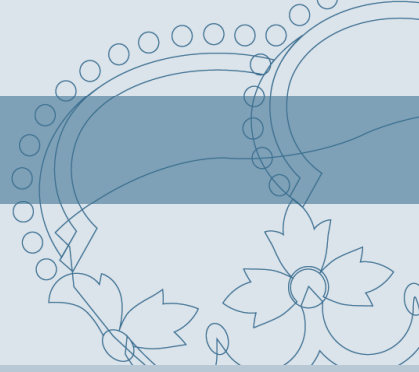


Colección

MONOGRAFÍAS



Evaluación de políticas y reformas educativas en España (1982-1992)

Tres experiencias de metaevaluación

ESTER GARCÍA SÁNCHEZ



INAP

INSTITUTO NACIONAL DE ADMINISTRACIÓN PÚBLICA

EVALUACIÓN DE POLÍTICAS
Y REFORMAS EDUCATIVAS EN ESPAÑA (1982-1992)

Tres experiencias de metaevaluación

EVALUACIÓN DE POLÍTICAS
Y REFORMAS EDUCATIVAS EN ESPAÑA (1982-1992)
Tres experiencias de metaevaluación

Ester García Sánchez

INSTITUTO NACIONAL DE ADMINISTRACIÓN PÚBLICA
MADRID, 2011

Colección: MONOGRAFÍAS

FICHA CATALOGRÁFICA DEL CENTRO
DE PUBLICACIONES DEL INAP

GARCÍA SÁNCHEZ, Ester

Evaluación de políticas y reformas educativas en España (1982-1992)
[Texto impreso] : tres experiencias de metaevaluación / Ester García Sánchez.
– 1ª ed. – Madrid : Instituto Nacional de Administración Pública, 2011. –
207 p. ; 24 cm. – (Monografías)

Bibliografía: p. 187-201

ISBN 978-84-7351-406-4. – NIPO 852-11-020-2

1. España-Política educativa-1982-1996. I. Instituto Nacional de Ad-
ministración Pública (España). II. Título. III. Serie

37.014(460)“1982/1996”

Primera edición: Septiembre 2011

Catálogo general de publicaciones oficiales:
<http://www.060.es>

Queda prohibida, salvo excepción prevista en la ley, cualquier forma de reproducción, distribución, comuni-
cación pública y transformación de esta obra sin contar con autorización de los titulares de propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual (arts. 270 y sigs. del Código Penal).

Edita:

INSTITUTO NACIONAL DE ADMINISTRACIÓN PÚBLICA

www.inap.es

ISBN: 978-84-7351-406-4 (formato papel); 978-84-7351-405-7 (formato electrónico)

NIPO: 852-11-020-2 (formato papel); 852-11-019-X (formato electrónico)

Depósito Legal:

Preimpresión: Composiciones RALI, S.A.

Impresión: Publidisa

En esta publicación se ha utilizado papel reciclado libre de cloro de acuerdo con los criterios medioambien-
tales de la contratación pública.

AGRADECIMIENTOS

Quisiera dejar constancia de mi sincero agradecimiento a todos cuantos, de una forma u otra, me han brindado su ayuda y apoyo a lo largo de este camino que comenzó con la elaboración de la tesis doctoral.

En primer lugar, desearía expresar mi gratitud al profesor Francisco J. Vnaclocha, director de la tesis, y a quien debo mucho más de lo que puedo expresar aquí. Siempre he podido contar con su confianza, sus consejos y sus enseñanzas siempre acertadas y, cómo no, con su estímulo y paciencia en los momentos de desánimo.

Los profesores Julián Santamaría Ossorio, Rafael Bañón i Martínez, José Ignacio Cases Méndez, Guillermo Márquez Cruz y Manuel Villoria Mendieta, miembros del tribunal que valoró la tesis doctoral, juzgaron mi trabajo con amabilidad e hicieron enriquecedores comentarios que confío en haber sabido aprovechar.

Mi reconocimiento también a quienes, generosamente, me dedicaron su tiempo y aceptaron colaborar en la investigación como entrevistados o informantes.

Un recuerdo muy especial a los profesores Barry MacDonald (University of East Anglia), Saville Kushner (University of the West of England) y María José Sáez Brezmes (Universidad de Valladolid) con quienes tuve la fortuna de iniciarme en el campo de la evaluación y a quienes debo gran parte de cuanto sé sobre el tema.

En el área de Ciencia Política y de la Administración de la Universidad Carlos III he encontrado siempre el respaldo necesario para desarrollar mis tareas docentes e investigadoras. Antonio Natera, Ruth Martín (hoy en la Universidad de la Laguna), Isabel Wences, Roberto Losada y Javier Redondo me han brindado, durante todos estos años, su ayuda y amistad. Su generosidad intelectual y su calidad humana han hecho que la vida académica, pese a sus rigores, fuera algo grato y entrañable.

AGRADECIMIENTOS

Por más que suene a lugar común, y como ellos bien saben, sin Maruja, Ana, Juan, Ana Victoria y Rodrigo, todos estos años de dedicación no habrían sido igual de llevaderos. Pablo, Clara y José Miguel, a quienes tantas y tantas horas ha robado mi trabajo, me regalan a diario el cariño y las sonrisas que necesito. A ellos dedico este libro.

ÍNDICE

Relación de abreviaturas	13
Índice de cuadros	17
Introducción	19

PRIMERA PARTE

Capítulo I. La pluralidad de concepciones de metaevaluación	27
1. Dibujando los contornos del concepto (de metaevaluación).	27
2. En busca de criterios para evaluar las evaluaciones. Un recorrido por los principales enfoques metaevaluativos.	31
2.1. El cuantitativismo en metaevaluación: la preocupación por la calidad técnica de las evaluaciones..	31
2.2. La metaevaluación desde una perspectiva cualitativista.. . . .	33
2.3. Las propuestas conciliadoras.	35
Capítulo II. Un esquema de metaevaluación desde la Ciencia Política y el análisis de políticas públicas	45
1. A vueltas con el carácter político de las evaluaciones	45
2. Las premisas de partida	50
2.1. Las dos dimensiones de la evaluación	50
2.2. Actores y decisiones	51
2.2.1. El concepto de <i>actor</i>	51
2.2.2. ¿Qué factores explican las decisiones y las actuaciones de los actores?	53
2.2.3. Hacia una clasificación de los actores: naturaleza y modalidad de participación.	55

3. El esquema propuesto: utilidad y rasgos generales	57
4. Las evaluaciones objeto de la metaevaluación	61
5. Elementos para la aplicación del esquema a los estudios de caso. .	63
5.1. Las fuentes de información y la selección de los informantes	63
5.2. Las fases de los procesos evaluativos	65
Capítulo III. De la reforma educativa a la evaluación de programas: los escenarios de la política educativa española entre 1982 y 1996	69
1. Del programa electoral al programa de gobierno: las primeras ac- tuaciones del ministerio Maravall	70
1.1. Las medidas de promoción educativa	71
1.2. La elaboración de la Ley Orgánica del Derecho a la Educación	73
1.3. El desarrollo de la LODE	75
2. Los antecedentes de la LOGSE	76
2.1. Los programas experimentales de reforma	76
2.2. El debate público sobre la reforma: el Proyecto para la Refor- ma de la Enseñanza	80
2.3. El Libro Blanco para la Reforma del Sistema Educativo	83
3. La Ley de Ordenación General del Sistema Educativo	84
4. Calidad y evaluación en la agenda de gobierno: la educación en España después de la LOGSE	86

SEGUNDA PARTE

Capítulo IV. La evaluación externa de la reforma experimental de enseñanzas medias	91
1. El programa de reforma experimental de enseñanzas medias	91
1.1. La situación de partida: las enseñanzas medias a comienzos de los años ochenta.	91
1.2. El proceso de reforma	92
1.2.1. Los principios y los objetivos de la experimentación .	93
1.2.2. La reforma del primer ciclo de enseñanzas medias . . .	94
1.2.3. La reforma del segundo ciclo	96
1.2.4. Balance del proceso experimental	100
2. La evaluación externa de la reforma experimental	101
2.1. El diseño de la evaluación	102
2.1.1. De la iniciativa a la formulación de objetivos	102

2.1.2.	La apuesta por un enfoque excluyente de evaluación por objetivos	104
2.1.3.	La aparente ausencia de criterios de valor	108
2.1.4.	¿Una evaluación externa?	109
2.1.5.	El procedimiento de muestreo	109
2.1.6.	La elección de las técnicas de recogida y análisis de datos	113
2.2.	El desarrollo de la evaluación: la obtención de datos, el tratamiento estadístico y el análisis de la información	115
2.3.	Los resultados	116
2.3.1.	Las principales conclusiones y el estilo de los informes	116
2.3.2.	La utilidad y el impacto de la evaluación.	118
Capítulo V.	La evaluación del programa de integración de alumnos con necesidades educativas especiales	121
1.	El programa de integración escolar de alumnos con necesidades educativas especiales	121
1.1.	De la segregación a la integración	121
1.2.	La regulación de la educación especial en España: la asunción del principio de integración	122
1.3.	El programa experimental de integración	123
1.3.1.	Los principios y las características del programa.	123
1.3.2.	El proceso de implantación	124
2.	La evaluación del programa de integración.	127
2.1.	Los comienzos	127
2.2.	El diseño de la evaluación	129
2.2.1.	La definición de los objetivos.	129
2.2.2.	Un enfoque metodológico supuestamente iluminativo	130
2.2.2.1.	El modelo teórico	132
2.2.2.2.	Las fases de la evaluación y las técnicas de recogida y análisis de datos	134
2.2.3.	Los criterios de valor	140
2.3.	Del diseño al análisis de datos	140
2.4.	La etapa final de la evaluación	144
2.4.1.	Un documento de legitimación del programa: el informe final de evaluación	144
2.4.2.	Los usos del proceso de evaluación	148

Capítulo VI. La evaluación del proyecto <i>Atenea</i> para la introducción de los ordenadores en las aulas	153
1. El proyecto <i>Atenea</i> para la introducción de los ordenadores en las aulas	153
1.1. Los antecedentes y la puesta en marcha del proyecto	153
1.2. Objetivos y ámbitos de aplicación	154
1.3. La formación del profesorado	157
1.4. La creación del Programa de Nuevas Tecnologías de la Información y la Comunicación. El proyecto <i>Atenea</i> en el marco del PNTIC	159
2. El proceso de evaluación del proyecto <i>Atenea</i>	162
2.1. La iniciativa y el impulso de la evaluación	162
2.2. El diseño: «pluralismo» y negociación	163
2.2.1. Los objetivos y el modelo de evaluación	163
2.2.2. Las técnicas de recogida y análisis de datos.	166
2.2.3. Las opiniones de los actores implicados como criterios de valor de la evaluación	168
2.3. El desarrollo del proceso evaluativo	168
2.3.1. El trabajo de campo	168
2.3.2. El informe de progreso	169
2.4. El final de la evaluación	171
2.4.1. La presentación del informe final y la ruptura de las relaciones entre clientes y evaluadores.	171
2.4.2. La influencia de los resultados en las percepciones sobre la utilidad de la evaluación.	173
Capítulo VII. Conclusiones	175
Anexo I. Referencias bibliográficas y fuentes documentales	187
Anexo II. Datos sobre las entrevistas realizadas	203

RELACIÓN DE ABREVIATURAS

AEDE	Asociación de Editores de Diarios Españoles
AIC	Coalición Agrupaciones Independientes de Canarias
ANECA	Agencia Nacional de Evaluación de la Calidad y Acreditación
ANPE	Asociación Nacional de Profesionales de la Enseñanza
AP	Alianza Popular
AP-PDP	Alianza Popular-Partido Demócrata Popular
AP-PDP-PL (CP)	Alianza Popular-Partido Demócrata Popular (Coalición Popular)
AREX	Acción Regional Extremeña
BUP	Bachillerato Unificado y Polivalente
CC.OO	Comisiones Obreras
CD	Coalición Democrática
CDC	Centro de Desarrollo Curricular
CDC	Convergencia Democrática de Catalunya
CEAPA	Confederación Española de Padres de Alumnos
CECE	Confederación Española de Centros de Enseñanza
CDS	Centro Democrático y Social
CENIDE	Centro Nacional de Investigaciones para el Desarrollo de la Educación
CEP	Centro de Profesores
CG	Coalición Galega
CIDE	Centro de Investigación y Documentación Educativa ¹ Centro Nacional de Investigación, Documentación y Evaluación ²
CiU	Convergencia i Unió
CNREE	Centro Nacional de Recursos para la Educación Especial

¹ Desde su creación, en 1983, hasta 1990.

² Desde 1990 a 1993.

RELACIÓN DE ABREVIATURAS UTILIZADAS

CONCAPA	Confederación Católica de Padres de Alumnos
COU	Curso de Orientación Universitaria
CSIF	Central Sindical Independiente y de Funcionarios
DCB	Diseño Curricular Base
EA-EUE	Eusko Alkartasuna-Euskal Ezquerria
EAT	Equipos de Atención Temprana
EE	Euskadiko Ezquerria
EE.MM.	Equipos Multiprofesionales
ERC	Esquerria Republicana de Catalunya
ERFN	Esquerria Republicana de Catalunya - Front Nacional
EGB	Educación General Básica
ESO	Enseñanza Secundaria Obligatoria
FED-EC	Federació de Esquerria Democràtica-Esquerria Catalana
FERE	Federación Española de Religiosos de la Enseñanza
FETE-UGT	Federación Española de Trabajadores de la Enseñanza- Unión General de Trabajadores
FP	Formación Profesional
FPDL	Federación de Partidos Demócratas y Liberales
FSD	Federación Socialdemócrata
HB	Herri Batasuna
ICE	Instituto de Ciencias de la Educación
IEA	The International Association for the Educational Achievement
INCE	Instituto Nacional de Calidad y Evaluación
INCIE	Instituto Nacional de Ciencias de la Educación
ITE	Instituto de Técnicas Educativas
IU	Izquierda Unida
KEC	Key Evaluation Checklist
LGE	Ley General de Educación y Financiamiento de la Reforma Educativa
LODE	Ley Orgánica Reguladora del Derecho a la Educación
LOECE	Ley Orgánica de Estatuto de los Centros Educativos
LOFAGE	Ley Orgánica de Organización y Funcionamiento de la Administración General del Estado
LOGSE	Ley de Ordenación General del Sistema Educativo
LOPEG	Ley Orgánica de Participación, Evaluación y Gobierno de los Centros
LRU	Ley de Reforma Universitaria
MEC	Ministerio de Educación y Ciencia
MEC	Meta-evaluation Checklist
MRP	Movimientos de Renovación Pedagógica
OCDE	Organización para la Cooperación y el Desarrollo Económico
OTAN	Organización del Tratado Atlántico Norte
PAR	Partido Aragonés Regionalista

RELACIÓN DE ABREVIATURAS UTILIZADAS

PCE	Partido Comunista de España
PDC	Pacte Democràtic de Catalunya
PDC	Partido Demócrata Cristiano
PDP	Partido Demócrata Popular
PGI	Partido Gallego Independiente
PNTIC	Programa Nacional de Nuevas Tecnologías de la Información y la Comunicación
PNV	Partido Nacionalista Vasco
PSA-PA	Partido Socialista Andaluz-Partido Andalucista
PSI	Partido Socialdemócrata Independiente
PSD	Partido Socialdemócrata
PSLA	Partido Social Liberal Andaluz
PSOE	Partido Socialista Obrero Español
PSP-US	Partido Socialista Popular-Unión Socialista
PP	Partido Popular
SITE	Servicio de Inspección Técnica de Educación
SOEV	Servicio de Orientación Escolar y Vocacional
UC	Unión Canaria
UCD	Unión de Centro Democrático
UC-DCC	Unió de Centre-Democràcia Cristiana de Catalunya
UDM	Unión Demócrata de Murcia
UEO	Unión Europea Occidental
UGT	Unión General de Trabajadores
UN	Unión Nacional
UNESCO	Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura
UPC	Unión del Pueblo Canario
UV	Unión Valenciana
WEU	Western European Union

