facilidad a alguno de los modelos de evaluación institucional definidos en la literatura, concretamente al de "coste-eficacia" (rendimiento efectivo de las inversiones hechas en enseñanza), que es un modelo superador del de "coste-beneficio" o de rentabilidad de la educación, vigente hasta los años 70. En este sentido, el modelo construido considera variables de input-output al estudiar los costes en un centro. Por otra parte, la experimentación que de dicho modelo se realiza en 22 centros privados de la provincia de Madrid, aunque es una parte del trabajo poco descrita en la memoria de investigación, puede inferirse que se llevó a cabo una evaluación global del modelo.

Del trabajo realizado en la evaluación del C.E.P. de Getafe-Parla (Sánchez Martín) se desprende que se ha diseñado una evaluación global del centro a través de su programación de formación, que contempla cuatro tipos de evaluación: de contexto (normativa legal de la reforma, características de los centros adscritos, demandas de los centros y de los profesores, etc.); de entradas (recursos económicos y

humanos, el propio proyecto del C.E.P. y su oferta); de proceso (funcionamiento, organización, gestión, selección de personal, difusión de la oferta, evolución de la oferta, dificultades, etc.) y de producto (formación dada, eficacia, costos, satisfacción de la demanda, satisfacción de los usuarios con la oferta, participación, alcance, imagen del C.E.P., etc.). Sin embargo, los cuatro tipos de evaluación se han llevado a cabo casi simultáneamente (analizando la información procedente de discursos y documentos), no realizándose una evaluación de los procesos sobre la marcha.

La investigación de Espinosa y colaboradores, que como ya se ha dicho no evalúa sino que propone un sistema de evaluación del contexto de centros de educación especial, define éste como: "conjunto de características físicas, arquitectónicas, organizativas, de relación, clima social, etc, delimitadas espacio-temporalmente y que pueden ejercer funciones condicionadoras de conductas y moduladoras de patrones de respuestas individuales". Sin embargo, por la complejidad y amplitud del sistema, éste analiza más ámbitos que los que en principio comprende la definición cotidiana de contexto al incluir, entre otros aspectos, la organización y el clima social del centro (consideradas en otros modelos variables de entrada o de proceso).

De hecho este modelo evaluativo contempla las tres dimensiones de un entorno o contexto que se describían en la presentación de este trabajo: 1) entorno del centro; 2) centro propiamente dicho; y 3) escenarios del centro. Esta dimensionalización del contexto, y por tanto del sistema para su evaluación, responde a la aplicación del principio de normalización personal y contextual de los sujetos con retraso mental. Por otro lado, los propios autores señalan que con este instrumento se puede hacer una evaluación tanto global como parcial del contexto de los centros de educación especial.

Patrón empleado para la emisión de juicios de valor

La instancia de referencia empleada por el equipo de Ibar Albiñana para interpretar los datos de la experimentación de su modelo M.E.C.E. fue criterial, en tanto en cuanto el referente eran los propios objetivos de adecuación, sencillez y utilidad en función de los que fue diseñado el modelo. Por otra parte, como modelo de estudio de costes, el M.E.C.E. puede aplicarse tanto para hacer una evaluación criterial de la administración y gestión de un centro escolar, como para llevar a cabo una evaluación normativa. De hecho, una de las utilidades del

modelo consiste en la posibilidad de comparar el coste de la educación entre distintos centros, entre un centro y un grupo de centros o, con las adaptaciones oportunas, entre un centro o centros y el conjunto de centros del sistema educativo que pertenecen a un mismo grupo normativo (titularidad pública o privada, de una etapa u otra del sistema, de distintos tamaños, etc.).

Por su parte, la evaluación del C.E.P. de Getafe-Parla (Sánchez Martín) consiste en una evaluación criterial en tanto que la instancia de referencia empleada para interpretar la información recogida se define muy claramente. Los criterios empleados fueron:

- Las funciones que la Administración asigna a los Centros de Profesores (el grado de cumplimiento de las mismas).
- El propio proyecto del C.E.P. de Getafe-Parla (su oferta de formación).
- La satisfacción de las demandas de los profesores.

Al mismo tiempo, se hace evaluación normativa, al compararse la oferta del C.E.P. de Getafe con la del Centro de Leganés y con el conjunto de la oferta a nivel provincial.

En el caso de la investigación dirigida por Espinosa, hay que decir que la instancia de referencia es claramente criterial, ya que se fundamentan muy cumplidamente (teórica y empíricamente) y a priori, los criterios o características que debería reunir un contexto de un centro de educación especial que promoviese el desarrollo, la autonomía y la integración de los sujetos en él escolarizados. En este caso son criterios de interpretación los aspectos ambientales y organizativos deseables para la integración y normalización de alumnos con retraso mental. De hecho, el conjunto de instrumentos (cuestionarios y escalas) que aporta el equipo, se han construido de acuerdo a esta referencia criterial.

— *Diseño y desarrollo de la evaluación*

Como ya se ha dicho anteriormente, el modelo de estudio de costes de la educación (M. Ibar) responde a un diseño de análisis estadístico del coste-eficacia de la enseñanza, que se experimentó en 22 centros privados y la muestra estudiada fue incidental, ya que participaron

centros que voluntariamente respondieron a una demanda de colaboración ofrecida a la totalidad del colectivo de 342 centros de la F.E.R.E. en la provincia de Madrid. Esta muestra no puede considerarse representativa de los centros privados integrados en la F.E.R.E., ni tampoco suficiente por el número de centros que la componen, pero sin embargo sí permitió a los autores poner a prueba su modelo, detectar posibles deficiencias y ajustarlo. Por otra parte, el tipo de información que se recoge, así como el análisis de la misma son de tipo cuantitativo-estadístico, como parece lógico dada la naturaleza de la evaluación que se propone.

El interés metodológico de la propuesta que realizan Sánchez Martín y colaboradores como modelo de evaluación de un C.E.P., estriba en el uso de técnicas cualitativas sociológicas (grupos de discusión y análisis del discurso), la creación de un esquema de evaluación y de unas determinadas fuentes de información, así como la exclusión del cuestionario como técnica principal de recogida de información. El modelo, se articula en torno a 4 ejes:

- Producir un índice o eje secuencial de la actividad que se trata de evaluar.
- Determinar las fuentes de información mínimas relevantes (Ministerio de Educación y Ciencia, Centros de Profesores, centros educativos y profesores) y de las técnicas más adecuadas para alcanzar el nivel informativo necesario (examen de documentos, observación directa y grupos de discusión).
- Vaciar la información en el índice o eje secuencial; tratamiento de la misma y construcción de un discurso/representación de base para su análisis.
- El análisis evaluativo del discurso.

Los investigadores han optado por una estructura metodológica conformada por técnicas cualitativas, por entender que son las únicas que permiten acercarse a lo diverso.

Como fuentes de información se emplearon las documentales y personales del C.E.P. (Director y Equipo directivo, el Consejo y el Equipo Pedagógico), en concreto una muestra de 6 Equipos directivos y Claustros de los 144 centros y de los 3.100 profesores adscritos al CEP de Getafe-Parla. Estos centros eran de titularidad pública y privada concertada de los niveles educativos de E.G.B., B.U.P. y Formación Profesional.

En consecuencia, el tipo de información que se recoge es:

- Documental (porque en ella se encuentran expresados los objetivos a alcanzar, las actividades ofrecidas, desarrolladas y sus evaluaciones):
 - a. Documentación procedente de la Administración: marco legal (L.G.E., Constitución, L.O.D.E., L.O.G.S.E.); normativa diversa sobre el tema expedida por la Dirección General de Renovación Pedagógica y la Subdirección General de Formación del Profesorado; el Plan de Investigación Educativa y de Formación del Profesorado de 1989 (el Plan Marco); las Memorias de Actividades de Formación Permanente del profesorado (documentos evaluativos), los Planes de Formación Permanente del Profesorado y los Planes Provinciales de formación.
 - b. Documentación procedente del C.E.P. de Getafe: sus Planes y Memorias, los documentos divulgativos a los centros y un amplio conjunto de instrucciones, circulares, informes, boletines, etc.
- Personal: discurso de los profesores, directores, equipos directivos, etc. (del C.E.P. y de los centros adscritos).

Con respecto a los procedimientos utilizados para recoger la información, se centran en la recogida y examen de documentos, la observación directa (de entrevistas y reuniones de trabajo) y los grupos de discusión (en estos dos últimos casos, los investigadores han completado guiones para registrar la información).

El análisis de la información recogida se ha adecuado al tipo de información:

- a. Tratamiento de la información documental: vaciado selectivo conforme a un plan, y sistematización en una estructura predeterminada (concretada en un cuestionario). Se empleó una técnica sociológica de tipo cualitativo: a través de la lectura y selección de documentos, se construye el discurso resultante y se analiza buscando el significado y el sentido del mismo). Como resultado se obtienen unos ejes discursivos, donde van encajando los datos obtenidos por las demás técnicas sociológicas aplicadas.

b. Trabajo de campo en el C.E.P.:

- Se realizó una observación directa de los centros (2 reuniones de trabajo y 5 entrevistas en profundidad en cada centro). Se grabó, transcribió y analizó la información, realizándose unos informes en función de la estructura del cuestionario prefijado. Esta información, añadida a la obtenida en el análisis documental, fue trabajada para construir el discurso.
- Se formaron grupos de discusión con profesores de cada centro: el análisis de la información recogida se hizo también sobre el índice o estructura general.