ÍNDICE DE CUADROS

I.1.	Propuestas y <i>checklists</i> para la metaevaluación. Cuadro-resumen.....	42
II.1.	Actores con influencia potencial en los procesos de evaluación, según su naturaleza	56
II.2.	Dimensiones y subdimensiones del esquema de metaevaluación	58
III.1.	Evolución del número de centros por nivel educativo y titularidad (cursos 1979-80 a 1989-90).....	000
III.2.	Evolución del número de becarios y de las inversiones en becas en enseñanzas medias y universitarias (cursos 1982-83 a 1991-92) ..	000
IV.1.	Evolución del alumnado matriculado en BUP y COU y del alumnado que terminó COU (cursos 1983-84 a 1991-92)	92
IV.2.	Evolución del número de alumnos matriculados en el plan experimental, por Comunidades Autónomas (cursos 1986-87 a 1991-92) .	98
IV.3.	Distribución porcentual del alumnado matriculado en el bachillerato experimental, por titularidad de centro (cursos 1984-85 a 1991-92).....	99
IV.4.	Distribución porcentual del alumnado matriculado en todos los centros de educación secundaria y FP, por enseñanza (cursos 1984-85 a 1991-92)	99
IV.5.	Distribución porcentual del alumnado matriculado en los centros públicos de educación secundaria y FP, por enseñanza (cursos 1984-85 a 1991-92)	100
IV.6.	Distribución porcentual del alumnado matriculado en los centros privados de educación secundaria y FP, por enseñanza (cursos 1984-85 a 1991-92)	100
IV.7.	Variables consideradas en la evaluación de la reforma experimental de enseñanzas medias	105

IV.8.	Mediciones <i>pretest</i> y <i>postest</i> realizadas a las generaciones estudiadas	110
IV.9.	Tipos de muestreo empleados en cada uno de los estudios	111
IV.10.	Número de centros experimentales y no experimentales considerados en cada estudio	112
IV.11.	Número de alumnos considerados en cada estudio	112
IV.12.	Instrumentos empleados para la recogida de datos en el 1 ^{er} y 2 ^o estudio de la evaluación	114
IV.13.	Instrumentos empleados para la recogida de datos en el 3 ^{er} estudio de la evaluación	114
IV.14.	Resumen de los resultados encontrados al comparar los alumnos experimentales y de control, según tipo de centro (BUP o FP) y ciclo (primero o segundo)	117
V.1.	Evolución del número de centros, profesores y alumnos del programa de integración (cursos 1985-86 a 1989-90)	125
V.2.	Ejecución presupuestaria del programa de integración entre 1986 y 1990 (en millones de pesetas)	125
V.3.	Variables analizadas en la evaluación (por sistema)	133
V.4.	Características de la muestra seleccionada para la fase cuantitativa (centros)	135
V.5.	Características de la muestra seleccionada para la fase cuantitativa (alumnos)	135
V.6.	Cuadro-resumen de los instrumentos de recogida de datos empleados en la fase cuantitativa	136
V.7.	Relación de variables más relevantes consideradas en las fases cualitativa y de intervención (por sistema)	139
V.8.	Momento de aplicación de las pruebas de la fase cuantitativa	141
VI.1.	Evolución del número de centros participantes y de equipos informáticos en el proyecto <i>Atenea</i> (1985-1990)	155
VI.2.	Coste total del Programa de Nuevas Tecnologías. Proyectos <i>Atenea</i> y <i>Mercurio</i> , 1985-1990 (en millones de pesetas)	156
VI.3.	Costes del proyecto <i>Atenea</i> , 1985-1990 (en millones de pesetas)	156
VI.4.	Valoración por parte de los monitores de la formación recibida.	158
VI.5.	Principales problemas encontrados por los monitores para el desempeño de sus funciones	161
VI.6.	Variables consideradas en la evaluación de progreso del proyecto <i>Atenea</i> e informantes consultados sobre ellas	164
VI.7.	Variables consideradas en la evaluación final del proyecto <i>Atenea</i>	166

INTRODUCCIÓN

La educación está lejos de ser un ámbito *pacífico*. No lo ha sido nunca y parece difícil que alguna vez pueda llegar a serlo. Si hay algún campo en el que la divergencia de valores, ideas y creencias (y, por qué no decirlo, también de intereses) sea la nota dominante, ese es, sin duda, el de la educación. La historia reciente de España proporciona abundantes ejemplos en ese sentido. Así se evidenció, por ejemplo, durante los debates en la ponencia constitucional sobre el artículo 27¹ que consagraba el derecho a la educación, con todo lo que ello implica; más tarde, durante el proceso de elaboración de la fracasada LOECE (Ley Orgánica de Estatuto de los Centros Escolares), la LODE (Ley Orgánica del Derecho a la Educación), la LRU (Ley de Reforma Universitaria), la LOGSE (Ley de Ordenación General del Sistema Educativo), la LOCE (Ley Orgánica de Calidad de la Educación) y la LOE (Ley Orgánica de Educación)², una y otra vez con ocasión de cada política propuesta y de cada ley tramitada. La controversia acerca de los principios que deben guiar las políticas y vertebrar el sistema educativo se reprodujo hace casi un año, con ocasión de las ne-

¹ Que, originalmente, fue tramitado como artículo 26. La divergencia de posturas entre los dos partidos con mayor representación parlamentaria en aquellos momentos, la Unión de Centro Democrático (UCD) y el Partido Socialista Obrero Español (PSOE), se fraguó en torno a dos principios: la libertad de enseñanza –defendida por el primero– y la igualdad de oportunidades –irrenunciable para el segundo–. Uno y otro se erigieron, respectivamente, en representantes de los intereses de los sectores conservadores –las formaciones políticas de derechas, la Conferencia Episcopal, la Confederación Católica Nacional de Padres de Alumnos (CONCAPA) o la Federación Española de Religiosos de la Enseñanza (FERE) y progresistas –los partidos de izquierdas, los movimientos sociales laicos, los sindicatos y las asociaciones de padres que defendían la enseñanza pública–, tanto dentro como fuera del Parlamento. Sobre las dificultades surgidas durante la elaboración del artículo constitucional y las posturas de los distintos sectores, véanse, entre otros, PUELLES (1991), FERNÁNDEZ-MIRANDA (1983), MARAVALL (1984) o SOLANA (1992). En GARCÍA SÁNCHEZ (2003) puede encontrarse un resumen de éste y de otros procesos, como los que afectaron a la LODE y la LOGSE.

² Sobre estas dos últimas pueden consultarse BAJO (2004), FEITO (2003, 2005) y TIANA FERRER (2009)

gociaciones que debían desembocar en la firma del denominado «Pacto por la Educación» y que resultaron infructuosas.

El educativo ha sido y es un ámbito privilegiado para la evaluación y ello por varias razones. En primer lugar, la evaluación de programas es, en gran medida, deudora de la evaluación educativa. Buena parte de los enfoques, modelos, incluso estándares, de los que se sirven hoy los evaluadores (de cualquier parte del mundo y de cualquier campo de actividad) se gestaron en el seno de la evaluación educativa³.

En segundo lugar, porque la evaluación resulta particularmente relevante en un escenario tan proclive a la presencia de posiciones encontradas como el educativo, al permitirnos decidir, sobre la base de evidencias científicas⁴, cuáles son las políticas o los programas educativos que arrojan o pueden arrojar resultados positivos (y, por tanto, han de ser mantenidos o implantados *ex novo*) y cuáles han de ser descartados por razón de los problemas que plantean o podrían llegar a plantear. Más aún, en un escenario fuertemente marcado por la crisis económica, la pertinencia de la evaluación resulta aún más clara, al ser mayor la necesidad de destinar los fondos públicos a aquellas políticas con mayores probabilidades de éxito (léase, aquellas que mejor resuelvan los problemas para los que fueron diseñadas). No se olvide tampoco que la legitimidad de los gobiernos depende hoy, más que nunca, no solo de su eficacia en la gestión sino de la puesta en marcha de mecanismos *reales* de rendición de cuentas. Ambos motivos refuerzan todavía más, si cabe, no ya la conveniencia sino la necesidad de impulsar la evaluación de políticas y programas educativos.

A su vez, todo ello hace de la educación un espacio también propicio para el trabajo metaevaluativo, esto es, para la evaluación de las evaluaciones. La metaevaluación nos permite determinar la «calidad» de las evaluaciones, distinguir las que han sido o son útiles de las que no lo son, contribuyendo, de ese modo, al desarrollo de la práctica evaluativa e, incluso, a fortalecer las bases teóricas en las que dicha práctica se sustenta y al avance y consolidación del campo profesional de la evaluación. Es evidente que la metaevaluación requiere de un cierto grado de progreso (cuantitativo y cualitativo) de la evaluación, que permita realizar análisis concluyentes a partir de los cuales resulte posible conocer el papel que, en la práctica, cumplen o han cumplido las evaluaciones, más allá de lo que se afirme en el plano del discurso.

En España, el desarrollo de la evaluación de políticas y programas de carácter educativo, al menos hasta fechas recientes, ha sido comparativamente mayor que el registrado en otros ámbitos de políticas. Puede incluso afirmarse que

³ Son muchos los trabajos que han dado cuenta de la contribución de la evaluación educativa a la génesis y desarrollo de la evaluación. Los de mayor interés siguen siendo, a nuestro juicio, el clásico de GUBA y LINCOLN (1989) y, más recientemente, el de NEVO (2006).

⁴ Lo que, en modo alguno, supone obviar los valores y los principios, tan importantes, a nuestro entender, como los «datos objetivos».

las ya innumerables experiencias de evaluación habidas en el mundo educativo han servido de acicate y fuente de inspiración en la puesta en marcha de planes evaluativos en otras políticas sectoriales.

Estas circunstancias no solo permitían sino que parecían aconsejar también la realización de un estudio metaevaluativo. Es esta la primera de las contribuciones de la obra que ahora se presenta, habida cuenta de la notable escasez de trabajos empíricos en la materia en el contexto español⁵. La metaevaluación realizada⁶ tuvo como objeto de estudio tres de las evaluaciones de programas de reforma educativa más destacadas y de mayor predicamento de cuantas se han llevado a cabo hasta el momento en España. Fueron, además, las primeras evaluaciones sistemáticas desarrolladas en el ámbito educativo y ello con ocasión de la puesta en marcha de los programas experimentales de reforma que constituyeron el embrión de la LOGSE⁷. Nos referimos a la *evaluación externa de la reforma experimental de las Enseñanzas Medias (1984-1992)*, la *evaluación del programa de integración escolar de alumnos con necesidades educativas especiales (1985-1989)* y la *evaluación del proyecto Atenea para la introducción de la informática en la escuela (1985-1990)*.

El lector no encontrará solo coincidencias entre esta obra y la originaria si bien, como es lógico, los objetivos y los principios que las inspiran son, en lo esencial, muy similares. La revisión crítica de la bibliografía especializada en la materia nos ha permitido actualizar, depurar y complementar las categorías de análisis y el esquema empleados en la primera investigación. Asimismo, recientemente hemos tenido la oportunidad de poder contrastar, en la práctica, la aplicabilidad y la utilidad de algunos de los componentes del esquema en la metaevaluación de las evaluaciones externas y del sistema de evaluación del Consejo de Evaluación del Desarrollo Social del Gobierno del Distrito Federal de México, EVALÚA DF⁸. La reflexión teórica y el trabajo empírico –además,

⁵ Por ejemplo, SANTOS GUERRA (2008), MURCIANO (2004) y BUSTELO (2001), además de GARCÍA SÁNCHEZ (2003).

⁶ La investigación no respondió a un encargo, como suele ser habitual en el campo de la evaluación, sino que se inscribió en un trabajo académico conducente a la elaboración de una tesis doctoral que llevó por título «Evaluación de programas de reforma educativa en España. Tres estudios de caso desde un enfoque de metaevaluación». Fue defendida en la Facultad de Ciencias Políticas y Sociología de la Universidad Complutense de Madrid en 2003.

⁷ La experiencia adquirida en la práctica de la evaluación fue, además, determinante en el proceso de institucionalización de la evaluación educativa. Véase, a este respecto, GARCÍA SÁNCHEZ (2003; 2004)

⁸ Este trabajo se ha desarrollado en el marco de la convocatoria internacional competitiva realizada por EVALÚA DF en 2010. El equipo de metaevaluación ha estado integrado por los profesores OSUNA (director científico), MURCIANO (director de las metaevaluaciones de la política de acceso domiciliario al agua potable y de la política dirigida a jóvenes), BUENO (director de la metaevaluación de la política de fomento del empleo y protección del trabajo) y GARCÍA SÁNCHEZ (directora de la metaevaluación de las políticas de infancia). El esquema aplicado en cada una de las metaevaluaciones se basa en MURCIANO (2007) y GARCÍA SÁNCHEZ (2003;

obviamente, del paso del tiempo— nos han aportado, en suma, una necesaria y enriquecedora distancia con respecto a la primera investigación.

La segunda aportación fundamental de esta obra es la propuesta que en ella se realiza *de un esquema novedoso para la evaluación de las evaluaciones*. Dicha propuesta, construida a partir de diversos trabajos en evaluación educativa, análisis de políticas públicas y Ciencia Política, descansa en el reconocimiento de una doble dimensión en las evaluaciones: la técnica, vinculada a las características más explícitas de las evaluaciones, aquellas relacionadas con la evaluación en tanto que proceso de investigación; y la política, en la que se integran los componentes éticos, ideológicos y cognitivo-simbólicos que están en la base del comportamiento de los actores. Cada una de las dimensiones se asocia, a su vez, a un conjunto de cuestiones o subdimensiones acerca de las cuales los actores toman determinadas decisiones configurando, de este modo, el proceso evaluativo. La aplicación del esquema consistirá, precisamente, en analizar y valorar esas decisiones y su incidencia sobre la evaluación. No se trata, pues, de emplear una lista de comprobación para verificar si la evaluación cumple o no con ciertos requisitos. Así, el esquema que proponemos pretende aproximarse más a la visión comprensiva y orientada a la experiencia de la que habla Stake (2006) que a los métodos más convencionales regidos por criterios y estándares.

La conjugación de distintas perspectivas, creemos, no fortalece solo a la metaevaluación sino también a la Ciencia Política, por cuanto le descubre unos nuevos y sugestivos objetos de estudio —las evaluaciones—, lo que implica, también, enfrentarle con nuevos desafíos. En las evaluaciones se reproduce y, en ocasiones, se intensifica, la dinámica *política* que se da en las políticas públicas; en la evaluación, las divergencias no se dan solo entre quienes tienen una visión diferente sobre el programa sino también entre quienes mantienen posiciones encontradas acerca de cómo este deba juzgarse y de quiénes puedan tomar parte en ese juicio. La evaluación es, pues, un ámbito aún más político, si es que eso es posible, que las propias políticas. De ello se deriva, precisamente, su interés para el politólogo.

* * *

El contenido de la obra se ha organizado en dos partes claramente diferenciadas.

La primera, que corresponde a los capítulos I, II y III, tiene una orientación eminentemente teórica. En el primero de ellos, después de trazar los límites del concepto de metaevaluación, se emprende un recorrido por las propuestas y los enfoques más destacados en busca de referencias útiles para construir un esquema con el que evaluar las evaluaciones. El segundo comienza con una cuestión

2009). Puede consultarse el informe final del proyecto en la dirección <http://www.evaluad.f.gob.mx/transparencia/2011/inf_finmet_09.pdf> .

esencial, como es el carácter político de las evaluaciones, que se aborda ya desde perspectivas propias de la Ciencia Política. A continuación se exponen las premisas de partida, la definición de actor de la que se parte y los factores que explican las decisiones que los actores adoptan, las dimensiones y subdimensiones que integran el esquema de metaevaluación y la relevancia y aplicabilidad del mismo. El tercer capítulo proporciona al lector las claves sobre la política educativa del período 1982-1992, escenario en el que se fraguaron los tres programas experimentales cuyas evaluaciones analizamos en esta obra.

Los capítulos IV, V y VI de la II parte presentan los estudios de caso, esto es, las metaevaluaciones realizadas. Todos ellos –cuya extensión, por razones obvias, es considerablemente mayor que la de los anteriores– siguen una misma lógica analítica y tienen una similar secuencia expositiva. Comienzan con una somera descripción de los programas de reforma a los que se refieren las evaluaciones que constituyen nuestro objeto de estudio. En los siguientes apartados se reconstruyen y analizan los procesos de evaluación (desde el momento en que se concibió la idea de evaluar el programa hasta que se difundieron y/o publicaron los últimos resultados) tomando como referencia cada una de las subdimensiones del esquema metaevaluativo.

Para finalizar, en las conclusiones se hace una recapitulación de los elementos analíticos de los que nos hemos servido y se incorporan las reflexiones surgidas a raíz de la investigación: sobre la relevancia de la metaevaluación como instrumento al servicio de las evaluaciones, las fortalezas del esquema metaevaluativo propuesto y los aprendizajes que entre esta y la Ciencia Política pueden darse, la centralidad del análisis del binomio actores-decisiones en la metaevaluación y las relaciones existentes entre utilidad, pluralismo y viabilidad de la evaluación.

Las referencias bibliográficas se han recogido en el anexo I, mientras que en el II se exponen los datos relativos a las entrevistas realizadas (relación de personas entrevistadas, breve curriculum vitae y fecha de la entrevista) durante el trabajo de campo y que han constituido nuestra fuente principal de información.

PRIMERA PARTE

CAPÍTULO I

LA PLURALIDAD DE CONCEPCIONES DE METAEVALUACIÓN

1. DIBUJANDO LOS CONTORNOS DEL CONCEPTO (DE METAEVALUACIÓN)

Nadie duda hoy de que la evaluación constituye una herramienta esencial para el perfeccionamiento de las políticas ni de su enorme potencial para fomentar la transparencia y la rendición de cuentas por parte de los poderes públicos. Pero poco podrá contribuir la evaluación a la mejora de los procesos decisorios que están en la base de políticas y programas y al fortalecimiento de la *accountability* si no se revisan y perfeccionan los instrumentos de los que se sirve, si no se contrasta la validez de sus juicios, si no se comprueba que resulta útil para conseguir «mejores» políticas. Una evaluación mal planteada –ya sea desde el punto de vista técnico, del de la gestión o del estrictamente político– puede llevar a elegir el programa o la política menos adecuados.

Es precisamente en este punto en el que la metaevaluación, en tanto que evaluación de las evaluaciones, cobra su verdadero sentido. La metaevaluación constituye «una obligación profesional de los evaluadores» (Stufflebeam 2001: 183). Pero la relevancia de la metaevaluación va más allá. Destinatarios de las políticas y clientes¹ tienen en ella una herramienta imprescindible para decidir acerca de la fiabilidad de las evaluaciones y si estas proporcionan una base sólida y rigurosa a partir de la cual se puedan conocer en profundidad las fortalezas y debilidades de una determinada política pública. Incluso, puede afirmarse que la metaevaluación comporta una obligación ética y no solo científica cuando las evaluaciones afectan al bienestar de las personas (Scriven 1991).

La metaevaluación desempeña un papel central en la práctica y en el desarrollo institucional de la evaluación de políticas públicas. Esta constatación, aunque imprescindible, resulta ser solo un primer paso en la aproximación al tema de la me-

¹ Entendidos como quienes encargan las evaluaciones o quienes están al frente de las unidades u organizaciones que las solicitan, sean o no responsables de su financiación.

taevaluación y no ofrece una respuesta a todos los interrogantes que pueden plantearse en torno al mismo: ¿qué es exactamente la metaevaluación?, ¿cuál es o debe ser su objeto?, ¿el proceso evaluativo, sus resultados o ambos?, ¿qué aspectos de una evaluación se deben valorar?, ¿qué criterios son los adecuados para hacerlo?

La primera tarea a la que hemos de enfrentarnos es la de establecer el perímetro, los contornos del concepto *metaevaluación*. Como suele ser habitual en las Ciencias Sociales, es posible encontrar muy distintas y dispares definiciones. Algunas son tan genéricas y tan «de mínimos», que pueden albergar concepciones muy diferentes; es el caso de la ofrecida por Michael Scriven, quien, en 1969, acuñó el término metaevaluación para designar la evaluación de las evaluaciones. Aunque una definición como esta apenas suscitaría rechazo entre los especialistas, parece claro que plantea problemas de aplicación práctica. Otras, en cambio, pretenden ser tan precisas —especificando incluso los criterios de metaevaluación— que apenas si dejan espacio para la existencia de enfoques distintos. La propuesta por Stufflebeam se encuentra en esta línea. Según este autor, la metaevaluación consiste en «definir, obtener y aplicar información descriptiva y valorativa acerca de la utilidad, viabilidad, pertinencia y rigor de una evaluación y sobre si se ha desarrollado de forma sistemática y competente, respetuosa, con integridad y honestidad y responsabilidad social» (2001: 185)².

Una definición algo más clara y, por ello, más útil, es la que se construye por analogía con la más comúnmente aceptada del término *evaluación*³ y que será la que tomemos como referencia en esta obra. Según esta, la *metaevaluación* es

² Veinte años antes, el autor había concebido la metaevaluación como «el proceso de diseño, obtención y empleo de información descriptiva y valorativa acerca de la utilidad, la viabilidad, la ética y el rigor técnico de una evaluación con el propósito de orientarla y dar cuenta públicamente de sus puntos fuertes y de sus puntos débiles» (1981: 151).

³ Según ROSSI y FREEMAN, la investigación evaluativa consiste en «la aplicación sistemática de los procedimientos de la investigación social, para valorizar la conceptualización y el diseño, la ejecución y la utilidad de los programas de intervención social» (1989: 14). Por su parte, PATTON (1997) considera la evaluación de programas como el proceso sistemático de recogida de información acerca de las actividades, las características y los resultados de un programa, de cara a su utilización para reducir incertidumbres, mejorar la efectividad y tomar decisiones sobre el programa. En opinión de STUFFLEBEAM y SHINKFIELD (1993), y en la línea de la definición adoptada por el *Joint Committee on Standards for Educational Evaluation*, la evaluación consiste en el enjuiciamiento sistemático de la valía o el mérito de un objeto. Para CAROL WEISS, la evaluación es «la valoración sistemática del desarrollo y/o los resultados de un programa o una política, en función de un conjunto implícito o explícito de estándares, como un medio de contribuir a la mejora del programa o la política» (1998: 4). En palabras de ALVIRA, la evaluación «conlleva la aplicación de procedimientos sistemáticos y rigurosos de recogida de información (con el objetivo) de emitir un juicio de valor sobre el objeto que se está evaluando». Para este autor, «el énfasis en la utilización de los resultados y en la participación de las llamadas “partes interesadas” a lo largo del proceso evaluativo constituyen otras tantas características distintivas de la metodología de la evaluación» (1991: 7). Como puede apreciarse, todas las definiciones de *evaluación* coinciden en destacar tres elementos: su carácter de investigación, su finalidad valorativa y su orientación aplicada. Sobre las diferencias entre investigación y evaluación pueden consultarse GLASS y WORTHEN (1971-1972), DE LA ORDEN (1985), COHEN y FRANCO (1993) o BUSTELO (1999).

aquella investigación sistemática cuyo objetivo es emitir un juicio acerca de la bondad, el valor y los méritos⁴ (o deméritos) de una evaluación⁵.

Sea cual sea la definición por la que se opte, vemos que la metaevaluación puede centrarse tanto en el proceso como en los resultados de las evaluaciones, en su diseño, en su desarrollo, en la forma en que se presenta y difunde la información producida o en su incidencia (léase utilidad) sobre las políticas o programas. Podemos decidir evaluar una evaluación una vez que esta ya ha concluido, adoptando de este modo una perspectiva sumativa (Scriven 1991) o retrospectiva (Stufflebeam y Shinkfield 2007). Estaríamos, en este último caso, ante lo que Schwandt y Halpern (1988) denominan «retroevaluación» o ante una evaluación *ex post facto* en palabras de Bustelo (2002). Es igualmente posible hacer que la metaevaluación se adelante al proceso evaluativo, en un claro intento de guiar al evaluador en la etapa de planificación del trabajo antes de que este comience a desarrollarse. Hablaríamos entonces de evaluación formativa (Scriven 1991), *ex ante* (Bustelo 2002), proactiva (Stufflebeam y Shinkfield 2007) o de «proevaluación» (Schwandt y Halpern 1988). De la misma forma, el juicio a la evaluación puede abordarse desde dentro, en el seno del equipo u organización que llevó a cabo la experiencia evaluativa (metaevaluación interna) o desde fuera, aportando un punto de vista externo que mantenga la distancia (al menos, teóricamente) sobre lo realizado (metaevaluación externa).

Más aún, y en relación con el objeto que puedan tener las metaevaluaciones y el papel que puedan desempeñar, hemos de destacar la oportunidad que representa la metaevaluación para comprender y enjuiciar los mecanismos y los resultados de la función de evaluación en un contexto político determinado (Nilsson y Hogben 1983; Bustelo 2001; 2002; García Sánchez 2003). La mirada de la metaevaluación no se ha de dirigir solo a evaluaciones concretas sino que puede y debe extenderse a las *políticas de evaluación* –a cómo esta es con-

⁴ Entre los expertos de habla inglesa suele ser habitual distinguir entre *merit*, término que se refiere al valor intrínseco que tiene un determinado objeto (ya sea un programa o una evaluación) y *worth*, con el que se designan aquellas características del objeto que le hacen tener un valor «de mercado» (no necesariamente económico) o para un propósito determinado. En castellano, las diferencias entre los dos términos, *valor* y *mérito*, aunque existen, son menos claras. Según la Real Academia de la Lengua, *mérito* es aquello que hace que tengan valor las cosas mientras que algunas de las acepciones de *valor* tienen una connotación más instrumental, en tanto se refieren al «grado de utilidad o aptitud de las cosas» o a la «cualidad de las cosas, en virtud de la cual se da por poseerlas cierta suma de dinero o equivalente». Véase Real Academia Española de la Lengua (2001), *Diccionario de la Lengua Española*, 22ª edición, Madrid. En línea: <http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=m%E9rito> (consulta: 30 enero de 2011).

⁵ Esta obra se centra exclusivamente en la evaluación y metaevaluación de programas y proyectos, aunque algunas de las reflexiones que se realizan podrían ser también aplicables a otros tipos de evaluación, como la de servicios, la del rendimiento o la de centros. Hemos de señalar también que, aun siendo conscientes de las diferencias que pueden existir entre los términos *programa* y *proyecto*, los tomaremos como equivalentes ya que, en la práctica, las evaluaciones en uno y otro caso se desarrollan de manera similar. Más allá de esta consideración concreta, no debe olvidarse que la naturaleza del objeto evaluado condiciona sustancialmente el proceso de evaluación.

cebida por los poderes públicos, de qué forma se articula el sistema de evaluación, qué medidas se adoptan para que las evaluaciones contribuyan a la planificación o qué control se pretende ejercer sobre la información producida—. A través de los estudios metaevaluativos podemos, pues, comprender el comportamiento que, en materia de evaluación, tienen los poderes públicos⁶.

Una vez establecidos los límites del concepto y aclarado qué aspectos o fases del proceso de evaluación se pueden evaluar, en qué momento y si ello puede hacerse adoptando una perspectiva interna o externa, procede ahora que nos ocupemos de los siguientes interrogantes, probablemente los de mayor trascendencia para la metaevaluación: ¿sobre qué base se pueden determinar la bondad, el valor o los méritos de una evaluación?, ¿cuáles son los criterios conforme a los cuales se pueden o deben evaluar las evaluaciones?, ¿hemos de recurrir a criterios o a estándares?⁷, en otras palabras, ¿necesitamos propuestas prescriptivas o analítico-descriptivas? o ¿debe el metaevaluador ocuparse de aquellos aspectos de la evaluación que no son estrictamente técnicos?

Con el fin de encontrar una respuesta a estas cuestiones, recorreremos el camino que, desde finales de los años 60⁸, han ido trazando los principales enfoques metaevaluativos.

2. EN BUSCA DE CRITERIOS PARA EVALUAR LAS EVALUACIONES. UN RECORRIDO POR LOS PRINCIPALES ENFOQUES META-EVALUATIVOS

Al igual que sucede en el caso de la evaluación, en el campo de la metaevaluación encontramos enfoques positivistas en los que la atención se dirige principal-

⁶ Trasladamos a las evaluaciones la conocida afirmación de ASHFORD (1976), para quien las políticas públicas vendrían a ser una forma de descripción del comportamiento institucional.

⁷ Aunque, en ocasiones, «criterios» y «estándares» son empleados como si se tratara de sinónimos, en esta obra subrayamos las diferencias existentes entre ellos. Así, denominamos criterio de valor a aquel aspecto o componente (del objeto evaluado) cuyo análisis se entiende imprescindible en la evaluación del mismo. Un estándar no es sino el nivel satisfactorio de rendimiento o cumplimiento de un determinado criterio. Véanse WHEELER, HAERTEL y SCRIVEN (1992).

⁸ Con el propósito de mantenernos fieles al tema central de esta obra, nos ocuparemos tan solo (con algunas excepciones) de las propuestas formuladas para la metaevaluación de políticas y programas, aunque somos conscientes de que los modelos y enfoques evaluativos pueden constituir una referencia útil para identificar los criterios con los que evaluar las evaluaciones. A su vez, y dentro de dichas propuestas, nos detendremos únicamente en aquellas que tienen vocación de generalidad (es decir, que han sido diseñadas no solo para un estudio de metaevaluación concreto) y, por razones obvias, en las que en España y en América Latina cuentan con un mayor reconocimiento. Algunas de estas no han sido consideradas, bien por centrarse no en el proceso de evaluación en su conjunto sino en la calidad y utilidad de los informes de evaluación, como es el caso de la realizada por LARRÚ (2003) sobre el sistema de evaluaciones de Europeaid, bien por ocuparse no de la evaluación de un programa sino de un sistema de evaluación de titulaciones universitarias, como la de FERNÁNDEZ-RAMÍREZ *et al.* (2010).

mente hacia los aspectos técnico-metodológicos de las evaluaciones, perspectivas cualitativas, que tienden a poner el acento en las cuestiones éticas y en el carácter intrínsecamente político de los procesos de evaluación⁹ como también propuestas que cabría calificar de eclécticas y en las que tanto métodos como valores tienen cabida aunque en un grado variable dependiendo de los autores¹⁰.

El protagonismo de estos enfoques y propuestas ha ido variando con el tiempo. El desarrollo de la metaevaluación ha seguido, en este sentido, un curso paralelo al de la evaluación¹¹. Así, la superación de la primacía del paradigma positivista¹² y el avance hacia la *cuarta generación*, por emplear la expresión de Guba y Lincoln (1989), se han dado también en el campo de la metaevaluación, si bien con mucha menor intensidad que en el caso de la evaluación. Ello no impide obviamente que, en las épocas marcadas por el cuantitativismo, se puedan encontrar propuestas alejadas del mismo o que, en la actualidad, en plena *cuarta generación*, podamos descubrir trabajos en los que se hace especial hincapié en lo técnico-metodológico.

2.1. El cuantitativismo en metaevaluación: la preocupación por la calidad técnica de las evaluaciones

Como es de sobra conocido, en los momentos de auge del paradigma positivista o en aquellos casos en que la propuesta sobre metaevaluación se sitúa en esa perspectiva, los aspectos metodológicos adquieren todo el protagonismo. Objetividad, validez y fiabilidad de los datos se convierten en los estándares con respecto a los cuales se juzga el trabajo de evaluación. Hay quienes incluso van más allá al equiparar uso de técnicas cuantitativas y «calidad» en evaluación y al considerar que esta solo queda plenamente garantizada con el empleo de procedimientos cuantitativos para la recogida y tratamiento de datos en el

⁹ Baste apuntar aquí, y aun a riesgo de adelantar algunas de las claves del esquema de metaevaluación que propondremos en el siguiente capítulo, que ello se deriva de la presencia en los procesos de evaluación de intereses contrapuestos que pugnan por conseguir o incrementar su cuota de poder en la toma de decisiones.

¹⁰ BUSTELO (2002) distingue entre enfoques que conciben la metaevaluación como control de la calidad de las evaluaciones y enfoques en los que aquella se entiende, esencialmente, como el proceso de descripción, análisis y valoración de los procesos de evaluación. Entre los primeros, la autora sitúa a SCHWANDT y HALPERN (1988), SCRIVEN (1991; 1993) y PATTON (1997) y entre los segundos, a VEDUNG (1997). A nuestro juicio, el propio trabajo de BUSTELO podría incluirse en esta última categoría.

¹¹ La obra de GUBA y LINCOLN (1989) sigue siendo una de las mejores referencias para conocer y comprender las transformaciones habidas en la evaluación de políticas públicas. Son de interés también CHRISTIE y ALKIN (2004), STUFFLEBEAM y SHINKFIELD (1993), MADAUS, SCRIVEN y STUFFLEBEAM (1983) y ESCUDERO (2003).