Tras la realización de estos dos pasos (a y b) se procedió al análisis evaluativo. Así, la estructura empleada, tanto para el análisis de la información documental, como para la procedente del trabajo de campo, es la siguiente:

- Marco general: objetivos y establecimiento de la red donde se sitúa el C.E.P. de Getafe-Parla.
- Organización de este C.E.P.
- Producción de las ofertas de formación procedentes del M.E.C., la Dirección Provincial, los Centros de Profesores, y los centros y profesores adscritos.
- Representación de las ofertas en el mercado (valor de uso y cambio).
- Demandas manifiestas y latentes.
- El mercado o el encuentro entre ofertas y demandas.
- Consumo/uso de las ofertas.

Al utilizar un diseño naturalista-cualitativo de evaluación, la interpretación de los resultados se hace en términos valorativos, descriptivos y comprensivos de la realidad del C.E.P.

Por su parte, el equipo dirigido por M.J. Espinosa parece haber seguido un proceso deductivo en la construcción del sistema de instrumentos de evaluación del contexto de centros que escolarizan sujetos con retraso mental. Se parte de la definición muy operativa y minuciosa del constructo "contexto", se identifican sus partes o elementos,

dentro de cada una de ellas se operativizan las variables que las conforman y, a partir de ahí, van surgiendo los ítems para recoger la información. De este modo se obtienen un conjunto de escalas y cuestionarios que, aunque constituyen un sistema, sus autores afirman que cada uno de ellos puede ser utilizado independientemente.

Los rasgos técnicos más destacables de este sistema de evaluación son, pues, los siguientes:

- Si un centro está interesado sólo en evaluar su entorno contexto externo y no su contexto interno (centro y escenarios), puede hacerlo porque el prototipo lo permite.
- El sistema es flexible: las puntuaciones no son generales, sino específicas para cada parte del sistema. Por tanto, si un centro decide emplear todo el sistema, dispondrá al final de un mapa de puntuaciones sobre el contexto global del centro; si desea emplear parte del sistema, puede hacerlo porque cada apartado dispone de instrucciones específicas de cumplimentación y corrección. Al final del sistema se proporcionan las plantillas de corrección, y las estructuras en las que el usuario puede elaborar las gráficas de exploración allí donde sea pertinente. Esto facilita, además, la toma de decisiones sobre las posibles o necesarias modificaciones contextuales.

Por otra parte, aunque en la memoria de investigación no se ofrece información de sus resultados y consecuencias, se sabe que los autores llevaron a cabo dos estudios piloto del sistema de evaluación, puesto que lo experimentaron en dos centros de educación especial.

Por otra parte, y con respecto a algunos aspectos metodológicos del sistema de evaluación de contextos elaborado, puede decirse que las fuentes de información serán directas y personales (o bien la visita a las instalaciones); el tipo de información que se puede recoger versa sobre tres elementos del contexto considerados básicos por los investigadores:

- La contextura del mismo: la compaginación y unión de las partes que componen el todo contextual.
- Sus escenarios: la estructura física y social diseñada y organizada para facilitar la ejecución de una determinada trama escénica.

- - Sus escenas: la manifestación del fenómeno social real, que contiene una trama, unos actores y espectadores, y determinadas condiciones.

El procedimiento utilizado para recoger la información es el cuestionario multidimensional en el que se puede recoger información muy pormenorizada del contexto del centro. Para facilitar el análisis de la información se adjunta un manual de corrección, en el que se asignan puntuaciones a los datos recogidos que luego se reflejan en una escala de valoración para cada indicador.

— *Difusión y uso de los resultados*

En el caso de la investigación dirigida por el profesor Ibar Albiñana, y a juicio de los autores, la utilidad del modelo de evaluación de costes elaborado (MECE) para los centros es muy grande, ya que permite:

- Hacer estudios de los propios costes (de cada nivel educativo y de cuatro servicios más).
- Un fácil tratamiento de los datos. El tiempo necesario para introducir y procesar los datos de un centro de tamaño medio (de entre 900 y 1000 alumnos), hasta llegar a imprimir los resultados es de unas 3 horas.
- Estudios comparativos de costes (con otros centros) que sirvan de orientación y referencia para la gestión del centro.

Como afirman los propios autores: "Todo Administrador o Director que quiera ser consciente y responsable de la repercusión que pueden tener sus decisiones en la asignación de recursos a las distintas actividades del centro, debe tener bien claras estas interdependencias. De lo contrario se verá sorprendido por el dispararse de los costes de la enseñanza o frenado por la falta de financiación. Los índices y ratios de gestión que vamos a desarrollar son el gran medio de que puede valerse un Director para controlar la marcha de su centro conforme a una orientación dada". En esta medida, los destinatarios naturales del modelo resultante son los administradores de los centros de cualquier tipo de titularidad (pública y privada en sus dos modalidades). No obstante, los propios autores consideran que el modelo podría ser también

utilizado por las Unidades de planificación y estadísticas del Ministerio de Educación y Ciencia, así como los responsables de las mismas a nivel Comunidad Autónoma y provincial.

Por otra parte, el modelo incluye una serie de fichas en las que se recapitulan y ordenan los datos de todas las variables que han sido objeto de análisis, así como de "fichas resumen" del estudio del centro para facilitar la elaboración de un Informe que podrá ser utilizado con facilidad para orientar la toma de decisiones a la dirección del centro.

En la evaluación del C.E.P. de Getafe (Sánchez Martín), las conclusiones tienen estructura y orden y la audiencia directa de los resultados es el propio Centro de Profesores evaluado, al que se le entrega un Informe que se espera sirva para la reflexión.

En este sentido, los autores prevén que el uso de los resultados de la evaluación sea del tipo siguiente:

- La mejora de la formación de los profesores que se canalice a través del C.E.P. evaluado (e indirectamente de otros).
- La aplicación del modelo de evaluación de C.E.P. elaborado, a otros centros.
- La asunción de las críticas realizadas por parte de las autoridades competentes en la materia (provinciales, comunitarias, nacionales, etc.).

En la investigación dirigida por Espinosa, no se formalizan vías de presentación (informe de evaluación u otras) de los resultados que se obtendrían a partir de la utilización del sistema de evaluación de contextos. Por otra parte, la audiencia a la que se dirige el prototipo elaborado por los investigadores son los responsables políticos, administrativos y educativos de los que dependen los centros de educación especial, así como sus propios órganos de dirección y participación.

— *Metaevaluación*

La investigación dirigida por M. Ibar, al consistir tanto en el diseño de un modelo evaluativo de corte administrativo, como en su experimentación, realiza una evaluación de un modelo de evaluación lo que, aunque no puede considerarse propiamente una metaevaluación, sí supone un esfuerzo por someter a prueba y al juicio de otros (los destinatarios el modelo) el trabajo realizado.

En el caso de la evaluación dirigida por Sánchez Martín no se prevé ni se lleva a cabo una metaevaluación. En todo caso, se evalúan las autoevaluaciones que había venido haciendo el propio C.E.P. de Getafe-Parla de sus actividades formativas.

Igualmente, en la memoria de investigación dirigida por M.J. Espinosa, como se ha dicho, tan sólo se menciona la existencia de dos estudios piloto del sistema de evaluación elaborado. Por eso, se constata la ausencia de una valoración de dichos estudios.

CAPÍTULO V

CONSIDERACIONES FINALES: VISIÓN PANORÁMICA DE LAS INVESTIGACIONES

El conjunto de investigaciones que se ha revisado en esta obra se puede caracterizar también de acuerdo a una serie de variables que permiten obtener una perspectiva distinta a la que se ha ofrecido en páginas anteriores con el análisis de su contenido y de la tipología de las evaluaciones realizadas. En las tablas que se presentan a continuación se exponen, de acuerdo a la estructuración y presentación por grupos de investigaciones que se ha seguido en el Capítulo IV, los datos de cada estudio respecto a las siguientes variables: dirección del trabajo, período de realización, institución de procedencia (Centro de Profesores, Universidad, Administración, etc.), nivel educativo, contenido o área de trabajo, y número de investigadores participantes. Asimismo, se ofrece información que facilita la localización del resumen de cada investigación en los Catálogos que publica cada dos años el C.I.D.E..