¹² Sobre las premisas en las que descansa este paradigma véase HAMMERSLEY (1995: 2 y ss.).

marco de un diseño de naturaleza experimental¹³. La metaevaluación tendría, pues, como objetivo calibrar en qué medida el estudio de evaluación se aproxima a ese «modelo».

El trabajo de metaevaluación de Bernstein y Freeman (1975) constituye un buen ejemplo de lo que los enfoques cientifistas consideran como «calidad» en la investigación evaluativa. Estos autores desagregaron el concepto de calidad en una serie de dimensiones –tipo de muestreo, análisis de datos, procedimientos estadísticos y diseño de la investigación– a las cuales atribuyeron valor en función de una escala en la que la máxima valoración correspondía al nivel máximo de calidad. Desde su perspectiva, solo los diseños experimentales o cuasi-experimentales, el muestreo aleatorio, las técnicas cuantitativas de análisis de datos y los procedimientos multivariantes merecen la máxima calificación.

Una de las primeras propuestas de Daniel Stufflebeam en materia de metaevaluación se sitúa también entre las que otorgan un destacado papel al método y a los métodos, si bien estos no tienen la centralidad que se les concede en la obra de Bernstein y Freeman antes comentada.

Tomando como punto de partida los criterios que son comúnmente aceptados para juzgar las investigaciones científicas, Stufflebeam (1974) identifica tres elementos que han de tenerse en cuenta a la hora de evaluar una evaluación: idoneidad técnica, utilidad y coste-eficacia, los dos primeros desagregados ya en algunos trabajos previos de este y otros autores¹⁴. Así, el listado queda compuesto por los siguientes criterios: validez interna, validez externa, fiabilidad, objetividad, relevancia, importancia de los datos recogidos, alcance de la información, credibilidad, oportunidad temporal, grado de difusión de la evaluación y coste-eficacia.

Quizá lo más sugerente y destacable de la propuesta de Stufflebeam es que apunta no solo los requerimientos a los que se debe ajustar la evaluación sino también los problemas a los que el evaluador puede tener que enfrentarse en el transcurso de su trabajo y que pueden comprometer el cumplimiento de dichos requerimientos. Stufflebeam considera que existen seis grandes «categorías» de dificultades: las conceptuales, las de carácter sociopolítico, las legales (o relacionadas con el contrato), las técnicas, las de tipo administrativo y las ético-morales. Siquiera de forma indirecta, los factores no estrictamente técnicos o

¹³ Recordemos que la finalidad de los diseños experimentales es determinar si los efectos producidos a raíz de la implantación de un programa son o no atribuibles al mismo. El experimento consiste en la formación, mediante procedimientos aleatorios, de uno o varios grupos de intervención (los que se someten al programa) y uno o varios de control (aquellos que no son beneficiarios) que son equivalentes a los primeros en sus características básicas. En los diseños cuasi-experimentales los grupos de intervención y de control no son equivalentes, puesto que para su formación no resulta posible aplicar procedimientos aleatorios. Para mayor abundamiento véanse CAMPBELL y STANLEY (1973) y COOK y CAMPBELL (1979).

¹⁴ Véanse GUBA y STUFFLEBEAM (1970) y STUFFLEBEAM *et al.* (1971).

metodológicos se abren paso en las reflexiones más tempranas de Stufflebeam sobre metaevaluación.

Las cuestiones de carácter técnico y metodológico adquieren también un claro protagonismo en la obra de Schwandt y Halpern (1988). Para estos autores, la metaevaluación, como metáfora de la auditoría fiscal, ha de servir para mejorar la calidad de las evaluaciones. Más concretamente, la metaevaluación debe, en primer término, comprobar si la metodología de la evaluación es o no la apropiada y, en segundo término, si la información producida es correcta. Para ello, ha de ser relevante, fiable, suficiente, representativa y haber sido recabada a tiempo. Solo si se cumplen estos criterios, las interpretaciones, juicios y recomendaciones serán adecuados.

2.2. La metaevaluación desde una perspectiva cualitativista

Los autores que se inscriben en los enfoques preocupados por los valores perciben la evaluación de forma diametralmente opuesta. En líneas generales, sus aportaciones tienden (i) a subrayar la importancia de la interpretación y comprensión de los hechos y los fenómenos sociales desde la perspectiva (subjetiva) de los actores-sujetos (ii) a poner el énfasis en las cuestiones éticas y en el carácter intrínsecamente político de las evaluaciones y (iii) a criticar la excesiva simplicidad de las argumentaciones positivistas que identifican «calidad» con utilización de técnicas cuantitativas (si bien es cierto que los propios cualitativistas¹⁵ suelen decantarse por los métodos de naturaleza cualitativa). Como es obvio, esta particular visión de los procesos evaluativos se refleja en las propuestas de metaevaluación que realizan.

«Diálogo» y «transparencia» constituyen los principios clave sobre los que descansa la reflexión de Uusikylä y Virtanen (2000). Sostienen estos autores que la bondad y el mérito de las evaluaciones han de ser determinados por todos los *stakeholders*¹⁶. Ciertas decisiones –como qué tipo de información se requiere o qué criterios han de emplearse para valorar la calidad de una evaluación–

¹⁵ En ocasiones, los enfoques que se sustentan en estas premisas suelen recibir la denominación de *cualitativistas*, *pluralistas* o, incluso, *naturalistas*, aunque, en puridad, el término *naturalismo* hace referencia al modelo de evaluación propuesto por GUBA y LINCOLN (1981).

¹⁶ Optamos por emplear el término en inglés porque su uso está ampliamente difundido entre los especialistas en evaluación. Dicho término, según recuerda MENDELLOW (1987, citado por PATTON 1997: 41), procede del mundo de la consultoría y se refiere a las personas que no poseen acciones o intereses directos en una compañía pero sin cuyo apoyo esta dejaría de existir. Según la definición del *Joint Committee on Standards for Educational Evaluation* (1994: 25) se considerarán *stakeholders* todos aquellos actores que están implicados en un proceso de evaluación y/o pueden afectar o verse afectados por el mismo. El concepto de *audiencias*, tal y como lo define SCRIVEN (1991), es bastante más amplio: en él se incluyen no solo quienes tienen un interés directo en la evaluación sino también aquellos que tienen un interés indirecto y que, pudiendo estar interesados en el proceso evaluativo, no pueden ni influir en el mismo ni verse influidos por él.

no pueden recaer en un único actor. La metaevaluación se concibe, pues, «como un proceso que propicia el diálogo abierto y el juicio colectivo sobre la utilización de los resultados de la evaluación y, de este modo, contribuye al aprendizaje organizativo mediante la investigación evaluativa» (2000: 52).

Pero, sin duda, es en el modelo de evaluación democrática y deliberativa de House y Howe (1999; 2000) donde la dimensión política de las evaluaciones adquiere mayor protagonismo. Este modelo es, en gran medida, deudor de las propuestas de House, uno de los autores que más ha insistido en la necesidad de que la evaluación sea una experiencia pluralista y de que se atenga a requerimientos de carácter ético¹⁷.

Inicialmente concebido para la evaluación, la versatilidad del modelo de House y Howe lo ha convertido en una magnífica guía para el metaevaluador, además de ser un auténtico referente para todos aquellos que se interesan por la dimensión ético-política de las evaluaciones. El modelo pretende facilitar la obtención de conclusiones válidas, sin sesgos, allí donde existan opiniones encontradas acerca de un programa. Para ello, las evaluaciones –o, en su caso, las metaevaluaciones– han de seguir tres principios o criterios básicos: la inclusión, el diálogo y la deliberación. En primer lugar, se entiende que la evaluación ha de *incluir* los intereses, valores y opiniones de los principales *stakeholders* implicados en la política o el programa, aunque no todos tengan idéntica participación en el proceso evaluativo. La evaluación, en segundo término, ha de propiciar el *diálogo* con y entre los *stakeholders*¹⁸, lo que no implica necesariamente que estos lleguen al final a un acuerdo. El último de los criterios alude al proceso de *deliberación* que ha de desarrollarse en las evaluaciones con el fin de que puedan extraerse conclusiones válidas y fundamentadas acerca de la política pública que se ha sometido a evaluación.

No podemos dejar de mencionar aquí la obra más reciente de Robert Stake¹⁹, una de las más influyentes y sugestivas de los últimos años. Concebida y desarrollada como una reflexión teórica sobre la evaluación, su tesis central es que en esta existen dos grandes perspectivas: la basada en mediciones, criterios y estándares y la que descansa en la experiencia, la interpretación y la comprensión de los hechos²⁰. Las premisas y el tipo de razonamiento de estas perspectivas son bien distintos. En el primer caso, en la *evaluación basada en estándares*

¹⁷ Ello se advierte tanto en sus principales reflexiones teóricas sobre evaluación (1980; 1993) como en sus trabajos de metaevaluación, entre los que destaca su crítica de la evaluación del programa *Follow Through* (HOUSE *et al.* 1977).

¹⁸ Por su parte, GREENE (2001) y RYAN y DE STEFANO (2001) han insistido en la idea del diálogo como fuerza y foro para la democratización. Para estos autores, el diálogo en la evaluación constituye un modo de mejorar la calidad democrática de los procesos decisorios al propiciar la implicación de las distintas perspectivas existentes sobre un programa.

¹⁹ Nos referimos a la publicada en inglés en 2003 y traducida al castellano en 2006.

²⁰ Estas dos formas de entender y abordar la evaluación vendrían a ser equivalentes, en términos generales, al positivismo y al cualitativismo de los que venimos hablando en esta obra.

res, se parte de la creencia de que el evaluando (políticas, programas) tiene un valor intrínseco, independiente de las percepciones de los actores y cuya determinación precisa (tan solo) de instrumentos objetivos como son los indicadores. En la denominada *evaluación comprensiva*, por el contrario, se considera que la valía del evaluando no es un rasgo inherente al mismo que está esperando a ser descubierta por el evaluador sino que se construye a través de los juicios y percepciones de las personas implicadas. El instrumental al que se recurre y el tipo de análisis que se realiza serán los habituales para este tipo de datos.

Como la evaluación, la metaevaluación –«una ética de cuestionamiento continuo»²¹– puede también basarse en estándares o ser *comprensiva*. En el primer caso, el objetivo es comprobar si, en el transcurso de la evaluación, los evaluadores se han atenido o no a una serie de estándares. Las propuestas de Daniel Stufflebeam y Michael Scriven se inscribirían, según Stake, en esta categoría. En el segundo caso, esto es, cuando se sigue una lógica comprensiva, la metaevaluación gira en torno a las personas, a sus experiencias y a los hechos del proceso evaluativo; el objetivo último, como apunta Murciano (2007), es comprender las decisiones e interpretaciones de los evaluadores teniendo en cuenta el contexto en el que estos se desenvuelven. *Quieting Reform*, la metaevaluación del programa *Cities in Schools* publicada por Stake en 1986, constituye un buen ejemplo de cómo el autor entiende la evaluación y la calidad en evaluación.

2.3. Las propuestas conciliadoras²²

La superación de la primacía del cuantitativismo primero y de la polémica entre este y el cualitativismo después ha permitido, con el paso del tiempo, dibujar un escenario en el que se advierten, mayoritariamente, las ventajas e inconvenientes de ambos paradigmas²³. Se reconoce hoy que «no hay estudios “perfectos”. Y no puede haberlos porque no hay acuerdo sobre qué es la perfección» (Patton 1986: 220). En la actualidad, se admite pues que el evaluador debe conocer y dominar todo tipo de técnicas y que la elección de unas u otras debe hacerse en función de la naturaleza del objeto de estudio y de los objetivos

²¹ STAKE (2006: 262).

²² No nos ocuparemos aquí de otras propuestas eclécticas realizadas por instancias internacionales como el Banco Mundial (2007) o la Organización para la Cooperación y el Desarrollo Económico OCDE (2010) ya que no han sido específicamente elaboradas para la metaevaluación.

²³ Sostiene BERICAT que los cuantitativistas, a pesar de la revalorización de la perspectiva cualitativa en los años setenta, siguieron manteniendo una actitud arrogante. Sabedores de su primacía, observaron «casi con indiferencia» las críticas que les dirigían los cualitativistas. (Los cuantitativistas) «seguros de sí mismos, siguieron realizando su labor, perfeccionando sus técnicas, sin apenas levantar la cabeza e, incluso hasta muy recientemente, sin apenas volver la vista con la intención de asimilar los beneficios que la emergencia del cualitativismo pudiera reportarles» (1998: 27).

de la investigación²⁴. De la misma forma, se asume la dificultad de hallar unos criterios universalmente válidos que guíen al metaevaluador en su análisis crítico de la evaluación. Así, se tiende a aceptar que dichos criterios habrán de ser adoptados en cada caso, dependiendo de las características del contexto²⁵.

En gran medida, este clima de pugna entre paradigmas es el responsable del surgimiento de aquellas contribuciones en el campo de la metaevaluación a las que hemos denominado eclécticas y que, sin olvidar la dimensión técnica, pretenden atender también a la más propiamente ético-política de los procesos de evaluación.

Las reflexiones de Michael Scriven (1969; 1991; 2007) constituyen aquí una referencia inexcusable. Ya en sus primeros escritos sobre metaevaluación se advierte una clara preocupación por incluir, entre los criterios para enjuiciar la evaluación, tanto las cuestiones relacionadas con el rigor de la investigación evaluativa y la utilidad de la información como aquellas otras vinculadas a la ética del proceso de evaluación. A juicio de Scriven, el metaevaluador puede y debe servirse no solo de aquellos estándares que han sido específicamente formulados para la metaevaluación sino también de aquellos otros que han sido concebidos para la realización de evaluaciones. De esta consideración surgieron sus dos propuestas: *la Meta-Evaluation Checklist* y *la Key Evaluation Checklist*.

La *Meta-Evaluation Checklist* (MEC), incluye una serie de ítems, que el autor obtiene de diversas fuentes, sobre el deber ser del estudio evaluativo. En la MEC se considera que la evaluación ha de explicitar y justificar los estándares de valor utilizados y ser comprensiva, comprensible, creíble, ética, factible, suficientemente precisa desde el punto de vista técnico –pero sin que ello haga olvidar cuestiones como la relevancia de la información proporcionada–, «sensibles» desde el punto de vista político –esto es, que atienda y responda a las necesidades de las distintas audiencias–, oportuna, segura, válida y rentable en términos de coste-eficacia.

La *Key Evaluation Checklist* (KEC), en principio elaborada como referencia para la realización de evaluaciones, puede también resultar de utilidad al metaevaluador. La KEC, que el autor ha revisado y actualizado recientemente (2007), no debe ser considerada como una lista de *desiderata* sino más bien como un catálogo de elementos a los que habría que prestar atención en un

²⁴ PATTON señala que la utilización conjunta de distinto tipo de técnicas ha producido resultados muy interesantes en el campo de la evaluación, como muestran sus propias aportaciones (1982) o las de autores como FETTERMAN (1984), REICHARDT y COOK (1979), MAXWELL *et al.* (1985), FIRESTONE y HERRIOT (1984) o, más recientemente, CARACELLI y GREENE (1997) o GREENE, BENJAMIN y GOODYEAR (2001; 2002).

²⁵ Pese a ello, PATTON (1986) considera que los métodos de la evaluación pueden ser juzgados según unos criterios de referencia de carácter general tales como su conveniencia, su utilidad, su grado de factibilidad, su credibilidad y su relevancia.

proceso metaevaluativo. El catálogo se estructura en cuatro apartados²⁶. Bajo la rúbrica «fundamentos», se recogen las cuestiones relativas a los antecedentes y al contexto de la evaluación, la descripción del programa y la definición de los términos empleados, la identificación de los usuarios (*impactees*), los costes y recursos y los criterios de valor. El siguiente apartado se dedica a la «subevaluación» de determinados aspectos, como son el propio proceso del programa, los resultados, los costes, la comparación con las alternativas existentes al programa evaluado y la posibilidad de generalización de los resultados de la evaluación. En la última parte, referida a las «conclusiones e impactos», se realiza una breve síntesis de lo expuesto en apartados anteriores, se recogen las explicaciones sobre los éxitos y fracasos del programa y las recomendaciones que se entienden pertinentes, la atribución de responsabilidades y la justificación correspondiente, la difusión de la información producida y, finalmente, la meta-evaluación.

Otra de las propuestas que podríamos considerar ecléctica y que ha gozado y goza de gran predicamento en evaluación es la realizada por el *Joint Committee on Standards for Educational Evaluation*.

En 1974, un comité nombrado por la *American Educational Research Association*, la *American Psychological Association* y el *National Council on Measurement on Education* y presidido por D.L. Stufflebeam, sugirió a estas organizaciones la elaboración de un conjunto de estándares con los cuales valorar la práctica de la evaluación educativa. El trabajo del comité –al cual se sumaron después nueve asociaciones profesionales más– fue arduo y complejo, debido a las importantes divergencias existentes entre sus miembros acerca de cuestiones como la utilización de tests estandarizados o la inclusión de estándares relativos al personal o a organizaciones de evaluación²⁷. El *Joint Committee* recogió los resultados de sus reuniones en la obra *Estándares para la evaluación de programas, proyectos y materiales educativos*, publicada en 1981. Estos estándares fueron sometidos a un proceso de contrastación por parte de especialistas en el campo de la evaluación y de las Ciencias Sociales, así como por profesionales de la educación, antes de que su utilización fuera recomendada.

En 1989, al tiempo que se procedía a la acreditación oficial de los estándares por el *American National Standards Institute*, se comenzó a percibir la necesidad de revisar su contenido y de ampliar su ámbito de aplicación a programas médicos, asistenciales, militares o sociales. Se llegó así a una nueva y más precisa formulación de los estándares, en la que estos fueron considerados como aspectos sustantivos de los cuatro rasgos característicos de la evaluación: utili-

²⁶ En el primero de ellos, titulado «cuestiones preliminares», se incluyen el informe ejecutivo, la introducción y la metodología de la evaluación. En puridad, sería más un epígrafe del informe de evaluación que un bloque de cuestiones que deban abordarse durante el proceso evaluativo.

²⁷ En 1988, el *Joint Committee* procedió a la publicación de un catálogo de estándares destinado a la evaluación de personal, *The Personnel Evaluation Standards*.

dad, viabilidad, rigor y adecuación o respeto a las normas²⁸. En 2011 se ha publicado la tercera edición de estos estándares²⁹, una vez concluido el proceso de revisión iniciado siete años atrás y en el que tomaron parte más de 400 *stakeholders*. La preocupación por incorporar las perspectivas y los intereses políticos de los distintos *stakeholders*, por los derechos humanos o por la honestidad del proceso de evaluación siguen ocupando, en la edición de 2011, un espacio muy destacado en la propuesta del *Joint Committee*, siquiera como subestándares. La versión de 2011 incluye un estándar adicional referido a la rendición de cuentas o a la transparencia de la evaluación (*evaluation accountability*) que se desglosa en tres subestándares, dos de los cuales giran en torno a la metaevaluación. Se considera que se deben desarrollar procesos internos de metaevaluación en los cuales, sobre la base de los estándares del *Joint Committee* o de otros que sean aplicables, se examine si la evaluación da cuenta (o justifica) de cuál es su diseño, de los procedimientos que ha seguido, de la información recogida y de los resultados obtenidos. Asimismo, se entiende que patrocinadores, clientes, evaluadores y otros *stakeholders* han de propiciar la realización de metaevaluaciones externas.

La denominada *Program Evaluations Metaevaluation Checklist* de Stufflebeam (1999) constituye una adaptación al ámbito metaevaluativo de los estándares de evaluación de programas del *Joint Committee* en su segunda edición: utilidad, viabilidad, adecuación y respeto a las normas y rigor. Su finalidad es servir de guía para verificar si un proceso evaluativo se atiene a estos estándares, que son considerados el paradigma de la calidad, el *deber ser* de la evaluación. Cada uno de los estándares se divide, a su vez, en una serie de subestándares –cuyo cumplimiento se valora a partir de una escala que va del 0-1 –insatisfactorio– al 6 –excelente–. Al igual que sucede en el caso de los estándares del *Joint Committee*, esta propuesta abre la puerta a las cuestiones ético-políticas incluyéndolas, también en este caso, como subestándares.

En una línea argumentativa similar se encuentra la denominada *Guiding Principles Checklist for Evaluating Evaluations* de Stufflebeam, Goodyear, Marquart y Johnson (2005), elaborada a partir de los *Guiding Principles for Evaluators* de la *American Evaluation Association* (2004)³⁰. No obstante, el peso que se otorga aquí a los elementos de naturaleza ético-política es mayor que en las propuestas anteriores: tres de los cinco grandes bloques de principios –investigación sistemática, competencia, integridad / honestidad, respeto por

²⁸ *Utility, feasibility, accuracy y propriety*, respectivamente, en inglés. Véase *Joint Committee on Standards for Educational Evaluation* (1994).

²⁹ YARBROUGH *et al.* (2011).

³⁰ La primera versión de estos *Guiding Principles* data de 1994, cuando la Junta de la *American Evaluation Association* (AEA) constituyó un grupo de trabajo con el propósito de elaborar un conjunto principios orientadores para la evaluación. En 2002-2003 se constituyó un nuevo equipo de trabajo que revisó la primera versión y cuyo informe, una vez aprobado por los miembros de la AEA, dio lugar a los *Guiding Principles* que fueron hechos públicos en 2004.

las personas y responsabilidad por el bienestar general y el de la sociedad— se refieren a dichos elementos, algo que también ocurre en el caso de los *Guiding Principles* en los que se basa.

En la metaevaluación de las políticas de igualdad de género de los gobiernos central y autonómicos de España realizada por Bustelo (2002) tampoco encontramos criterios que aludan explícitamente a la política o a la ética de las evaluaciones. La autora agrupa los criterios de metaevaluación en tres grandes categorías referidas a (i) cómo se plantean las evaluaciones y cuáles son las estrategias evaluativas empleadas, (ii) el papel de algunos elementos clave en la evaluación y (iii) la utilización de las evaluaciones. En la primera se encuadran criterios tales como el contexto evaluativo, la claridad o ambigüedad en los propósitos y las funciones de las evaluaciones, las estructuras institucionales de la evaluación, los tipos de evaluación por los que se opta y los recursos empleados en los procesos de evaluación. En la segunda se incluyen el estudio de agentes y participantes, los momentos y los tiempos, los criterios e indicadores, así como los procedimientos e instrumentos de recopilación y análisis de datos utilizados. La tercera y última engloba el examen de la pertinencia y utilidad de la información producida, la comunicación de los resultados y el impacto de los resultados de la evaluación en la política pública. Sin embargo, pese a no aparecer de forma expresa en ningún criterio, la política —en especial en su dimensión de estructura o *polity* y de resultado o *policy*— está constantemente presente en su esquema de análisis y en la aplicación de este en el trabajo empírico³¹.

El trabajo de metaevaluación de Cardozo Brum (2006) sobre los programas de desarrollo social en México se sustenta sobre una visión de la evaluación que cabría calificar de integral. La premisa de partida es que «la comprensión y el perfeccionamiento de la evaluación debe llevarnos a enjuiciar el valor y el mérito no solo de los trabajos de los evaluadores profesionales sino de todo el sistema en que se organiza y desarrolla el anterior, incluyendo el papel jugado por otros actores o instancias que inciden en la calidad del proceso desarrollado y del informe final logrado» (2006: 138). La lista de comprobación que propone la autora integra elementos procedentes de distintos enfoques de metaevaluación, con la lógica adaptación a la realidad social y política mexicana. Los elementos se organizan en dos grandes categorías: la referida al contexto en el que se desarrollan las evaluaciones y la relativa a la metodología y los procesos internos de la evaluación. En la primera se integrarían cuestiones como la planificación de las evaluaciones, el origen de la decisión de evaluar, las instancias intervinientes (y el papel formal que les corresponde), el funcionamiento real del sistema (relaciones entre instancias, licitaciones y términos de referencia, mercado de la evaluación, cumplimiento de obligaciones legales, equipos eva-

³¹ Por ejemplo, en la estructura institucional o en el impacto de la evaluación sobre la política pública.

luadores y sus motivaciones, problemas administrativos), los recursos para evaluar, la difusión de los resultados de las evaluaciones o la utilización de los resultados en la adopción de decisiones. La segunda categoría englobaría aspectos como las fuentes de información y análisis de datos, la pertinencia y el diseño del programa, los recursos utilizados, el análisis de procesos, la medición de la cobertura, los resultados alcanzados, los impactos, la participación y la satisfacción social, la utilización de resultados y recomendaciones, los informes finales, el coste de la evaluación y los escenarios prospectivos. Como puede advertirse, aunque la propuesta no descuida la importancia que los actores (y las relaciones que se dan entre ellos) tienen a la hora de conformar los procesos de evaluación, la política se sitúa en ella en un plano de igualdad con respecto a otros aspectos de la evaluación.

El mundo de las valoraciones, los intereses y las actitudes adquiere un destacado protagonismo en las reflexiones de Santos Guerra (1999) y Santos Guerra y Moreno Olivos (2004). Los autores distinguen seis planos o contextos, por expresarlo con sus palabras, en la evaluación, cada uno de los cuales corresponde a un conjunto de elementos que resultan críticos a la hora de evaluar las evaluaciones. El *contexto de constitución* se refiere esencialmente a la génesis de la evaluación y a los procesos de negociación que se desarrollen en ella. El *contexto de justificación* englobaría todo lo relativo a los métodos empleados para la obtención de datos, su aplicación, la información resultante y el tratamiento que se dé a la misma. Con el *contexto de presentación* se alude al contenido y forma de presentación del informe, su estructura lógica y su coherencia interna, su adecuación a las necesidades de las distintas audiencias, la extensión o la claridad expositiva. El *contexto de negociación* hace referencia a las negociaciones en torno a la información que se produce el transcurso de la evaluación. El *contexto de valoración* se conforma a partir de «las motivaciones, las intenciones, las actitudes, las presiones, los cambios y las vicisitudes que han manifestado los protagonistas durante el proceso de evaluación» (2004: 917). Por último, al fijarse en el *contexto de difusión*, el metaevaluador podrá descubrir si se ha difundido suficientemente el informe, salvaguardándose así el derecho de los ciudadanos a conocer qué ha sucedido con la evaluación y el de los profesionales en el ejercicio de su actividad, si ha habido intentos de maquillarlo o de presentar solo las conclusiones más «amables» con el programa. La propuesta resulta extremadamente sugerente, si bien la distinción entre contextos no resulta sencilla ni siquiera en el plano analítico.

En la metaevaluación realizada por Santos Guerra *et al.* (2008) sobre el Plan de Evaluación de Centros Escolares de Andalucía, la política ocupa también un lugar destacado. Los autores, partiendo desde una perspectiva etnográfica y comprensiva, dirigen su atención no tanto a los resultados del Plan de Evaluación cuanto al desarrollo mismo del proceso evaluativo y, muy en particular, a la divergencia de posiciones e intereses y a las relaciones entre los actores implicados. La importancia que se atribuye a estos elementos es tal que, incluso, se convierten en centrales para explicar algunos de los problemas surgidos en

el transcurso de la metaevaluación. La investigación metaevaluativa no se construyó a partir de unos criterios previos sino sobre la base de profundizar en ciertas cuestiones que los miembros del equipo consideraron esenciales para comprender el Plan de Evaluación. Como puede apreciarse, el listado de cuestiones es extenso: cultura en evaluación entre los participantes, finalidades del Plan, tanto explícitas como implícitas, actitud y disposición de los evaluadores y su visión de la evaluación, adecuación del enfoque evaluativo a la realidad evaluada, forma en que se concibe la evaluación en el Plan, información proporcionada a los «evaluados», jerarquías en la evaluación, necesidad y obligatoriedad de la evaluación, su carácter real o de mero trámite, indefinición o mala definición de los roles asignados a ciertos actores, existencia de negociación y debate, prioridades en cuanto a la información que se recaba, pertinencia de las técnicas de recogida de datos, importancia otorgada a los informantes, tiempo invertido, conflictos surgidos durante la evaluación, presentación y difusión de la información, negociación y utilización del informe, efectos de la evaluación, participación o ausencia de los evaluados y el sentido descendente de la evaluación.

En el panorama latinoamericano, resulta también de interés detenerse en el esquema de metaevaluación propuesto por Díaz Rosales (2001). La metaevaluación consiste, a juicio de este autor, en el análisis, la interpretación y la sistematización científica del proceso de evaluación. Ello implica atender a sus resultados, a las categorías y modelos que se utilicen, a la metodología y la epistemología en la que se sustentan dichos modelos y a la ideología y los paradigmas científicos en los que se inscriben. La hermenéutica constituye la principal herramienta de la que se sirve la metaevaluación. A través de ella y fijando la atención en los elementos mencionados, es posible obtener una comprensión rica o completa –esto es, ética y política– del proceso evaluativo, que vaya más allá de sus resultados (considerados esencialmente como «síntomas» o expresiones de la evaluación).

Cuadro I.1
Propuestas y *checklists* para la metaevaluación
Cuadro-resumen

	Dimensiones, ítems, estándares
Bernstein y Freeman (1975)	La metaevaluación ha de juzgar... – el tipo de muestreo – el análisis de datos – los procedimientos estadísticos y el diseño de la investigación
Stufflebeam (1974)	La metaevaluación debe ocuparse de... – la validez interna – la validez externa – la fiabilidad – la objetividad – la relevancia – la importancia de los datos recogidos – el alcance de la información – la credibilidad – la oportunidad temporal – el grado de difusión de la evaluación y – el coste-eficacia
Schwandt y Halpern (1988)	La metaevaluación debe comprobar... – si la metodología empleada en la metaevaluación es la adecuada – si las evidencias encontradas son correctas (y, por tanto, relevantes, fiables, suficientes, representativas y han sido recabadas a tiempo)
Uusikilä y Virtanen (2000)	Los criterios empleados en la metaevaluación se agrupan en... – Guía (definición del proceso de evaluación y expectativas de la agencia) – Selección de los evaluadores – Planificación y desarrollo del proceso evaluativo – Calidad de los informes de evaluación – Contactos y colaboración entre los distintos actores durante el proceso de evaluación – Experiencias de las agencias y los evaluadores sobre el proceso evaluativo – Valoración global por parte de agencias y evaluadores acerca del éxito del proceso evaluativo
Evaluación democrática deliberativa (House y Howe 1999; 2000)	La metaevaluación debe comprobar si la evaluación se ha atendido a los principios de... – Inclusión – Deliberación – Diálogo
Stake (2006)	La propuesta parte de dos premisas: – la valía del evaluando no es un rasgo inherente al mismo que está esperando a ser descubierta por el evaluador sino que se construye a través de los juicios y percepciones de las personas implicadas. – la metaevaluación gira en torno a las personas, a sus experiencias y a los hechos del proceso evaluativo.

	Dimensiones, ítems, estándares
Metaevaluation Checklist (Scriven 1991)	<p>La evaluación debe...</p> <ul style="list-style-type: none"> - explicitar y justificar los estándares de valor utilizados - ser comprensiva - ser comprensible - ser creíble - ser ética - ser factible - ser precisa desde el punto de vista técnico –pero sin que se olviden cuestiones como la relevancia de la información proporcionada - ser «sensible» desde el punto de vista político –esto es, que atienda y responda a las necesidades de las distintas audiencias - ser oportuna - ser segura - ser válida - ser rentable en términos de coste-eficacia.
Key Evaluation Checklist (Scriven 1991; 2007)	<p>La metaevaluación debe analizar...</p> <ul style="list-style-type: none"> - los antecedentes y el contexto de la evaluación - la definición de los términos empleados - las descripciones de la evaluación que hagan sus responsables - los criterios de valor - la identificación de las audiencias y de los <i>stakeholders</i> - la determinación de los costes y recursos - la valoración del proceso de investigación y de sus resultados - las alternativas existentes a la evaluación - la posibilidad de generalización de la evaluación - la significación global de la investigación evaluativa - la existencia de recomendaciones sobre el uso o el destino de la evaluación.
Joint Committee on Standards for Educational Evaluation (1994; 2011)	<p>Las evaluaciones deben cumplir los siguientes requisitos...</p> <ul style="list-style-type: none"> - Utilidad - Viabilidad - Pertinencia (adecuación a las normas)¹ - Rigor (precisión)
Program Evaluations Metaevaluation Checklist (Stufflebeam 1999)	<p>Las evaluaciones deben cumplir los siguientes requisitos...</p> <ul style="list-style-type: none"> - Utilidad - Viabilidad - Pertinencia (adecuación a las normas)¹ - Rigor (precisión)
Guiding Principles Checklist for Evaluating Evaluations (Stufflebeam, Goodyear, Marquart y Johnson 2005)	<p>La metaevaluación debe analizar...</p> <ul style="list-style-type: none"> - Si la evaluación es una investigación sistemática - El grado de competencia de los evaluadores - La integridad y honestidad del proceso - Si ha existido respeto por las personas - Si los evaluadores han asumido responsabilidad por el bienestar de la sociedad
Bustelo (2002)	<p>La metaevaluación debe analizar...</p> <ul style="list-style-type: none"> - cómo se plantean las evaluaciones y cuáles son las estrategias evaluativas empleadas - el papel de algunos elementos clave en la evaluación y - la utilización de las evaluaciones

¹ El término en inglés es *propriety*.