EVALUACIÓN DE PROGRAMAS

| Evaluación de reformas y de programas experimentales | | | | | | |
|---|----------------|--------------------------------------|-------------------------------|-------------------------------------|---------------|----------------------|
| <i>Identificación del estudio</i> | <i>Periodo</i> | <i>Institución procedencia</i> | <i>Contenido</i> | | <i>Equipo</i> | <i>Catálogo CIDE</i> |
| | | | <i>Nivel educativo</i> | <i>Área de trabajo</i> | | |
| Aguado Odina, T | 1988-89 | U.N.E.D. | E. Infantil | Currículo | 1 | 1991-92 |
| Aguilera, H.J. et al. | 1985-89 | CIDE y Centro de Recursos para la EE | Preescolar C. Inicial | Integración alumnos con nee | 9 | 1989-90 |
| Álvaro Page, M. et al. | 1984-92 | C.I.D.E. | Enseñanzas Medias | Reforma experimental de la etapa | 21 | 1986-88 y 1989-90 |
| Álvaro Page, M. et al. | 1989-92 | C.I.D.E. | Enseñanza técnico-profesional | Reforma experimental de los módulos | 5 | 1991-92 |
| Escudero Muñoz, J.M. et al. | 1992-en curso | ICE de la U de Murcia | Primaria | Reforma | 7 | — |
| Domínguez Girón, M. et al. | 1989 | CIDE (contratación) | EGB/BUP/FP | Prensa como recurso didáctico | 3 | 1989-90 |

EVALUACIÓN DE PROGRAMAS CURRICULARES DE APLICACIÓN RESTRINGIDA

| Evaluación de programas curriculares para distintas áreas y niveles educativos | | | | | | |
|---|----------------|------------------------------------|------------------------|--------------------------|---------------|----------------------|
| <i>Identificación del estudio</i> | <i>Período</i> | <i>Institución procedencia</i> | <i>Contenido</i> | | <i>Equipo</i> | <i>Catálogo CIDE</i> |
| | | | <i>Nivel educativo</i> | <i>Área de trabajo</i> | | |
| Caldevilla Fdez., M.A. et al | 1987-89 | Universidad de Oviedo | Formación Profesional | Valoración de resultados | 4 | 1989-90 |
| Martín Hernández, F. et al. | 1989-91 | CEP de Madrid-Norte | Bachillerato | Filosofía | 6 | 1991-92 |
| Zabalza Beraza, M.A. et al | 1988-91 | Universidad Santiago de Compostela | Formación Profesional | Prácticas en alternancia | 12 | 1991-92 |

EVALUACIÓN DE PROGRAMAS CURRICULARES DE APLICACIÓN RESTRINGIDA

| Evaluación de programas instructivos para Primaria y Secundaria centrados en procesos didácticos | | | | | | |
|---|----------------|--------------------------------|------------------------|---------------------------|---------------|----------------------|
| <i>Identificación del estudio</i> | <i>Periodo</i> | <i>Institución procedencia</i> | <i>Contenido</i> | | <i>Equipo</i> | <i>Catálogo CIDE</i> |
| | | | <i>Nivel educativo</i> | <i>Área de trabajo</i> | | |
| Alonso Tapia, J. et al. | 1986-91 | Universidad Autónoma (Madrid) | E.S.O. | Lengua | 3 | 1991-92 |
| Garrido Gil, C.F., et al. | 1990-92 | CEP de Murcia | E.S.O. | Matemáticas | 4 | 1991-92 |
| Martínez Martín, M. et al. | 1989-91 | Universidad de Barcelona | E.S.O. | Educación Ética | 4 | 1989-90 |
| Mesa Franco, M. ^a .C. et al. | 1991-94 | CEP de Melilla | Primaria | Lengua castellana/materna | 6 | — |
| Rivera Otero, A. et al. | 1990-92 | CEP de Ciudad Lineal (Madrid) | E.S.O. | Ciencias Sociales | 5 | 1991-92 |
| Sánchez Miguel, E. et al. | 1988-89 | Universidad de Salamanca | E.S.O. | Lengua | 6 | 1989-90 |

EVALUACIÓN DE PROGRAMAS CURRICULARES DE APLICACIÓN RESTRINGIDA

| Evaluación de programas instructivos para Primaria y Secundaria centrados en el uso de recursos didácticos | | | | | | |
|---|----------------|---------------------------------|------------------------|----------------------------|---------------|----------------------|
| <i>Identificación del estudio</i> | <i>Período</i> | <i>Institución procedencia</i> | <i>Contenido</i> | | <i>Equipo</i> | <i>Catálogo CIDE</i> |
| | | | <i>Nivel educativo</i> | <i>Área de trabajo</i> | | |
| Bautista García Vera, A. et al. | 1985-87 | ICE de la U. Complutense-Madrid | EGB/BUP | Matemáticas | 22 | 1986-88 |
| Carmona Cerrato, J. et al. | 1991-94 | CEP de Don Benito (Badajoz) | EGB/BUP/COU | Lengua y Ciencias Sociales | 5 | — |
| Pérez de Landazabal Expósito, M. ^a C. et al. | 1987-91 | C.S.I.C. | BUP/COU | Ciencias Físicas | 3 | 1991-92 |

EVALUACIÓN DE PROGRAMAS CURRICULARES DE APLICACIÓN RESTRINGIDA

Evaluación de programas instructivos para Primaria y Secundaria contruidos sobre modelos teóricos definidos

| Identificación del estudio | Periodo | Institución procedencia | Contenido | | Equipo | Catálogo CIDE |
|--------------------------------|---------------|--------------------------------|------------------------|-----------------|--------|---------------|
| | | | Nivel educativo | Área de trabajo | | |
| Alonso Tapia, J. et al. | 1984-86 | ICE de la U Autónoma de Madrid | E.S.O. | Literatura | 13 | 1982-86 |
| Álvarez González, J. et al. | 1984-86 | ICE de la U.N.E.D. | E.S.O. | Literatura | 1 | 1986-88 |
| Gutiérrez Rodríguez, A. et al. | 1989-91 | Universidad de Valencia | Secundaria | Matemáticas | 7 | 1991-92 |
| Ramos Fernández, F. et al. | 1990-en curso | ICE de la U de Cantabria | E.S.O. y universitario | Ciencias | 5 | — |

EVALUACIÓN DE PROGRAMAS CURRICULARES DE APLICACIÓN RESTRINGIDA

| Evaluación de programas instructivos desarrollados en Educación superior | | | | | | |
|---|----------------|--------------------------------|------------------------|----------------------------------|---------------|----------------------|
| <i>Identificación del estudio</i> | <i>Período</i> | <i>Institución procedencia</i> | <i>Contenido</i> | | <i>Equipo</i> | <i>Catálogo CIDE</i> |
| | | | <i>Nivel educativo</i> | <i>Área de trabajo</i> | | |
| Caballero Sahelices, M.C. et al. | 1992-en curso | ICE de la U de Valladolid | Universitario | Ciencias Físicas | 3 | — |
| Bartolomé Pina, A. | 1985-87 | Universidad de Barcelona | Universitario | Estadística y nuevas tecnologías | 1 | 1986-88 |
| Sánchez Alciturri, J.M. et al. | 1985-87 | ICE de la U de Cantabria | Universitario | Geotecnia | 5 | 1986-88 |

EVALUACIÓN DE PROGRAMAS

| Evaluación de programas de Formación Inicial y Permanente del Profesorado | | | | | | |
|---|---------------|---------------------------------|---|----------------------------------|--------|---------------|
| Identificación del estudio | Periodo | Institución procedencia | Contenido | | Equipo | Catálogo CIDE |
| | | | Nivel educativo | Área de trabajo | | |
| Alvarez Castillo, J.L. et al. | 1992-95 | Vicerrectorado U. de Valladolid | Universitario | Formación inicial | 5 | — |
| Barrón Ruiz, A. et al | 1992-en curso | ICE de la U. de Salamanca | Profesores de EGB | Formación permanente | 2 | — |
| García Sallán, J. et al | 1992-94 | Universidad de Barcelona | Equipos directivos | Formación permanente | 9 | — |
| Marcelo García, C. et al. | 1988-90 | Universidad de Sevilla | Universitario (Magisterio) | Formación inicial | 4 | 1989-90 |
| Marcelo García, C. et al. | 1992-94 | Universidad de Sevilla | Profesores de Infantil, Primaria y Secundaria | Formación permanente | 6 | — |
| Nicolás Prieto, F. et al. | 1992-en curso | CEP de Valladolid | Profesores de Infantil y Primaria | Formación permanente | 3 | — |
| Portán Ariza, R. et al. | 1992-en curso | Universidad de Sevilla | Profesores no universitarios | Programas A.C.D. (F. permanente) | 3 | — |
| San Fabián Maroto, J.L. et al. | 1992-94 | Vicerrectorado U. de Oviedo | Profesores no universitarios | Formación permanente | 4 | — |

EVALUACIÓN DE PROGRAMAS

| Evaluación de programas de prevención de drogodependencias en la escuela | | | | | | |
|--|---------------|----------------------------------|-----------------|-------------------------|--------|---------------|
| Identificación del estudio | Periodo | Institución procedencia | Contenido | | Equipo | Catálogo CIDE |
| | | | Nivel educativo | Área de trabajo | | |
| Amengual Munat, M. et al. | 1987-90 | Universidad de Islas Baleares | E.S.O. | Educación para la salud | 4 | 1989-90 |
| March Cerdá, M. et al. | 1992-en curso | Vicerrectorado U. Islas Baleares | E.S.O. | Educación para la salud | 2 | — |

EVALUACIÓN DE CENTROS

| Identificación del estudio | Período | Institución procedencia | Contenido | | Equipo | Catálogo CIDE |
|-----------------------------------|---------------|---------------------------|--------------------------------------|-----------------------|--------|---------------|
| | | | Nivel educativo | Área de trabajo | | |
| Del Barrio del Campo, J.A. et al. | 1984-86 | ICE de la U. de Santander | Preescolar Primaria Secundaria | Centros ordinarios | 7 | 1982-86 |
| González Enriquez, C. et al. | 1985-86 | C.I.D.E. (contratación) | EGB/BUP/FP | Centros ordinarios | 2 | 1982-86 |
| Marín Ibáñez, R. et al. | 1984-86 | ICE de la U.N.E.D. | Universitario | Universidad | 6 | 1982-86 |
| Moreno Herrero, I et al. | 1992-en curso | CEP de Getafe | Formación permanente | Centros de Profesores | 19 | — |
| Morgenstern Picovsky, S. et al. | 1989-93 | U.N.E.D. | Formación permanente | Centros de Profesores | 2 | — |
| Palomares Aguirre, M.C. | 1992-en curso | CEP de Cuenca | Primaria | Centros ordinarios | 1 | — |
| Tejedor Tejedor, F.J. et al. | 1991-en curso | ICE de la U. de Salamanca | Universitario | Universidad | 6 | — |

INVESTIGACIONES CENTRADAS EN LA ELABORACIÓN DE MODELOS PARA LA EVALUACIÓN DE PROGRAMAS O CENTROS EDUCATIVOS

| <i>Identificación del estudio</i> | <i>Período</i> | <i>Institución procedencia</i> | <i>Contenido</i> | | <i>Equipo</i> | <i>Catálogo CIDE</i> |
|---|----------------|---|------------------------------------|--|---------------|----------------------|
| | | | <i>Ámbito</i> | <i>Modelo de evaluación</i> | | |
| Antúnez Marcos, S. et al. | 1992-en curso | Universidad de Barcelona | Formación permanente de profesores | Evaluación demorada de los efectos de los programas de F. permanente | 5 | — |
| De Juan Espinosa, M. et al. | 1985-90 | Universidad Autónoma de Madrid | Educación Especial | Evaluación del contexto de centros de e.e. | 5 | 1989-90 |
| Ibar Albiñana, M. et al | 1984-87 | Vicerrectorado. U. Pontificia Comillas (Madrid) | Centros no universitarios | Evaluación de costes de la enseñanza | 4 | 1986-88 |
| Sáez Brezmes, M. ^a J. et al. | 1992-en curso | CEP de Valladolid | Formación permanente | Autoevaluación de cursos de Actualización Científico-Didáctica | 5 | — |
| Sánchez Martín, E. et al. | 1991-92 | CEP de Getafe-Parla (Madrid) | Formación permanente | Evaluación de un CEP a través de sus ofertas y de la demanda | 11 | 1991-92 |
| Villar Angulo, L.M. et al. | 1992-94 | Universidades de Granada y Sevilla | Formación permanente | Sistema integrado de evaluación de formación en centros | 9 | — |

Entre las variables que identifican a cada investigación, hay algunas con las que es posible obtener una panorámica más general y descriptiva del conjunto de 48 investigaciones incluidas en esta obra. Se trata de variables que proporcionan un tipo de información que permite extraer algunas conclusiones adicionales y de interés, tanto para la propia Administración educativa, como para las personas o instituciones interesadas.

Tipo de convocatoria del C.I.D.E.

Hay diversos tipos de convocatorias en los que han sido financiadas las investigaciones presentadas: las Ayudas a la investigación, el Concurso Nacional de investigación educativa, o el decimotercero y último Plan Nacional de Investigación que gestionaba este tipo de convocatorias antes de la creación del C.I.D.E. Asimismo, se van a considerar aquí las investigaciones que han sido financiadas y realizadas por la propia Administración educativa, en este caso por el C.I.D.E. o, en colaboración con él, por otras unidades del Ministerio de Educación y Ciencia, y también aquellas que han sido contratadas por el C.I.D.E. y llevadas a cabo por empresas o investigadores independientes.

TIPO DE CONVOCATORIA O DE FINANCIACIÓN (1982-92)

| <i>Ayudas</i> | <i>Concurso</i> | <i>XIII Plan de Investigación</i> | <i>Realizadas o contratadas por la Administración educativa</i> |
|---------------|-----------------|-----------------------------------|---|
| 58,3% | 25% | 6,2% | 10,4% |

El hecho de que los proyectos financiados a través de las convocatorias de Ayudas doblen a los de los Concursos, hace pensar que, puesto que en estos diez años no todas las convocatorias del Concurso Nacional de Investigación han incorporado como tema preferente de estudio la evaluación de programas o centros educativos, puede suponerse que aquellos investigadores que deseaban trabajar sobre estos temas, han acudido a las sucesivas convocatorias de Ayudas a la investigación que han venido siendo casi siempre de tema libre.

Institución de procedencia

En las distintas convocatorias que periódicamente se realizan desde el C.I.D.E. se establece que los investigadores pueden optar por presentar su solicitud de diversas formas: a) a título particular; b) desde el Instituto de Ciencias de la Educación (ICE) o desde el Vicerrectorado de alguna Universidad española; c) desde un Centro de Profesores (CEP); d) o desde alguna Institución dedicada al fomento de la investigación, como ha sido el caso del Consejo Superior de Investigaciones Científicas (CSIC).

A los efectos descriptivos que se persiguen, también se van a considerar aquellas evaluaciones llevadas a cabo por la propia Administración educativa (por el C.I.D.E. u otras unidades).

En la Tabla que se presenta a continuación se ofrece el porcentaje de investigaciones que corresponde a cada una de estas situaciones.

INSTITUCIÓN DE PROCEDENCIA

| <i>Universidad (vicerrectorados, ICE, etc.) 66,6%</i> | <i>Centros de Profesores</i> | <i>Administración educativa (CIDE y otros)</i> |
|---|------------------------------|--|
| 66,7% | 20,8% | 12,5% |

Las conclusiones que pueden obtenerse de estos datos han de ser necesariamente prudentes, pues se desconocen las circunstancias que han llevado a cada equipo de investigación a canalizar su solicitud de una u otra manera. Sin embargo, aún desconociendo las razones que explican el hecho de que los vicerrectorados universitarios arrojen un total de proyectos de investigación financiados mayor que el de los Institutos de Ciencias de la Educación, sí parece claro que la casi totalidad de las investigaciones financiadas en este período de diez años procede de instancias universitarias (en torno al 67%), lo que las sitúa a gran distancia de las presentadas desde los Centros de Profesores (aproximadamente un 21%). Quizá el análisis de la variable que se comentará a continuación, el nivel educativo en el que trabajan los equipos de investigación o sus directores/as, puede dar una información complementaria que permita extraer alguna conclusión más aproximada sobre esta situación.

Composición del equipo de investigación (por niveles educativos)

En este caso, los equipos de investigación se caracterizan atendiendo al ámbito profesional en el que realizan su labor los investigadores que los componen. Así, se observa que se han formado distintos tipos de equipos de investigación:

a) *Equipos mixtos*, en los que pueden distinguirse:

- a.1. Equipos formados por profesores de Universidad y por profesores de niveles no universitarios (de Educación Infantil, Primaria y/o Secundaria). En este caso estarían más del 32% de los equipos de investigación que se incluyen en esta revisión.
- a.2. Equipos formados por profesores de Universidad, profesores de niveles no universitarios y por otros profesionales (procedentes de la Inspección, Equipos Psicopedagógicos, etc.). En este caso se encuentran sólo el 7% de los equipos de investigación.

b) *Equipos constituidos por investigadores de un sólo ámbito profesional*:

- b.1. Universitario: se trata de equipos formados exclusivamente por profesores de Universidad, lo que sucede en más del 37% de los equipos.
- b.2. No universitario: equipos formados por profesores de Educación Infantil, Primaria o Secundaria (bien de un sólo nivel, bien de varios niveles distintos), que es la composición del 21% de los equipos de investigación.

c) *Equipos de investigadores procedentes de empresas dedicadas, exclusiva o parcialmente, a la investigación o evaluación educativa*. En el conjunto de las investigaciones incluidas en esta obra, tan sólo hay un caso, el de un equipo procedente del Gabinete de Estudios y Aplicaciones (A.R.E.A), que realizó entre 1985 y 1986, contratado por el C.I.D.E., una evaluación de la utilización del equipamiento didáctico en centros públicos de Enseñanza Básica y Media.

COMPOSICIÓN DEL EQUIPO DE INVESTIGACIÓN

| <i>Mixto (varios ámbitos profesionales)</i> | | <i>Un mismo ámbito profesional</i> | | <i>Empresa</i> |
|---|---|------------------------------------|---------------------------------------|----------------|
| <i>Universitario/No universitario</i> | <i>Universitario/No universitario/Otros</i> | <i>Universitario</i> | <i>Infantil, Primaria, Secundaria</i> | |
| 32,5% | 7% | 37,2% | 21% | 2,3% |

Tema de investigación

Atendiendo al tema objeto de la investigación se han encontrado tres grandes grupos de investigaciones:

- A. Evaluación de programas:
 - a.1. Reformas y programas experimentales.
 - a.2. Programas curriculares de aplicación restringida.
 - a.3. Programas de formación del profesorado.

- B. Evaluación de centros.

- C. Elaboración de modelos de evaluación de programas y centros educativos.