	Dimensiones, ítems, estándares
Cardozo Brum (2006)	<p>La metaevaluación analiza y enjuicia...</p> <ul style="list-style-type: none"> – el contexto en el que se desarrollan las evaluaciones (planificación de las evaluaciones, origen de la decisión de evaluar, instancias intervinientes (y el papel formal que les corresponde), funcionamiento real del sistema, recursos, difusión de los resultados de las evaluaciones o utilización de los resultados en la adopción de decisiones) – la metodología y los procesos internos de la evaluación (fuentes de información y análisis de datos, pertinencia y diseño del programa, recursos utilizados, análisis de procesos, medición de la cobertura, resultados alcanzados, impactos, participación y satisfacción social, utilización de resultados y recomendaciones, informes finales, coste de la evaluación y escenarios prospectivos)
Santos Guerra (1999) y Santos Guerra y Moreno Olivios (2004)	<p>El metaevaluador explora seis contextos...</p> <ul style="list-style-type: none"> – contexto de constitución (génesis de la evaluación y los procesos de negociación que se desarrollen en ella) – contexto de justificación (métodos empleados para la obtención de datos, su aplicación, información resultante y tratamiento que se ha hecho de ella) – contexto de presentación (contenido y forma de presentación del informe, su estructura lógica y su coherencia interna, su adecuación a las necesidades de las distintas audiencias, extensión o la claridad expositiva) – contexto de negociación (negociaciones en torno a la información que se produce en el transcurso de la evaluación) – contexto de valoración («motivaciones, intenciones, actitudes, presiones, cambios y vicisitudes que han manifestado los protagonistas durante el proceso de evaluación») – contexto de difusión (¿se ha difundido suficientemente el informe, salvaguardándose el derecho de los ciudadanos a conocer qué ha sucedido con la evaluación y el de los profesionales en el ejercicio de su actividad?, ¿ha habido intentos de maquillarlo o de presentar solo las conclusiones más «amables» con el programa?)
Díaz Rosales (2001)	<p>La metaevaluación contiene los siguientes apartados...</p> <ul style="list-style-type: none"> – abordaje general (evaluabilidad, importancia, pertinencia, autoridad de la evaluación. – interpretación y comprensión (ética, ¿trata lo que tiene que tratar?, ¿es honesta y responsable socialmente?, epistemología, ¿es rigurosa?, ¿existen sesgos?) – calidad de los datos, información e informe (honestidad e imparcialidad de las fuentes, triangulación) – crítica (comparación de marcos de valores, equidad, calidad de la evaluación y responsabilidad profesional del evaluador) – comentario y conclusiones (utilidad y aplicaciones, valor social)

Fuente: Elaboración propia a partir de las obras citadas.

El camino recorrido a lo largo de estas últimas páginas y sus principales hitos (léase propuestas o autores) constituyen las referencias de las que nos serviremos para construir, en el siguiente capítulo, nuestro esquema de metaevaluación.

CAPÍTULO II

UN ESQUEMA DE METAEVALUACIÓN DESDE LA CIENCIA POLÍTICA Y EL ANÁLISIS DE POLÍTICAS PÚBLICAS

1. A VUELTAS CON EL CARÁCTER POLÍTICO DE LAS EVALUACIONES

Como hemos tenido oportunidad de comprobar en el capítulo anterior, la gran mayoría de los enfoques sobre metaevaluación incorpora criterios sobre la conveniencia de fomentar la participación de los *stakeholders*, hacer de la evaluación un espacio representativo de la pluralidad de intereses en juego¹, propiciar el diálogo² y la negociación y atenerse a unos «mínimos» estándares o principios éticos durante el proceso evaluativo. La consideración de todas estas cuestiones se explica por la forma en que se concibe la evaluación, que es vista no sólo como una empresa técnica sino como un tipo de investigación fuertemente impregnado de connotaciones políticas.

Son varias las razones por las que se puede afirmar la naturaleza política de las evaluaciones:

- (i) Los elementos que constituyen su objeto de estudio –las políticas y los programas públicos– son propuestos, elaborados, debatidos, financiados y aprobados en el seno del proceso político (Weiss 1987; Palumbo 1987). Su dinámica, sus estructuras y sus actores –en especial, aquellos implicados en la política pública evaluada– inciden notablemente en su realización.

¹ Señala PITKIN que el ámbito político –y la evaluación es netamente política– es «la clase de contexto en el que se hace relevante la representación como actividad sustantiva. Pues la representación no es necesaria allí donde esperamos soluciones científicamente verdaderas, allí donde no están involucrados compromisos de valor, decisiones ni juicios (...) Necesitamos la representación precisamente allí donde no nos contentamos con abandonar las cuestiones en manos del experto; podemos tener una representación sustantiva solo allí donde el interés se vea implicado» (1985: 236).

² Y es que, como recuerda MAJONE, «no se necesita ahora tanto desarrollar medidas “objetivas” de los resultados –la vieja meta de la investigación para la evaluación– como facilitar un diálogo profundo entre quienes defienden criterios distintos» (1997: 230).

- (ii) El impulso hacia su consolidación procede, fundamentalmente, de los poderes públicos, en forma de institucionalización y/o de financiación. Los gobiernos ven en ella un elemento de legitimación al tiempo que una oportunidad para el control. Al fin y al cabo, como afirma House, «el gobierno legitima la evaluación y la evaluación legitima la actuación gubernamental» (1993: x).
- (iii) Su propósito –tácito o explícito– es informar e influir, en un sentido o en otro, en el proceso de toma de decisiones políticas. No quiere ello decir que se orienten siempre a la mejora de la política o el programa en cuestión. Con frecuencia, se emplean como medio de rentabilizar el «esfuerzo político» o, incluso, se convierten en un arma arrojada entre opciones de distinto signo. Son pues especialmente proclives a ser instrumentalizadas y utilizadas para fines políticos o partidistas³.

Hasta hace unas décadas, sólo los autores próximos a enfoques «cualitativistas» –MacDonald (1977), House (1980) o Guba y Lincoln (1981; 1987; 1989), por citar algunos– hubieran suscrito estas afirmaciones. Hoy en día, sin embargo, prácticamente nadie cuestionaría su validez, poniendo en duda que la evaluación «es una práctica tanto política como científica» (Schwandt 2005: 5). Es cierto que, como nos recuerda con cierta sorpresa este mismo autor, aún hay quienes creen que la política representa una amenaza para la evaluación, el lado de la subjetividad frente a la dimensión técnico-científica en la que reinan los hechos objetivos y las explicaciones contrastadas y que los evaluadores deberían intentar minimizar la influencia de la política en sus evaluaciones. Sin embargo y, por fortuna, cada vez son más los autores que reconocen que las evaluaciones son procesos inherentemente políticos, aunque entre ellos existan diferencias acerca de cómo se deba gestionar la «política de la evaluación» (Greene 2003)⁴. Buena prueba de ello es que, en los últimos años, el interés por la relación política-evaluación –y, más recientemente, por la contribución de la evaluación a la democracia– no ha hecho sino acrecentarse y no solo en el ám-

³ Ello no nos debe llevar, como nos previene MORGAN, «a ver política por todas partes y a buscar “agendas escondidas” donde no hay agendas» (1993: 212). Así pues, el hecho de que las evaluaciones tiendan o puedan ser instrumentalizadas políticamente, no debe llevar al estudioso a partir de una aproximación *patológica* en su análisis. El que una evaluación no sea utilizada tal y como fue previsto o el que su diseño técnico no resulte ser el más apropiado para los objetivos propuestos no quiere decir necesariamente que detrás de todo ello haya una intencionalidad política ilegítima.

⁴ Apunta esta autora que, entre los partidarios de esta tesis, hay dos posturas muy claras. Por un lado, la de quienes consideran que el evaluador debe comprometerse explícitamente con alguna de las opciones ideológicas que afloran en el proceso de evaluación, discriminar positivamente a aquellos *stakeholders* que se encuentren en posición de desventaja e intentar influir sobre el contexto político induciendo cambios en el mismo (el enfoque de *empowerment* o la evaluación democrática). Por otro, aquellos que apuestan por el pluralismo, por integrar las distintas perspectivas en la evaluación pero *sin tomar partido* por ninguna de ellas, con objeto de evitar sesgos que puedan poner en entredicho la credibilidad del proceso (el modelo de evaluación «sensible» o *responsive evaluation* o los centrados en la utilización).

bito de la evaluación⁵. El tema ha logrado hacerse un espacio destacado, como hemos visto, en los enfoques de metaevaluación pero también en los catálogos de estándares y principios-guía de asociaciones o colectivos profesionales⁶. Y es que en un contexto como el actual, marcado por la preocupación por el buen gobierno y la calidad de la democracia, la clásica tesis «cualitativista» solo podía cobrar nueva fuerza.

Pero no son estos los únicos argumentos que pueden aducirse. Como ya señalamos en otro lugar (García Sánchez 2003; 2009), para profundizar en el tema resulta oportuno acudir al potencial analítico que puede ofrecernos la Ciencia Política. Esta disciplina y, en particular, los estudios sobre los actores, los procesos de toma de decisiones o los más recientes de políticas públicas⁷ resultan ser, para nuestros propósitos, una referencia inexcusable. Desde la Ciencia Política, la evaluación se puede concebir:

- (i) Como una actividad intrínsecamente política desde el momento en que en ella se ven involucrados actores con intereses contrapuestos, políticamente relevantes, que luchan por obtener una mayor cuota de poder e influencia en el proceso con el propósito de modularlo de acuerdo a sus preferencias. Y es que las evaluaciones, lejos de constituir escenarios eminentemente racionales, «asépticos» e integrados en torno a unos objetivos comunes, están presididas por la diversidad de intereses. Los distintos actores tienen percepciones y aspiraciones distintas acerca del programa o la política, de cómo deban juzgarse, cómo deba ser la evaluación y del papel que en ella les deba corresponder. El escenario de la evaluación se define, así, a partir de la interacción política de los actores implicados y/o con intereses en la evaluación.

⁵ Como muestran los trabajos más recientes de HOUSE (2006), STAKE (2006), KUSHNER (2005), KUSHNER y ADELMAN (2006), MACDONALD y KUSHNER (2005), SCHWANDT (2003; 2005) o GREENE (2001; 2003), entre otros.

⁶ Por ejemplo, los de las sociedades suiza, alemana o norteamericana de evaluación disponibles, respectivamente, en las siguientes direcciones de Internet >http://www.seval.ch/en/documents/SEVAL_Standards_2000_en.pdf (consulta: 27 diciembre 2010), <http://www.degeval.de/calimero/tools/proxy.php?id=71> (consulta: 27 diciembre 2010) y <<http://www.eval.org/Publications/aea06.GPBBrochure.pdf>> (consulta: 11 febrero 2011).

⁷ Actores y decisiones han merecido una atención preferente por parte de la Ciencia Política, tanto en sí mismos cuanto por su importancia en las políticas públicas. Concretamente, los procesos decisivos han ocupado una posición central entre los intereses de los estudiosos de políticas desde Lasswell (1951/2000); de hecho, el *decision making approach* (que se nutre de la teoría de juegos, las investigaciones sobre administración pública y la teoría de las organizaciones) constituye una aportación fundamental al análisis de políticas públicas (JEREZ MIR 1999, cit. en MARTINÓN 2006: 36). Las propuestas en este campo son enormemente variadas: desde los trabajos de BUCHANAN y TULLOCK (1962) y BUCHANAN y TOLLISON (1972), que se sitúan en una perspectiva racionalista, pasando por los de SIMON (1957), DAHL (1957), CROZIER y FRIEDBERG (1977), MARCH (1974) o HIRSCHMAN (1967) que ofrecen sugerentes críticas del modelo racional, hasta el incrementalismo de LINDBLOM (1959; 1991).

- (ii) configurada a partir de las decisiones (o no decisiones) que adoptan los distintos actores en el transcurso del proceso evaluativo. Estas decisiones –por ejemplo, qué actores participan, de quién es la propiedad de la información producida o a qué enfoque metodológico se recurre– implican un determinado posicionamiento sobre la política pública que se evalúa y sobre el propio proceso de evaluación, en definitiva sobre las asignaciones de valor que en ellos se realizan⁸. La evaluación siempre «defiende» o «promueve» unos determinados valores e intereses y descarta otros⁹.

El poder, el clásico objeto de estudio de la Ciencia Política, es visto por el *policy analysis* desde la perspectiva de sus resultados, las políticas públicas. En la conformación de estos productos no intervienen únicamente las estructuras estatales aunque a estas corresponda el papel de principales impulsoras. Los partidos políticos, las burocracias, las asociaciones de empresarios, los sindicatos, los movimientos sociales presionan a los gobiernos para que incorporen a su agenda los asuntos que les preocupan y para que las políticas se reorienten en función de sus intereses. Las políticas públicas se convierten entonces en el resultado de complejos procesos de ajuste y negociación entre actores políticos, económicos y sociales. El análisis de políticas públicas recupera, de este modo, la preocupación por los actores en detrimento de las estructuras¹⁰.

Algo similar sucede en el caso de las evaluaciones. Los procesos de evaluación se van construyendo a partir de las decisiones (o no decisiones) y actuaciones de ciertos actores. La importancia del *actor* es tal que es precisamente su intervención la que nos permite distinguir y reconocer, al menos analíticamente, las diferentes etapas del proceso evaluativo. El escenario de la evaluación –que no es sino una reproducción parcial y a pequeña escala del de las políticas o los programas a los que aquella se vincula– se define, pues, a partir de la interacción de los actores. Pero el escenario determina también cuáles serán los actores participantes y cuál será el grado de agregación de los actores colectivos.

Lo dicho hasta ahora no ha de entenderse como una forma soterrada de afirmar que las decisiones y los comportamientos de los actores sean las únicas variables que haya de manejar el metaevaluador. Conscientes de las debilidades

⁸ Nos servimos aquí de la clásica definición de EASTON (1969/2006) según la cual la política es aquella actividad orientada a la regulación de los conflictos sociales mediante decisiones que tienen carácter vinculante para el conjunto de la sociedad. A través de esas decisiones se otorga valor a determinadas demandas o intereses en detrimento de otros. Así, un acto se convierte en político cuando se asocia o tiene que ver con dichas decisiones, con las asignaciones imperativas de valor en una sociedad o en un contexto dado.

⁹ Tal y como recuerda JENNIFER GREENE. Véase la entrevista a la autora en *Harvard Family Research Project* (2005).

¹⁰ Así lo consideran, entre otros, DYE (1987) y SUBIRATS (1990-1991). Esto es cierto incluso en enfoques como el redes, el de coaliciones promotoras o el centrado en las *policy community*: en última instancia, las coaliciones promotoras, las redes o las comunidades políticas no son sino conjuntos –más o menos articulados– de actores.

del «enfoque de actor» (von Beyme 1994) creemos, con Marsh y Smith (2000), que tan importante es atender a los actores como a los elementos estructurales a la hora de explicar los fenómenos sociales y políticos. De hecho, las estructuras, aunque en ocasiones operen como actores, suelen ser los escenarios en los que estos despliegan su actuación. La apuesta que hacemos en este trabajo por los actores frente a las estructuras no es casual sino que se deriva de la propia fuerza de los hechos. La ausencia de estructuras de evaluación en el ámbito y el período estudiados otorgó un protagonismo claro a los actores. Sus decisiones contribuyeron no solo a conformar los procesos evaluativos sino a «construir» las estructuras mismas en lo que fue el proceso de institucionalización de la evaluación educativa en España. Las evaluaciones daban comienzo, en aquellos momentos, a raíz de la iniciativa personal de determinados actores y se desarrollaban como actividades periféricas, casi anecdóticas, al margen de las funciones habituales de las estructuras organizativas en las que se encuadraban, por lo que la influencia de estas era ciertamente limitada. La evaluación se encontraba aún en un estadio de desarrollo incipiente y no se contemplaba siquiera la posibilidad de que los resultados de las evaluaciones pusieran en peligro la continuidad de un programa cuando este era una opción de política educativa¹¹.