OBJETO DE ESTUDIO DE LA INVESTIGACIÓN

| <i>EVALUACIÓN DE PROGRAMAS (73%)</i> | | | <i>EVALUACIÓN DE CENTROS</i> | <i>ELABORACIÓN DE MODELOS DE EVALUACIÓN (DE PROGRAMAS O CENTROS)</i> |
|--|---|---|------------------------------|--|
| <i>Reformas y programas experimentales</i> | <i>Programas Curriculares de aplicación restringida</i> | <i>Programas de Formación del Profesorado</i> | | |
| 17% | 60% | 22,8% | 14,5% | 12,5% |

A la vista de estos datos se puede afirmar que, del total de investigaciones revisadas (48), una gran mayoría (el 73%) se han centrado en

la evaluación de programas y, dentro de esta modalidad, más de la mitad de ellas (el 60%) han emprendido la evaluación de programas curriculares de aplicación restringida (a niveles educativos o grupos de alumnos concretos). A este grupo le siguen las evaluaciones de centros universitarios o no universitarios (más de un 14% de las investigaciones), a las que habría que añadir las investigaciones que, construyendo un modelo específico de evaluación, este es también un modelo destinado a evaluar algún tipo de centro educativo, lo que arroja un porcentaje de más del 20% de las evaluaciones.

Caracterización de las evaluaciones

En relación a los rasgos y criterios empleados en esta revisión para caracterizar las evaluaciones, se puede concluir lo siguiente.

En el ámbito de la evaluación de centros predominan las evaluaciones externas, con función formativa, y realizadas en su mayoría con un enfoque del diseño y de la metodología de tipo experimental/cuantitativo o mixto. En lo que respecta a su alcance, se acometen evaluaciones tanto de los centros globalmente considerados, como de aspectos parciales de los mismos. Destaca el hecho de que ningún equipo haya llevado a cabo una metaevaluación de la evaluación de centros por ellos realizada.

En evaluación de programas se han realizado con la misma frecuencia evaluaciones externas e internas, y tanto con referencia criterial como normativa, si bien predominan las que tienen función formativa —aún habiendo también evaluaciones con propósitos selectivos o sumativos—, las parciales en su alcance, y las realizadas con un diseño y una metodología de tipo cuantitativo/experimental. Tan sólo dos equipos han realizado una metaevaluación de su estudio.

En el conjunto de las evaluaciones de programas revisadas, son las evaluaciones de reformas y de programas experimentales (6 estudios evaluativos), las que manifiestan unos rasgos más diferenciados con respecto a los criterios empleados para el análisis. En prácticamente todos los casos de este grupo se han realizado evaluaciones externas, formativas, globales, con instancia de referencia mixta —evaluaciones normativas y criterios al mismo tiempo— y con enfoque metodológico fundamentalmente cuantitativo/experimental o, en algunos casos, mixto (cualitativo en sus primeras fases/cuantitativo en

la última fase del estudio). Por otra parte, sólo hay constancia de que se haya llevado a cabo una metaevaluación, y fué la realizada por el equipo que llevó a cabo la evaluación del Programa de Integración de alumnos con necesidades educativas especiales.

Conclusiones generales

La información derivada de este sencillo análisis descriptivo, ofrece una perspectiva del conjunto de investigaciones revisadas que se caracteriza por los siguientes rasgos: son mayoría las evaluaciones realizadas sobre programas educativos, y especialmente sobre programas curriculares específicos, frente a las aplicadas a centros.

Por otra parte, aunque la composición de los equipos ha tendido a equilibrarse a lo largo de este período entre 1982 y 1992, todavía parece haber más investigación —en este caso evaluación— realizada o dirigida por profesores universitarios que por profesionales (docentes o no) de otros niveles educativos.

La afirmación anterior, sin duda, ha de ser matizada o completada con la consideración de circunstancias muy diversas: la función investigadora que se exige a todo profesor universitario, la disponibilidad de tiempo y condiciones para investigar, o el propio tema de investigación que se contempla en esta revisión de diez años de investigación financiada por el C.I.D.E. Ciertamente, la evaluación de programas y centros educativos podría no haber estado entre los intereses o preocupaciones prioritarios de investigación de los profesores de niveles educativos como Infantil, Primaria y Secundaria durante estos diez años. De hecho, éste es un ámbito que tradicionalmente parece haber estado reservado a los “expertos” y a los “técnicos” en la materia, situación que casi con seguridad se percibe de forma distinta desde la puesta en marcha de procesos institucionales de autoevaluación en los últimos años.

Otro dato que parece apoyar lo que se viene afirmando es que, de hecho, las investigaciones centradas en la evaluación de programas curriculares de aplicación restringida desarrollados para los niveles de Infantil, Primaria o Secundaria que se incluyen en esta obra, son realizadas en muchos casos por equipos mixtos de investigadores (formados por profesores de niveles educativos universitarios y no universitarios) o, incluso, por equipos de profesores procedentes de Primaria o Secundaria.

En otro orden de cosas, el tamaño de los equipos está en torno a 5 o 6 investigadores por término medio. Sin embargo, la dispersión es muy grande, puesto que pueden encontrarse desde investigaciones realizadas por una sola persona, hasta equipos formados por 22 investigadores.

Con respecto a la evaluación que se práctica, ésta no parece ser muy diferente cuando el objeto de la evaluación es un programa o un centro. En general, predominan las evaluaciones formativas y criterios, aunque en el caso de la evaluación de reformas o programas experimentales la referencia al grupo normativo (por ej.: el rendimiento de los alumnos de centros de reforma comparado con el de los alumnos de centros con enseñanza vigente, etc.) ha sido muy frecuente. Asimismo, y sobre todo en el caso de los programas, parecen predominar las evaluaciones parciales centradas sólo en algunas dimensiones del programa o del centro en cuestión -normalmente entradas y resultados-, frente a evaluaciones globales (de entradas, contexto, procesos y resultados).

Destaca el hecho de que la mayoría de las evaluaciones de centros, aunque también muchas de las evaluaciones de programas, se hayan realizado con diseños cuasi-experimentales y análisis de tipo estadístico. Sólo en algunas evaluaciones de programas se han combinado enfoques y metodologías: cualitativa en las primeras fases del estudio, y cuantitativa en la última o últimas. Asimismo, se constata en la mayoría de los trabajos revisados la ausencia de una metaevaluación.

En todo caso, sea cual sea el carácter de los equipos de investigación en cuanto a ámbito profesional, estructura o tamaño, e independientemente de la institución que canalice los esfuerzos investigadores, lo importante será poder constatar en futuras revisiones que la investigación educativa que centra sus esfuerzos en la evaluación crece en número y en calidad. En todo caso, el balance que puede hacerse del período analizado es realmente positivo en relación a ambos parámetros.

BIBLIOGRAFÍA

- ÁLVARO PAGE, M. y col. (1988): *Evaluación Externa de la Reforma Experimental de las Enseñanzas Medias (I)*. C.I.D.E. Ministerio de Educación y Ciencia.
- ÁLVARO PAGE, M. y col. (1990): *Evaluación Externa de la Reforma Experimental de las Enseñanzas Medias (II)*. C.I.D.E. Ministerio de Educación y Ciencia.
- ÁLVARO PAGE, M. y col. (1992): *Evaluación Externa de la Reforma Experimental de las Enseñanzas Medias (III)*. C.I.D.E. Ministerio de Educación y Ciencia.
- ÁLVARO PAGE, M. y col. (1993): *Evaluación de los Módulos Profesionales: Estudio de la Reforma Experimental de la Enseñanza Técnico-Profesional*. C.I.D.E. Ministerio de Educación y Ciencia.
- ÁLVARO PAGE, M. (1993): *Seminario sobre metodología de investigación y evaluación en las Ciencias Sociales*. Instituto de la Mujer. Material Multicopiado
- ALVIRA, F. (1991): *Metodología de la Evaluación de Programas*. Madrid, CIS, Cuadernos Metodológicos n.º 2.
- ANGULO, J.F. (1990): *Innovación y Evaluación educativa*. Málaga, EAC. Universidad de Málaga.
- ASTIN, A.W. y PANOS, R.J. (1971): "The evaluation of educational programs". En THORNDIKE, L. (Ed.): *Educational measurement*, 2.ª ed. Washington: American council of Education, p.733.
- BARTOLOMÉ PINA, M. (1990): "Evaluación y Optimización de los diseños de intervención". En *Revista de Investigación Educativa*. Vol. 8, n.º 16. pp 39-59
- CARBALLO SANTAOLALLA, R. (1990): "Evolución del concepto de evaluación: desarrollo de los modelos de evaluación de programas". *Bordón*, 42 (4)

- CARBALLO SANTAOLALLA, R. (1991): "Introducción a la evaluación de programas de acción social". En *Revista Complutense de Educación*, Vol. 2 (1) pp. 111-126.
- CONSEJERÍA DE SALUD (1985): *Propuesta de educación para la salud*. Madrid, Comunidad de Madrid.
- CHOPPIN, B. H. (1985): "Assesment, Evaluation and Measurement". En HUSEN, T. y POSTLETHWAITE, T. N. (Dir.): *The International Encyclopedia of Education*. Oxford, Pergamon Press.
- COHEN, E. y FRANCO, R. (1988): *Evaluación de Proyectos Sociales*. Buenos Aires, Grupo Editor Latinoamericano S.R.L.
- COOK, T.D. y SHADISH, W.R. (1986): "Program Evaluation: the worldly science". *Annual Review of Psycology*, 37, pp 139-232
- CRONBACH, L.J. et. al. (1980): *Toward Reform of Program Evaluation*. San Francisco, Jossey Bass.
- DE LA ORDEN, A. (1985): "Investigación Evaluativa". En DE LA ORDEN, A. (Dir.): *Investigación Educativa. Diccionario de Ciencias de la Educación*. Madrid, Anaya.
- DE LA ORDEN, A. (1985b): "Modelos de evaluación universitaria". *Revista Española de Pedagogía*, 169-170, julio-diciembre, p.531.
- DICCIONARIO DE CIENCIAS DE LA EDUCACIÓN (1985). Madrid, Santillana.
- GAIRÍN SALLÁN, J. (1992): "Evaluación de Programas de Formación". En ALVAREZ FERNÁNDEZ, M. (Coord.) (1992): *La dirección escolar, formación y puesta al día*. Madrid, Escuela Española.
- GARCÍA RAMOS, J. M. (1989): *Bases pedagógicas de la evaluación. Guía práctica para educadores*. Madrid, Síntesis.
- GONZÁLEZ SOLER, A. (1989): "Evaluación de Programas Educativos". En ABARCA PONCE, M.P. (Coord.): *La evaluación de Programas Educativos*, Madrid, Escuela Española.
- GUBA, E.G. y LINCOLN, Y.S. (1982): *Effective evaluation*. California, Jossey-Bass.
- JOINT COMMITTEE ON STANDARS FOR EDUCATIONAL EVALUATION (1981): *Standars for Evaluations of Educational Programs, Projects and Materials*. New York, McGraw-Hill.
- LÁZARO MARTÍNEZ, A. (1990): "Problemas y Polémicas en torno al Diagnóstico Pedagógico". *Bordón*, 42(1).
- PÉREZ JUSTE, R. y MARTÍNEZ ARAGÓN, L. (1989): *Evaluación de centros y calidad educativa*. Madrid, Cincel.

- REICHARDT, C.H. y COOK, T. (1979): *Qualitative and quantitative methods en Evaluation Research*. London, Sage Publ.
- RUTMAN, L. (1984): *Evaluation Research Methods. A Basic Guide*. London, Sage Publ.
- SCRIVEN, M. (1967): "The methodology of Evaluation". En TYLER, R. W.; GAGNE, R.M. y SCRIVEN, M. (Eds.): *Perspectives of Curriculum Evaluation*. Chicago, Rand McNally.
- SCRIVEN, M. (1991): *Evaluation Thesaurus*. London, Sage.
- STUFFLEBEAM, D. y SHINKFIELD, A.J. (1987): *Evaluación sistemática. Guía teórica y práctica*. Madrid, Paidós/MEC.
- TALMAGE, H. (1982): "Evaluation of Programs". En MITZEL, H.E.; HARDIN BEST, J. y RABINOWITZ, W. (Dirs.) *Encyclopedia of Educational Research*. AERA. The Free Press, New York. 5.ª ed.
- TEJEDOR, F.J.; GARCÍA-VALCÁRCEL, A. y RODRÍGUEZ CONDE, M.J. (1993): "*Perspectivas metodológicas actuales de la Evaluación de Programas en el ámbito educativo*". VI Seminario de Modelos de Investigación Educativa. Madrid, 23-25 de septiembre. Documento Multicopiado.
- TIANA FERRER, A. (1993): "Evaluación de centros y evaluación del sistema educativo". *Bordón*. 45(3).
- VV.AA. (1992). "Evaluación de Programas". Número monográfico. *Bordón*, 43(4).
- VV.AA. (1990): *Evaluación del Programa de Integración Escolar de Alumnos con Deficiencias*. C.I.D.E. Ministerio de Educación y Ciencia.
- WORTHEN, B. R. (1985): "Programs Evaluation". En HUSEN, T. y POSTLETHWAITE, T. N. (Dirs): *The International Encyclopedia of Education*. Oxford, Pergamon Press.

LISTADO DE INVESTIGACIONES REVISADAS

- 1) AGUADO, M.^a T. (1991): *Evaluación de programas pertenecientes al Plan Experimental para la reforma de la Educación Infantil*. Universidad Nacional de Educación a Distancia (U.N.E.D.).
- 2) AGUILERA, M.^a J., ÁLVAREZ, K., BABÍO, M., COLL, C., ECHEÍTA, G., GALÁN, M., MARCHESI, A., MARTÍN, E., MARTÍNEZ, R. y E.P.I.(1989): *Evaluación del Programa de Integración Escolar de Alumnos con Deficiencias*. CIDE. Centro Nacional de Recursos para la Educación Especial.
- 3) ALONSO, J., GUTIERREZ, F., GONZÁLEZ, E., MATEOS, M.^a, HUERTAS, J.A., REY, A., OLEA, J., SÁNCHEZ, J., SÁNCHEZ, J.C., ARCE, E., PARDO, A., MONTERO, I. y URZAIZ, V. (1986): *Entrenamiento de habilidades cognitivas y enriquecimiento motivacional: nuevas tecnologías para la Educación Compensatoria*. ICE de la Universidad Autónoma de Madrid.
- 4) ALONSO, J., MATEOS, M.^a M. y GUTIÉRREZ, F. (1991): *Entrenamiento metacognitivo: desarrollo de programas para la mejora de la comprensión lectora y los procesos de razonamiento*. Universidad Autónoma de Madrid.
- 5) ÁLVAREZ, J. L., SCHARAMM, C., NIETO, M., PINTO, A. y GORDALIZA, M.^a A. (1995): *Evaluación de un programa experimental de entrenamiento cognitivo-social para profesores en formación*. Universidad de Valladolid.
- 6) ÁLVAREZ, E. (1986): *Enseñanza de la Literatura Española Contemporánea en el Ciclo Superior de E.G.B., tomando como base el Método de Tipos (Miguel Delibes), y el Sistema de Unidades*

- Modulares de Aprendizaje*. ICE de la Universidad Nacional de Educación a Distancia.
- 7) ÁLVARO, M., FERNÁNDEZ, M.^ª J., GARCÍA, C., IZQUIERDO, M.^ª C. y LÓPEZ, B. (1993): *Evaluación de los Módulos Profesionales: estudio de la Reforma experimental de la Enseñanza Técnico-Profesional*. CIDE.
 - 8) ALVARO, M., CALLEJA, J., ECHEVARRÍA, M.^ª J., ESCOBAR, M., MARCOS, S., MARTÍN, E., MÍNGUEZ, A., MUÑOZ, J., NAVAS, M.^ª J., PELETEIRO, I. y TRILLO, C. (1988): *Evaluación Externa de la Reforma Experimental de las Enseñanzas Medias (I)*. CIDE.
 - 9) ÁLVARO, M., BUENO, M.^ª J., CALLEJA, J., CERDÁN, J., ECHEVARRÍA, M.^ª J., GARCÍA, C., GÓMEZ, C., JIMÉNEZ, S., MARTÍN JAVATO, L., MÍNGUEZ, A. y NAVAS, M.^ª J. (1990): *Evaluación Externa de la Reforma Experimental de las Enseñanzas Medias (II)*. CIDE.
 - 10) ÁLVARO, M., ALONSO, M., GIL Y TRAVER, F., CERDÁN, J., GÓMEZ, C., CALLEJA, J. y VÉLAZ DE MEDRANO, C. (1992): *Evaluación Externa de la Reforma Experimental de las Enseñanzas Medias (III)*. CIDE.
 - 11) AMENGUAL, M., CALAFAT, A., MEJÍAS, G. y BORRRAS, M. (1990): *Evaluación del programa de prevención escolar "Tú decides"*. Universidad de Baleares.
 - 12) ANTÚNEZ, S. (en curso): *Análisis de los resultados de los programas de formación permanente del profesorado del MEC mediante la evaluación demorada de sus efectos. Determinación de criterios y elaboración de instrumentos*. ICE Universidad Barcelona.
 - 13) BARRIO, J.A., CALVO, G., DIEGO, R., GONZÁLEZ, M., LÓPEZ-DÓRIGA, M.^ª J., MEDRANO, P. y ORNILLA, I. (1986): *Evaluación de criterios ecológicos sobre percepción de centros escolares y estrategias para un plan de intervención*. ICE de la Universidad de Santander.
 - 14) BARRÓN, A., y DEL RÍO, J. (en curso): *Propuesta, experimentación y evaluación de un programa de formación del profesorado en la aplicación e investigación de la enseñanza constructiva*. ICE Universidad Salamanca.