A la vista de todas estas consideraciones, parece evidente que la metaevaluación no puede quedar reducida al mero examen de la calidad técnica de una evaluación. Así pues, «no basta interrogarse sobre la calidad de los datos que se recogen en el informe o sobre el rigor de los métodos que se han aplicado para extraer información, sino que es preciso preguntarse por la naturaleza de la iniciativa, por la finalidad del proceso, por las reglas éticas que lo han inspirado, por la utilización que se ha hecho de la información, por el control democrático de los informes, por las secuelas que la evaluación ha dejado, etc.» (Santos Guerra 1999: 267). Ciertamente es que las evaluaciones son, en lo esencial, procesos de investigación, pero también lo es que son empresas eminentemente políticas que se desarrollan en un contexto también político. Parece lógico, entonces, que nuestro análisis de los procesos evaluativos incorpore el estudio del binomio actores-decisiones. Nos interesará, por ejemplo, determinar qué actores participaron en cada uno de los procesos de evaluación, cuál es su naturaleza, cómo se distribuyó el poder entre ellos, esto es, de qué modo tomaron parte, qué posición ocuparon en el proceso decisorio y qué rol desempeñaron en la evaluación. Es evidente que algunas de estas cuestiones tienen además una trascendencia técnica¹². Pero si se acepta que en una democracia los ciudadanos tienen un derecho legítimo a valorar la actuación de los poderes públicos, la importancia de esos temas adquiere una dimensión que excede el ámbito de lo metodológico. ¿Es

¹¹ Ello hacía también que la evaluación no despertara aún el interés de actores relevantes como los partidos políticos. Sobre el proceso de institucionalización educativa en el ámbito no universitario véase GARCÍA SÁNCHEZ (2003; 2004).

¹² Por ejemplo, la falta de representatividad en las evaluaciones puede hacer que la imagen que estas proyecten del programa sea incompleta o esté sesgada.

admisible excluir a ciertos actores de un proceso de evaluación? ¿Ha de ocuparse el metaevaluador de comprobar si los poderes públicos u otros actores han obstaculizado su participación? Creemos, con Patton (1997), que la metaevaluación se hace más necesaria cuanto más politizado esté el contexto y cuanto más «visibilidad» e importancia hayan alcanzado las evaluaciones en él.

2. LAS PREMISAS DE PARTIDA

Las reflexiones previas constituyen la base a partir de la cual se formulan las dos premisas en las que descansa nuestro esquema para la evaluación de las evaluaciones:

- (i) La metaevaluación ha de reconocer, analizar y enjuiciar tanto la dimensión técnico-metodológica como la dimensión política de las evaluaciones.
- (ii) Las decisiones que los distintos actores adoptan y las relaciones que se establecen entre ellos afectan e inciden en ambas dimensiones, contribuyendo a configurar los procesos de evaluación.

2.1. Las dos dimensiones de la evaluación

Así pues, nuestra propuesta se construye, en primer lugar, desde el reconocimiento de la doble dimensión de las evaluaciones.

En la denominada *dimensión técnico-metodológica* cabría incluir las características más evidentes y explícitas de las evaluaciones, aquellas relacionadas con la evaluación en tanto que proceso de investigación: los objetivos formales, el enfoque evaluativo y el metodológico, los métodos y las técnicas de recogida y análisis de datos, etc. Tendrían aquí encaje rasgos habitualmente asociados con la calidad de la investigación científica como la validez, la fiabilidad, el rigor o la consistencia del análisis.

La *dimensión política* de la evaluación, por su parte, integra aquellos componentes ideológicos y cognitivo-simbólicos asociados a las ideas, valores, opiniones e intereses de los distintos actores así como los vinculados a las relaciones (de poder o de búsqueda del poder) que se establecen entre ellos. Es la presencia de intereses contrapuestos –no solo con respecto a la política evaluada sino también al propio proceso de evaluación– lo que confiere a las evaluaciones esa dimensión política. Al dirigir nuestra atención hacia este plano podemos advertir algunas de las claves que explican por qué fracasan algunos procesos evaluativos –tremendamente costosos e «impecables» desde el punto de vista técnico– o por qué sus resultados no son utilizados, descubrir los desequilibrios de poder y qué *stakeholders* controlan la adopción de las principales decisiones y cómo todo ello afecta al desarrollo y los resultados de las evaluaciones.

La distinción entre las dos dimensiones es meramente analítica, no «real». De hecho, muchas de las decisiones adoptadas por los *stakeholders* afectan, en mayor o menor medida, a las dos dimensiones. Así, resulta extremadamente difícil encontrar una cuestión técnica que no tenga implicaciones políticas –por ejemplo, la elección del enfoque metodológico–, de la misma forma que lo es hallar una cuestión considerada *a priori* como política que no tenga consecuencias en el plano técnico –la participación o exclusión de determinados *stakeholders*.

2.2. Actores y decisiones

2.2.1. *El concepto de actor*

Pocos conceptos resultan tan centrales para la Ciencia Política contemporánea como el de actor. El término ha sido y es profusamente utilizado por teóricos y analistas y aparece, cada vez con más frecuencia, en nuestros discursos cotidianos. Quizá por ello resulta tan sorprendente que sean tan escasas las propuestas para definirlo: muchas de las investigaciones que afirman partir del *enfoque de actor*, como lo denomina von Beyme (1994), se limitan a señalar su importancia o se aventuran al análisis empírico sin haber aclarado previamente los contornos del concepto.

Así las cosas, entendemos imprescindible comenzar por hacer explícita una definición que nos permita superar el actual estado de sobreentendidos y nos sirva de guía a la hora de identificar y clasificar en la práctica, en el análisis empírico, a los actores que tomaron parte en cada uno de los procesos evaluativos. La identificación de los actores individuales, de las personas implicadas en un fenómeno social, en un proceso de evaluación, no presenta mayores dificultades. Sin embargo, no sucede lo mismo en el caso de los actores colectivos: evidentemente, no todo grupo de individuos puede ser considerado, de manera automática, un actor. No hay duda de que una empresa, un parlamento, un colegio profesional o un centro escolar merecen tal calificativo. Pero ¿es posible aplicar la definición de actor a aquellos otros colectivos, como el profesorado, los movimientos educativos o los estudiantiles, que carecen de una organización interna estable y de una «cabeza visible» y a los que resulta difícil atribuir alguna responsabilidad?

En la definición por la que aquí optamos se considera actor (colectivo) a aquella entidad (i) cuyos miembros están integrados en torno a similares –o, al menos, convergentes– percepciones e intereses, (ii) que cuenta con un cierto grado de organización y recursos y con mecanismos para la resolución de conflictos internos, (iii) que tiene los medios y la capacidad para decidir y/o actuar intencionada y estratégicamente para la consecución de un objetivo común¹³

¹³ SCHARPF (1997: 58).

como unidad suficientemente cohesionada¹⁴, lo que le identifica y diferencia frente al resto y (iv) a la que, por tanto, se le puede atribuir una responsabilidad por sus decisiones y/o actuaciones. En otras palabras, un actor es una *unidad de decisión-acción responsable*¹⁵.

Ahora bien, ¿qué sucede con aquellos colectivos –por ejemplo, los ciudadanos que inician una revuelta, los maestros o los médicos– de perfiles desdibujados, que son identificables en la práctica pero cuya escasa o nula formalización nos hace cuestionarnos su carácter de actor? La clave del problema parece residir, como sugiere Scharpf (1997), en el extremadamente variable grado de integración que presentan¹⁶. Pero es justo esa variabilidad la que nos permite advertir la existencia de, al menos, dos grandes tipos de *actores colectivos*: los denominados *actores analíticos* y los *actores colectivos propiamente dichos*.

En el caso de los primeros, la integración tan solo se da en el plano del análisis. Entre sus miembros no existe relación alguna: su único nexo de unión son las similitudes que el investigador les atribuye con objeto de poder operar con ellos como si de una unidad se tratase y de incluirlos en su esquema explicativo. Ciertamente, estas *colectividades taxonómicas*, como las denomina Harré (1981) son meras ficciones pero, sin ellas, sería imposible abordar el estudio de ciertos fenómenos sociales. Muchos de los colectivos a los que haremos referencia en este trabajo –los *equipos multiprofesionales* (EE.MM.) en la evaluación del programa de integración o los *monitores* en la evaluación del proyecto *Atenea*, por citar algunos ejemplos– tendrán esta consideración.

Los *actores analíticos* se distinguen claramente de los *actores colectivos propiamente dichos* –o *compuestos* en palabras de Scharpf (1997)–, aquellos en los que se da algún grado de integración, por pequeño que sea. Dentro de este segundo grupo existe también una gran diversidad¹⁷. En él encontraríamos des-

¹⁴ La capacidad de actuación unitaria del actor colectivo no es, sin embargo, algo permanente e inquebrantable. En ocasiones, la unidad de acción puede debilitarse o ceder ante un problema específico, por la divergencia de opiniones o ante la capacidad de liderazgo de uno de sus miembros.

¹⁵ Los referentes teóricos a los que hemos recurrido para elaborar esta definición pueden encontrarse en GARCÍA SÁNCHEZ (2003; 2007).

¹⁶ Como es obvio, el grado de integración de los actores no nos interesa en sí mismo sino tan solo cuando sea tal que permita otorgar a un colectivo la denominación de *actor*.

¹⁷ SCHARPF (1997) distingue tres tipos de *actores compuestos*, dependiendo de si la acción, el control de los recursos, el objetivo o propósito y las decisiones de sus miembros se definen individual o colectivamente: los *actores agregados*, los *actores colectivos* y los *actores corporativos*. Los primeros son los que presentan un menor grado de integración. Cada uno de sus miembros persigue sus propios objetivos, decide y actúa individualmente y controla sus propios recursos. Dentro de la categoría de *actores colectivos*, SCHARPF distingue cuatro tipos: las «coaliciones» (los objetivos se definen de manera individual y el control de los recursos permanece en manos de sus miembros), los «movimientos» (persiste el control individual de los recursos por parte de sus miembros pero estos persiguen un objetivo común), los «clubes» (el control sobre los recur-

de colectivos altamente estructurados, como las organizaciones, hasta colectivos escasamente formalizados, como los movimientos sociales, que se caracterizan casi exclusivamente por perseguir un objetivo común¹⁸.

En lo que a las decisiones se refiere, hemos de señalar que no todas las que afectan al proceso de evaluación parecen igualmente importantes. La revisión de los distintos enfoques metaevaluativos nos ha permitido identificar las más relevantes, aquellas que –planteamos aquí como hipótesis– definen el marco de principios esenciales de la evaluación y que denominamos, por su relevancia, decisiones «críticas» o «configuradoras». Las *decisiones configuradoras* afectan a: (i) a la determinación del objeto y los objetivos de la evaluación, (ii) a la elección del enfoque metodológico, (iii) a la especificación de los criterios de valor, (iv) a la selección de los actores participantes y, en su caso, a la determinación del papel que les corresponde en el proceso y (v) al estilo¹⁹ de los informes y al grado de difusión la información producida. De confirmarse la importancia de las decisiones configuradoras, los actores que logran tomar parte en las mismas serían quienes ocuparan una posición de mayor relevancia y quienes controlarían el proceso de evaluación.

2.2.2. *¿Qué factores explican las decisiones y las actuaciones de los actores?*

A pesar de que la complejidad de los fenómenos sociales y políticos (rara vez) puede explicarse a partir de un único elemento, es posible identificar ciertos factores que, inicialmente, parecen tener una mayor influencia en la conformación de las decisiones y/o actuaciones de los actores. Evidentemente, no cabe ningún *a priori* en este sentido: solo a través del análisis empírico se podrá determinar cuál o cuáles de esos factores tienen, en la práctica, más peso sobre el comportamiento del actor.

Entendemos aquí que las decisiones y actuaciones de los actores dependen de la acción combinada de cinco variables: i) su posición, ii) sus intereses, iii) sus sistemas de creencias, iv) sus habilidades y v) el entorno de oportunidad en el

... es colectivo aunque no exista un propósito común) y, finalmente, las «asociaciones» (tanto el control sobre los recursos como la finalidad de la acción se definen en términos conjuntos). La denominación de *actor corporativo* es únicamente aplicable a las organizaciones de estructura jerárquica en las que los miembros, aunque no se implican activamente en la definición del curso de acción, tienen la posibilidad de seleccionar y reemplazar a sus líderes. Su grado de integración es máximo.

¹⁸ La identificación de un actor colectivo dependerá siempre del nivel de análisis en el que nos situemos. Así, en un caso el actor puede ser una institución y, en otro, la persona que está al frente de dicha institución o un subconjunto de la misma.

¹⁹ Nos referimos a si el informe está redactado con un estilo técnico, comprensible solo para quienes posean conocimientos especializados o, por el contrario, si sus términos son asequibles incluso para no iniciados. Véanse ESCUDERO MUÑOZ (1995) y SÁEZ BREZMES (1995).

que despliega su actuación²⁰. La influencia de algunos de estos factores va más allá: así, es la convergencia en torno a unos intereses y creencias (además de la unidad de acción) lo que permite atribuir a un colectivo el carácter de actor.

La *posición* del actor es aquella que se deriva de su cargo (en el caso de los actores individuales) o de su situación en la estructura informal de la organización. Cabe esperar que cuanto mayor sea la responsabilidad del actor por razón de su cargo, mayor será su margen de maniobra. Sin embargo, el que el actor sepa aprovechar las ventajas que le reporta estar en determinado puesto dependerá también, en buena medida, de la habilidad que sea capaz de desplegar.

Bajo la denominación de *intereses* se incluyen las predisposiciones y expectativas que inducen a una persona a actuar y a hacerlo en un sentido determinado²¹. Los intereses (ya sean personales, profesionales o políticos²²) son, sin duda, una de las claves de la viabilidad del proceso de evaluación. El interés de uno o varios actores puede hacer que una evaluación prospere pese a las circunstancias desfavorables. Del mismo modo, su desinterés o una actitud reticente pueden lograr que el proceso de evaluación no llegue a iniciarse e incluso que, una vez iniciado, se paralice.

La realidad política, lejos de fraguarse en escenarios eminentemente racionales e integrados en torno a unos objetivos comunes, está presidida por la diversidad de intereses. Cada actor lucha por hacer valer sus pretensiones y por obtener la mayor cuota posible de poder en sus transacciones. La diversidad preside no solo las relaciones entre actores sino que es inherente al propio individuo. En otras palabras, los individuos albergan también intereses contrapuestos: sus metas personales pueden colisionar con lo que son sus objetivos profesionales. El sujeto, en función de la prioridad que en un momento dado tenga para él una determinada cuestión, decidirá cuál de sus intereses será el que oriente su actuación (Morgan 1993). Téngase en cuenta que la disparidad de criterios y motivaciones, lejos de tener efectos disfuncionales, puede llegar a convertirse en un reto y en una oportunidad para el cambio²³.

²⁰ En los estudios de caso, analizaremos la influencia de estos cinco elementos solo en aquellos actores que tomaron parte activa en los procesos de evaluación, esto es, en calidad de decisores.

²¹ Definición basada parcialmente en la ofrecida por MORGAN (1993: 129) y en la reflexión de STENMARK (2000-2001).

²² Desde nuestro punto de vista, la afiliación partidista, la afinidad ideológica o la vinculación al programa de los promotores de una evaluación deben ser considerados por el investigador como posibles factores explicativos del comportamiento de los actores, aunque ello no debe llevar a suponer que los mismos introducen necesariamente sesgos en la evaluación. En otras palabras, el que una evaluación sea promovida o realizada por personas próximas al programa no dice nada, *a priori*, acerca de su rigor y objetividad.

²³ Coincidimos, pues, con la visión del conflicto organizativo que se tiene desde la denominada corriente política o de poder y conflicto en las organizaciones. Pueden consultarse, a este respecto, CROZIER (1974), BLAU (1963), PFEFFER (1981) y PERROW (1990).