- 15) BARTOLOMÉ, A. R. (1987): *Evaluación de Programas Didácticos y sus procesos de realización*.
- 16) BAUTISTA, A., ARREDONDO, J. M.^º, AVILÉS, M., BAENA, A., CENALMOR, M.^º I., CRIADO, P., ESTRADO, I., GARCÍA, M.^º G., GARCÍA, M.^º I., IGLESIAS, R., LUCAS, P., MARTÍN A., MARTÍNEZ, A., MUÑOZ, J., PÉREZ, C., PÉREZ, A., RODRÍGUEZ, J.A., ROJAS, F., SAN MILLÁN, C., SÁNCHEZ, A., SEBASTIÁN, A., y SORIANO, M.A.(1987): *Análisis de la eficacia del microordenador en la enseñanza de las matemáticas en 7.º de E.G.B. y 2.º DE B.U.P.* ICE de la Universidad Complutense de Madrid.
- 17) CABALLERO, M. (en curso): *Diseño, puesta en práctica y evaluación de un método didáctico con un enfoque constructivista para la enseñanza de la Física a nivel Universitario*. Vicerrectorado Universidad Valladolid.
- 18) CALDEVILLA, M.A., CASTAÑÓN, C., DÍEZ, P.P. y GONZÁLEZ, I. (1989): *¿Qué hacen los alumnos después de F.P.?* Universidad de Oviedo.
- 19) CARMONA, J., SÁNCHEZ, R., ANDRADA, J. L. y GALLEGO, M., (1994): *Investigación acerca de los recursos sonoros como materiales didácticos en las áreas curriculares de Lengua-Literatura y Ciencias Sociales. (Aplicación metodológica y evaluación de su utilidad en los procesos de aprendizaje)*. CEP de Don Benito-Villanueva de la Serena (Badajoz).
- 20) DOMÍNGUEZ, M., MÉNDEZ, A.M., y VILLALAÍN, J.L. (1989): *Evaluación del programa Prensa-Escuela*. CIDE.
- 21) ESCUDERO MUÑOZ, J. M. (en curso): *La disseminación y el desarrollo de la LOGSE como reforma educativa: un estudio evaluativo*. ICE Universidad Murcia.
- 22) GAIRÍN, J., DOMÍNGUEZ, G., GIMENO, J., RUÍZ, J. M.^º, TEJADA, J., NAVÍO, A., RUÍZ, C., TURÓN, J. y ANTÚNEZ, S. (1994): *Estudio de las necesidades de formación de los equipos directivos de los centros educativos*. Universidad Autónoma de Barcelona.
- 23) GARRIDO, C. F., PRIETO, M.^º D., VILLA, M.^º D. y BERMEJO, M.^º R. (1992): *Diseño e implementación de un programa sobre habilidades cognitivas en el área de Matemáticas*. CEP de Murcia.

- 24) GONZÁLEZ, C. y DELGADO, A. (Area) (1986): *Estudio sobre la utilización del equipamiento didáctico en centros públicos de Enseñanza Básica y Media*. CIDE.
- 25) GUTIÉRREZ, A., CORBERÁN, R., HUERTA, M., JAIME, A., MARGARIT, J., PEÑAS, A. y RUÍZ, E. (1991): *Diseño y evaluación de una propuesta curricular de aprendizaje de la Geometría en Enseñanza Media basada en el Modelo de Razonamiento de Van Hiele*. Universidad de Valencia.
- 26) IBAR, M., UROSA, B., y VELASCO, D. (1987): *Modelo estudio de costes de la educación (M.E.C.E.)*. ICE Universidad Pontificia de Comillas.
- 27) IÑÍGUEZ, C., SORIA, M., GAYOSO, M. y AGUDO, J. (1988): *La enseñanza de la morfología médica con módulos de autoaprendizaje y sus efectos sobre el recuerdo a largo plazo*. ICE de la Universidad de Valladolid.
- 28) DE JUAN, M., COLOM, R., FLORES, M., RODRÍGUEZ F. y SABATÉ, J. (1990): *Contexto y retraso mental: S.E.C.-U.A.M. Sistema de Evaluación de Centros*. Universidad Autónoma de Madrid.
- 29) MARCELO, C., GARCÍA, E., LÓPEZ, J., y CORONEL, J. (1990): *Diseño desarrollo y evaluación de un programa de inducción a la práctica profesional para profesores principiantes basado en la reflexión y la supervisión*. Universidad de Sevilla.
- 30) MARCELO, C., CORONEL, J., ESTEBARANZ, A., MAYOR, C., MINGORANCE, P., PARRILLA, A., RODRÍGUEZ, J., y SÁNCHEZ, M. (1994): *Facilitar el cambio en la escuela: Análisis y evaluación de dimensiones personales, didácticas, organizativas y de asesoramiento en el proceso de innovación y formación en la escuela*. Universidad de Sevilla.
- 31) MARCH, M. (En curso.) *Evaluación de programas de prevención de drogas en la escuela*. Universidad de Baleares.
- 32) MARÍN, R., RUBIO, M., GARCÍA, L., MURGA, M., PÉREZ, G., y LEBRERO, P. (1986): *Evaluación de la Enseñanza Superior a Distancia*. ICE de la Universidad Nacional de Educación a Distancia.
- 33) MARTÍN, F., BARBANCHO, E., HERNÁNDEZ, F., JIMÉNEZ, J., TORO, J. y BOTE, A. (1991): *Evaluación del curricu-*

- lum y de la didáctica de la Filosofía en 3.º de BUP*. CEP de Madrid-Norte.
- 34) MARTÍNEZ, M., PUIG, J.M., BUXARRAIS, R. y TRILLA, J.(1990): *Diseño y evaluación de programas de educación ética*. Universidad de Barcelona.
 - 35) MESA, C., SÁNCHEZ, S., VÁZQUEZ, M., TORTOSA, A., ARROYO, R., MOLINA, L. y MACHADO, I. (1994): *Exploración de la situación bilingüe en los escolares de Melilla. Propuestas y estrategias de intervención educativa*. CEP Melilla.
 - 36) MORENO HERRERO, I. (en curso) *No llorarse las mentiras, sino cantarse las verdades: el cuestionamiento de nuestro quehacer institucional*. CEP Madrid. Getafe.
 - 37) MORGENSTERN, S. y MARTÍN, E. (1993): *La evaluación de los CEP. Análisis de su contribución a la reforma educativa*. Universidad Nacional de Educación a Distancia (U.N.E.D.).
 - 38) NICOLÁS, F. (en curso) *Investigación operativa de los proyectos de formación del profesorado en centros de escuelas rurales para el desarrollo curricular en el marco de la Reforma del Sistema Educativo*. CEP N.º 1 de Valladolid.
 - 39) PALOMARES, M. (en curso) *Evaluación del Programa de Colegios Rurales Agrupados en la provincia de Cuenca*. CEP Cuenca.
 - 40) PÉREZ DE LANDAZABAL, M.ª C., GARCÍA-GALLO, J. y MORENO, J. (1991): *Eficacia didáctica de una metodología en el diseño de unidades didácticas de Física con ordenador*. Universidad Autónoma de Madrid.
 - 41) PORLÁN, R. (en curso) *Evaluación formativa de un proyecto institucional de Formación Permanente del Profesorado: el Programa de Actualización Científica y Didáctica (modalidad A)*.
 - 42) RAMOS, F. (en curso) *Estudio experimental basado en la Teoría de Lakatos acerca del progreso de la Ciencia y sus implicaciones didácticas*. Universidad Cantabria. Santander.
 - 43) RIVERA, A., FRANCO, F., GARCÍA, F. y JÓRGE, M. (1992): *La investigación del alumno desde la perspectiva de las Ciencias Sociales*. CEP de Ciudad Lineal (Madrid).

- 44) SÁEZ, M. J. (en curso) *Virtualidad de un modelo de evaluación para cursos de actualización científico-didáctica de los CEPs*. CEP N.º 1 de Valladolid.
- 45) SAN FABIÁN, J., PÉREZ, R., RODRÍGUEZ, B. y RUANO, E. (1994): *Condiciones y posibilidades de la formación del profesorado en centros. Hacia un modelo de evaluación*. Universidad de Oviedo.
- 46) SÁNCHEZ, J., SAGASETA, C., ARRO, R., CAÑIZAL, J. y CAPELLÁN, F. (1987): *Evaluación de métodos de simulación informática en Enseñanzas Técnicas Superiores*. ICE de la Universidad de Cantabria.
- 47) SÁNCHEZ, E., FERNÁNDEZ DE CASTRO, I., DOMÍNGUEZ, P., MÉNDEZ, M., DE RON, E., FERNÁNDEZ, M.^a, DE ELEJABEITIA, C., VAQUERIZO, P., SAN MIGUEL C., GUMPERT, L. y VALERO, A. (1992): *Evaluación del CEP Getafe-Parla a través del estudio de sus ofertas y de las demandas de los profesores y centros educativos de su competencia*. CEP de Getafe.
- 48) SÁNCHEZ, E., ORRANTÍA, J., ROSALES, J., DE MIGUEL, E., FERNÁNDEZ, E. y GARCÍA, E. (1989): *Un programa para instruir en la comprensión de textos expositivos en la Etapa Secundaria Obligatoria*. Universidad de Salamanca.
- 49) TEJEDOR, F. (en curso) *Evaluación de las condiciones personales, materiales y funcionales en las que se desarrolla la docencia en la Universidad de Salamanca*.
- 50) VILLAR, L. M., VICENTE, P., BOLÍVAR, A., FERNÁNDEZ, M., GALLEGO, M.^a J., LEÓN, M.^a J., MOLINA, E., MORAL, C. y PÉREZ GARCÍA, P. (1994): *Marco conceptual para el diseño de sistemas integrados de evaluación de programas de desarrollo profesional del docente en centros educativos*. Universidades de Granada y Sevilla.
- 51) ZABALZA, M. A., RIAL, A., SALGADO, F., ALONSO, A., GONZÁLEZ, R., CERDEIRIÑA, E., GONZÁLEZ, C., BRAVO, J., ROUCO, J., DOMINGO, M.^a F., CEPEDA, O. e IGLESIAS, L. (1991): *Evaluación del Programa de Prácticas en Alternancia en la Comunidad Autónoma Gallega*. Universidad de Santiago de Compostela.

RELACIÓN ALFABÉTICA DE AUTORES

- Aguado Odina, M.^a Teresa
- Agudo, Javier
- Aguilera, M.^a Jesús
- Alonso Blanco, Albano
- Alonso de la Hera, Miriam Helena
- Alonso Tapia, Jesús
- Álvarez Castillo, José Luis
- Álvarez, Katia
- Álvarez González, Elvira
- Álvaro Page, Mariano
- Amengual Munar, Miguel
- Andrada Barrantes, José Luis
- Antúnez Marcos, Serafín
- Arce Sáez, Enrique
- AREA
- Arredondo Rodríguez, José M.^a
- Arro, Rafael
- Arroyo González, Rosario
- Avilés Martínez, Manuel
- Babío, Mercedes
- Baena Carbajal, Ángel
- Barbancho Blanco, Eusebio
- Barrio del Campo, José Antonio
- Barrón Ruiz, Ángela
- Bartolomé Pina, Antonio R.
- Bautista García-Vera, Antonio
- Bermejo García, M.^a Rosario
- Bolívar Botía, Antonio

- Borrás Ripoll, Matilde
- Bote Bonilla, Andrés
- Bravo López, Javier
- Bueno Monreal, M.^a José
- Buxarraís Estrada, M.^a Rosa
- Caballero Sahélices, M.^a Concesa
- Calafat Far, Amador
- Caldevilla Fernández, Miguel Ángel
- Calvo Díaz, Gonzalo
- Calleja Sopeña, José Ángel
- Cañizal, Jorge
- Capellán, Federico
- Carmona Cerrato, Julio
- Castañón Muñiz, Cirilo
- Cenalmor Rodríguez, M.^a Isabel
- Cepeda Romero, Olga
- Cerdán Victoria, Jesús
- Cerdeiriña Vázquez, Elvira
- Colom Marañón, Roberto
- Coll Salvador, César
- Corberán Salvador, Rosa
- Coronel Llamas, José Manuel
- Criado Sarabia, Pedro
- Delgado Solís, Ana M.^a
- Diego Cabarga, Robustiano
- Díez Suárez, Pedro Pascual
- Domingo Ouvard, M.^a Fernanda
- Domínguez Fernández, Guillermo
- Domínguez Girón, Mar
- Domínguez, Pilar
- E.P.I.
- Echeíta Sarrionandia, Gerardo
- Echeverría Cubillas, M.^a José
- Elejabeitia Tavera, Carmen de
- Escobar Mercado, Modesto
- Escudero Muñoz, Juan Manuel
- Estebaranz García, Araceli
- Estirado Martín, Isabel
- Fernández, Esther

- Fernández Cruz, Manuel
- Fernández de Castro Sánchez de Cueto, Ignacio
- Fernández Galleguillos, M.^a José
- Fernández García, M.^a Teresa
- Flores Lucas, M.^a del Valle
- Franco Franco, Felipe
- Gairín Sallán, Joaquín
- Galán, Marisa
- Gallego Arrufat, M.^a Jesús
- Gallego Cidoncha, Manuel
- García Aretio, Lorenzo
- García Arribas, M.^a Gregoria
- García Cejudo, M.^a Isabel
- García Jiménez, Eduardo
- García López, Carmen
- García Peña, Francisco
- García-Gallo Peñuela, Joaquín
- García Pérez, Emma
- Garrido Gil, Carlos Fulgencio
- Gayoso, Manuel J.
- Gil y Traver, Flora de
- Gimeno Soria, Xavier
- Gómez Bueno, Carmuca
- González Alonso, Elena
- González de la Vega, Manuela
- González de los Santos, Isidoro
- González Enríquez, Carmen
- González Iglesias, Rosa
- González Prieto, Carmen
- Gordaliza Rodríguez, M.^a Ángeles
- Gumpert Castillo, Leopoldo
- Gutiérrez Martínez, Francisco
- Gutiérrez Rodríguez, Ángel
- Hernández González, Francisco
- Huerta Palau, Manuel Pedro
- Huertas Martínez, Juan Antonio
- Ibar Albiñana, Mariano
- Iglesias Forneiro, Lina
- Iglesias Pérez, Rafael

- Iñiguez Lobeto, Carlos
- Izquierdo Nebreda, M.^a Carmen
- Jaime Pastor, Adela
- Jiménez Ruiz, José María
- Jiménez Suárez, Sofía
- Jorge Campos, M.^a Luz
- Juan Espinosa, Manuel de
- Lebrero Baena, M.^a Paz.
- León Guerrero, M.^a José
- López-Dóriga Fragua, M.^a José
- López Pérez, Beatriz
- López Yáñez, Julián
- Lucas Padín, Paz
- Machado Cabrerizo, Irene
- Marcelo García, Carlos
- March Cerdá, Martí Xavier
- Marchesi Ullastres, Álvaro
- Marcos Pérez, Susana
- Margarit Garrigues, Juan Bautista
- Marín Ibañez, Ricardo
- Martín Checa, Elena
- Martín Donaire, Antonio
- Martín Hernández, Felipe
- Martín Javato Romero, Laura
- Martín Ortega, Elena
- Martín Rodríguez, Eustaquio
- Martínez Arias, Rosario
- Martínez Carballo, Ángel
- Martínez Martín, Miguel
- Mateos Sanz, M.^a del Mar
- Mayor Ruiz, Cristina
- Medrano González, Paloma
- Mejías García, Gema
- Méndez Conde, M.^a Dolores
- Méndez Miras, Ana M.^a
- Mesa Franco, M.^a Carmen
- Miguel, Elena de
- Mingorance Díaz, Pilar
- Mínguez Ceballos, Ana

- Molina Domínguez, Lucía
- Molina Ruiz, Enriqueta
- Montero García-Celay, Ignacio
- Moral Santaella, Cristina
- Moreno Rebollo, José M.^a
- Moreno Herrero, Isidro
- Morgenstern Pitcovsky, Sara
- Muñoz Durán, Juan
- Muñoz Yangüas, José Luis
- Murga Menoyo, M.^a Ángeles
- Navas Ara, M.^a José
- Navío Gámez, Antonio
- Nicolás Prieto, Fernando
- Nieto Bedoya, Margarita
- Olea Díaz, Julio
- Ornilla García, Inmaculada
- Orrantía Rodríguez, José
- Palomares Aguirre, M.^a del Carmen
- Pardo Merino, Antonio
- Parrilla Latas, Ángeles
- Peleteiro Ramos, Isabel
- Peñas Pascual, Antonio
- Pérez de Landazabal Expósito, M.^a Carmen
- Pérez García, Purificación
- Pérez Pérez, Ramón
- Pérez Rodríguez, Alicia
- Pérez Rodrigo, Carmen
- Pérez Serrano, Gloria
- Pinto Martín, Amparo
- Porlán Ariza, Rafael
- Prieto Sánchez, M.^a Dolores
- Puig Rovira, José M.^a
- Ramos Fernández, Francisco
- Rey González, Asunción
- Rial Sánchez, Antonio
- Río Sánchez, José del
- Rivera Otero, Alicia
- Rodríguez Díez, Blanca
- Rodríguez López, José A.

- Rodríguez López, José M.^a
- Rodríguez Santos, Francisco
- Rojas Ruíz, Francisco
- Ron, Elena de
- Rosales Pardo, Javier
- Rouco Ferreiro, Javier
- Ruano Hernández, Elvira
- Rubio, M.^a José
- Ruíz Bueno, Carmen
- Ruíz Pérez, Enrique
- Ruíz Ruíz, José M.^a
- Sabaté Mur, Josefina
- Saez Brezmes, M.^a José
- Sagaseta, César
- Salgado López, Faustino
- San Fabián Maroto, José Luis
- San Miguel, Carmen
- San Millán García, Carmen
- Sánchez Alciturri, José Manuel
- Sánchez Calvo, Valentín
- Sánchez Fernández, Sebastián
- Sánchez Ferrer, Javier
- Sánchez García, José Carlos
- Sánchez Gómez-Coronado, Reyes
- Sánchez Martín, Eduvigis
- Sánchez Miguel, Emilio
- Sánchez Moreno, Marita
- Sánchez Torija, Alberto
- Schramm Martín, Carlos
- Sebastián Gómez, Alberto
- Soria, Miguel A.
- Soriano Jiménez, Miguel A.
- Tejada Fernández, José
- Tejedor Tejedor, Fco. Javier
- Toro Mérida, Joaquín
- Tortosa Pérez, Amelia
- Trilla Bernet, Jaime
- Trillo Marco, Carmen
- Turón García, Javier

- Urosa Sanz, Belén
- Urzaiz Celigueta, Victoria
- Valero García, Alfonso
- Vaquerizo Dulce, Pilar
- Vázquez Tortosa, M.^a del Mar
- Velasco, Dessirée
- Vélaz de Medrano Ureta, Consuelo
- Vicente Rodríguez, Pedro de
- Villa Luna, M.^a Dolores
- Villalaín Benito, José Luis
- Villar Angulo, Luis Miguel
- Zabalza Beraza, Miguel A.



Ministerio de Educación y Ciencia

Secretaría de Estado de Educación

Dirección General de Renovación Pedagógica