En los *sistemas de creencias*²⁴ se agrupan los denominados paradigmas cognitivos –esto es, las asunciones causales sobre ciertos fenómenos que limitan las alternativas que los actores consideran eficaces o útiles– y los marcos normativos –creencias o valores– que restringen las opciones de actuación que los decisores barajan como aceptables y legítimas desde el punto de vista moral. Las preferencias metodológicas de un actor se sitúan, claramente, en el marco de su sistema de creencias.

Siguiendo a Yukl (1994), entendemos por *habilidades* el conjunto de capacidades (derivadas, en el caso de los actores individuales, de su formación académica, su experiencia profesional y su personalidad) que llevan a un actor a poner en práctica sus ideas de manera eficaz y exitosa. Las habilidades no se refieren solo a la competencia para resolver problemas o para planificar actividades de manera estratégica, sino también a la aptitud que se demuestre en las relaciones con otros actores o a la destreza para hacer del entorno un escenario favorable a los intereses propios, aprovechando las oportunidades que brinde y salvando las restricciones y obstáculos que ofrezca. *A priori*, cuanto mayor sea el grado de integración de un actor colectivo en un momento dado, mayores serán sus posibilidades de alcanzar un resultado ventajoso.

El *entorno de oportunidad*²⁵ define y condiciona las posibilidades de actuación de los actores. En él tienen cabida una multiplicidad de factores –algunos perdurables en el tiempo (aunque no inmutables) y otros fortuitos o accidentales– que van desde los apoyos con los que cuenta un actor o la posición que ocupan los demás actores hasta el «momento» político, económico o social.

2.2.3. *Hacia una clasificación de los actores: naturaleza y modalidad de participación*

El análisis exploratorio de los tres procesos de evaluación seleccionados nos permitió identificar, siguiendo una lógica inductiva, los criterios a partir de los cuales se pueden clasificar los actores presentes en las evaluaciones.

El primer criterio –la naturaleza de los actores– convencional en Ciencia Política, nos permite distinguir tres categorías básicas: *actores político-administrativos*, *actores técnicos* y *actores corporativo-asociativos*.

El segundo de los criterios se refiere a la *modalidad de la participación* en el proceso evaluativo y distingue, dentro del grupo de actores participantes, entre quienes toman parte de manera *activa* y quienes lo hacen de forma *pasiva*.

²⁴ Tomamos el término «sistema de creencias» del enfoque de las coaliciones promotoras. Véase, en especial, el artículo publicado por SABATIER y HUNTER en 1989.

²⁵ Esta idea coincide, en parte, con las de *estructura* y de *oportunidad social* de las que habla SIBEON (1999).

Cuadro II.1
**Actores con influencia potencial en los procesos de evaluación,
 según su naturaleza**

Actores político-administrativos	Instituciones Partidos políticos Autoridades político-administrativas ²⁶
Actores técnicos	Unidades organizativas a las que se adscribe el programa Unidades organizativas vinculadas a la evaluación Personal técnico del programa (asesores, monitores, equipos multi-profesionales...) Personal técnico de la evaluación (diseñadores y aplicadores de pruebas, analistas de datos...) Equipos de evaluación
Actores corporativo-asociativos	Centros escolares Profesorado Asociaciones profesionales de profesores Sindicatos Movimientos educativos Familias Asociaciones de padres de alumnos Colectivos o asociaciones de alumnos

Fuente: Elaboración propia.

Así, los *actores activos* son aquellos cuyas actuaciones y/o decisiones conforman el proceso de evaluación. Su importancia en la evaluación se deriva precisamente de su participación activa en la evaluación, ya sea en calidad de decisores (por ejemplo, quienes toman la iniciativa de evaluar un programa, elaboran las pruebas para la recogida de datos o se ocupan de los diseños muestrales) o de aplicadores. Los *actores pasivos*, por su parte, intervienen en la evaluación en calidad de informantes, de meros receptores de las decisiones y/o actuaciones de los *actores activos*. Ello no quiere decir que no tomen decisiones sino que estas no son un elemento definitorio de su participación en la evaluación. Permanecen al margen del proceso decisorio por decisión propia (porque no tienen *interés* en participar), por la debilidad de su posición formal o informal, porque carecen de la *habilidad* necesaria para «hacerse un hueco» en el escenario de actores o porque su *entorno de oportunidad* no les es favora-

²⁶ Inicialmente, nos planteamos la posibilidad de distinguir dos categorías: «altos cargos» y «autoridades administrativas». En la primera se encuadrarían lo que la Ley 6/1997, de 14 de abril, de Organización y Funcionamiento de la Administración General del Estado, LOFAGE, considera «altos cargos» (los llamados *órganos* superiores –Ministros y Secretarios de Estado y los *órganos directivos* –Secretarios Generales, Subsecretarios, Secretarios Generales Técnicos y Directores Generales). En la segunda se incluirían únicamente los subdirectores generales. Sin embargo, con objeto de salvar las dificultades que plantea la distinción entre el nivel político y el administrativo y, en la medida en que creemos que ello no afecta de manera sustancial a nuestro análisis, optamos por emplear esta denominación genérica para designar tanto a los *altos cargos* como a los subdirectores generales.

ble (falta de recursos o de apoyos, afán de otros actores por vetar su acceso al proceso decisorio). En ocasiones, ni siquiera tienen conocimiento de que el programa está siendo evaluado.

Como se recordará, no todas las decisiones que se adoptan durante el proceso de evaluación tienen idéntica importancia; por ello, resulta conveniente diferenciar, dentro de los *actores activos*, aquellos que participan en la adopción de las *decisiones configuradoras* de la evaluación²⁷ –y a los que denominaremos *actores centrales*– de aquellos otros que, al no estar implicados en este tipo de decisiones, se sitúan en la periferia del proceso decisorio –los llamados *actores periféricos*.

Esta clasificación incorpora cierto matiz de gradación: las dos posiciones presentadas no son pues sino los extremos de un *continuum*. Con ello, queremos decir que el concepto de *centralidad* ha de ser definido no en términos absolutos sino relativos. Un actor que participe en la adopción de todas las *decisiones configuradoras* de la evaluación ocupará una posición de mayor *centralidad* que aquel otro que solo tome parte en alguna de estas decisiones.

3. EL ESQUEMA PROPUESTO: UTILIDAD Y RASGOS GENERALES

Se trata ahora, llegados a este punto, de conjugar las variables y los elementos –dimensiones, actores, decisiones– en las que nos hemos ido deteniendo a lo largo de este y del anterior capítulo con vistas a presentar nuestro esquema de metaevaluación. La utilización de este esquema en el análisis empírico permitirá que demos respuesta a la siguiente pregunta: ¿cuáles son las decisiones y actuaciones de los actores con respecto a cada una de las cuestiones consideradas en el esquema? En otras palabras, se trata de saber qué decisiones se tomaron (y sobre qué cuestiones no se adoptó ninguna decisión), quiénes las tomaron y a qué aspectos de las evaluaciones afectaron.

De esta forma, el esquema que proponemos no pretende ser un catálogo de estándares que la evaluación deba cumplir sino, más bien, un conjunto de cuestiones o subdimensiones (asociadas a la dimensión técnico-metodológica y a la dimensión política de las evaluaciones) que han sido identificadas a partir de distintos enfoques de metaevaluación y cuyo análisis permite reconstruir, comprender y enjuiciar el proceso evaluativo. Nos situamos, pues, más cerca de la visión comprensiva y orientada a la experiencia de la que habla Stake (2006) que de los métodos más clásicos y convencionales de metaevaluación. Entendemos que esta se construye a partir de la experiencia de los actores, del diálogo estrecho, la reflexión y la relación del metaevaluador «con las personas, con los hechos de la evaluación y con las opiniones de la gente» (Stake 2006: 257).

²⁷ Cfr. *supra*, apartado 2.2.1.

Asimismo, el esquema permite orientar tanto a los metaevaluadores como a los evaluadores a la hora de identificar qué decisiones pueden contribuir o han contribuido al éxito o al fracaso de un proceso evaluativo. Como es lógico, no se trata, en ningún caso, de buscar a los «culpables» o a los «héroes» de las evaluaciones sino de detectar aquellas decisiones que afectan o pueden afectar positiva o negativamente la utilidad, la pertinencia, el rigor o la viabilidad de las evaluaciones. El objetivo último es extraer lecciones que permitan una mejor «gestión» de los procesos evaluativos. Se trata, además, de una propuesta flexible, aplicable tanto en la metaevaluación formativa como en la sumativa y abierta a la inclusión de nuevos ítems y a la utilización de distintas técnicas de recogida y análisis de datos.

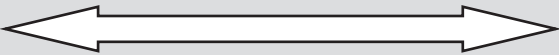
El siguiente cuadro recoge las subdimensiones que entendemos se han de tener en cuenta en la metaevaluación:

Cuadro II.2
Dimensiones y subdimensiones del esquema de metaevaluación

DIMENSIÓN POLÍTICA	←————→	DIMENSIÓN TÉCNICA
	Selección del equipo evaluador	
	Existencia y contenido del contrato de evaluación	
	Descripción y contextualización del programa	
	La evaluación ¿tiene en cuenta el contexto en el que se desarrolla?	
	¿Se prevén mecanismos de metaevaluación?	
	¿De qué modo se articulan evaluación y metaevaluación?	
	¿Quiénes participan en el proceso de evaluación?	
	¿De qué modo? ¿Quiénes toman qué decisiones?	
	¿Quiénes quedan excluidos del proceso decisorio?	
	¿Cuáles son las razones que se aducen?	
	¿Hay desequilibrios de poder entre <i>stakeholders</i> ?	
	¿De qué modo influye esa circunstancia en la evaluación?	
	¿Se establecen con claridad los roles de cada <i>stakeholder</i> ?	
	Las distintas perspectivas e intereses sobre el programa, ¿están adecuada y suficientemente representadas?	
	¿Cuáles son los mecanismos empleados para ello?	
	¿Se han explicitado los distintos intereses y visiones sobre el programa?	
	Carácter externo o interno de la evaluación	
	¿Cuáles son los objetivos explícitos de la evaluación?	
	¿Son los objetivos formulados los más adecuados para enjuiciar el programa público?	

UN ESQUEMA DE METAEVALUACIÓN DESDE LA CIENCIA POLÍTICA
Y EL ANÁLISIS DE POLÍTICAS PÚBLICAS

DIMENSIÓN POLÍTICA	←————→	DIMENSIÓN TÉCNICA
	¿Cuáles son los objetivos implícitos? ¿Qué intereses representan?	
	Oportunidad técnica de la evaluación con respecto al programa	
	¿Son adecuadas y suficientes la formación, experiencia y competencia técnica de los evaluadores?	
	El enfoque metodológico elegido ¿es el más apropiado para analizar y enjuiciar el programa?	
	¿Cuáles son los criterios de valor? ¿Se han hecho explícitos?	
	Las técnicas de recogida y análisis de datos ¿son las apropiadas teniendo en cuenta los objetivos de la evaluación?	
	La calidad de las técnicas elegidas ¿justifica el coste que entrañan?	
	Rigor en la fase de trabajo de campo	
	Fiabilidad, rigor y consistencia del análisis de datos	
	Las conclusiones ¿se derivan de los datos obtenidos?	
	Suficiencia y pertinencia de los recursos empleados	
	Los clientes ¿admiten comentarios y críticas con respecto al programa?	
	¿Es receptivo el equipo evaluador a las sugerencias y comentarios críticos?	
	¿Cuáles fueron las relaciones de los distintos <i>stakeholders</i> con el equipo evaluador?	
	Oportunidad política de la evaluación (en relación al contexto y al programa evaluado)	
	Grado de compromiso de los distintos <i>stakeholders</i> hacia el proceso de evaluación	
	¿Cuál es el estilo decisorio predominante? ¿Negociación y búsqueda del consenso o, por el contrario, estilo impositivo?	
	¿Cuál es la actitud del equipo evaluador con respecto a la participación de los <i>stakeholders</i> ?	
	Imparcialidad del equipo evaluador con respecto al programa y a los clientes	
	¿Se han arbitrado mecanismos para la gestión, el seguimiento del proceso evaluativo y/o para la resolución de conflictos?	
	¿Funcionan satisfactoriamente los circuitos de información y comunicación?	
	Los cambios introducidos en el plan inicial de evaluación (en su caso) ¿han sido oportunamente comunicados a los <i>stakeholders</i> ?	

DIMENSIÓN POLÍTICA		DIMENSIÓN TÉCNICA
	Transparencia con respecto a los costes del proceso de evaluación	
	Respeto al principio de confidencialidad	
	Claridad y estilo de los informes de evaluación	
	¿Existen presiones para cambiar el sentido de los datos o el de los informes de evaluación?	
	Difusión y/o publicación de los resultados de la evaluación	
	¿Se han ocultado datos importantes a ciertos <i>stakeholders</i> ?	
	Los informes de evaluación ¿incluyen información suficiente?	
	La información producida ¿ha llegado a tiempo de ser utilizada?	
	¿Proporcionan los informes una descripción y un juicio equilibrado sobre el programa u ofrecen, por el contrario, una imagen unitaria, parcial o distorsionada sobre el mismo?	
	Utilidad de la evaluación para los distintos <i>stakeholders</i>	
	¿Ha servido la evaluación –tanto si es sumativa como formativa– para mejorar o conocer con mayor profundidad el programa?	
	¿Cómo ha influido la evaluación sobre el contexto del programa?	
	Mejoras introducidas en el programa a raíz de la evaluación	
	¿Qué medidas se han adoptado para incrementar las posibilidades de utilización de las evaluaciones y propiciar el aprendizaje?	
	Impacto de la evaluación en las actitudes y percepciones acerca del programa	
	¿Cuál es el principal uso al que se ha destinado la evaluación?	
	¿Cuáles son los principales “efectos colaterales” de la evaluación?	
	¿De qué modo han condicionado los “efectos colaterales” el proceso de evaluación?	
	¿Se reconocen y admiten los principales fallos y limitaciones de la evaluación?	
	¿Es el proceso de evaluación representativo?	
	Relaciones entre evaluadores y equipo metaevaluador	
	Impacto de la metaevaluación en el proceso evaluativo	

4. LAS EVALUACIONES OBJETO DE LA META-EVALUACIÓN

La investigación que presentamos se centró en el análisis de tres de las evaluaciones de programas o proyectos educativos de reforma más significativas de las realizadas en España en el período 1982-1992.

La decisión de dirigir nuestra atención hacia el ámbito educativo estuvo motivada por el importante grado de desarrollo que en él había alcanzado la práctica evaluativa, no solo en los países del área anglosajona sino también en España, muy superior al que se registraba en el caso de otras políticas sectoriales. La delimitación temporal del estudio se justifica porque fue en la etapa 1982-1992²⁸ cuando se pusieron en marcha las primeras evaluaciones sistemáticas al hilo de los programas de reforma experimental.

La selección de los casos se hizo a partir de tres criterios convergentes:

- (i) *Que los programas objeto de las evaluaciones perteneciesen a niveles de enseñanza no universitaria*

Dos fueron las razones que aconsejaron seguir este criterio. En primer lugar, el hecho de que la evaluación de programas tuviese, en la etapa considerada, un recorrido mucho menor en la enseñanza superior que en la no universitaria²⁹. En segundo lugar, los problemas que la inclusión de evaluaciones de programas de distintos niveles educativos hubiera supuesto para el análisis, no solo por la autonomía que ser reconoce a las universidades, sino también por la especial naturaleza de las actividades que en ellas se desarrollan, por las peculiaridades de su estructura organizativa y por el perfil de sus docentes y sus estudiantes.

- (ii) *Que el diseño de dichos programas hubiese correspondido en exclusiva a la administración educativa central*

Se descartó, pues, la posibilidad de considerar también aquellas evaluaciones impulsadas por administraciones autonómicas, locales o por organismos internacionales porque las diferencias entre los contextos

²⁸ Este período coincide solo parcialmente con la etapa socialista (octubre de 1982- marzo de 1996). La investigación concluye en 1992 porque fue entonces cuando finalizó una de las evaluaciones sometidas a estudio.

²⁹ La situación en la actualidad es radicalmente distinta. La evaluación se ha instalado de forma definitiva desde hace ya algunos años en el ámbito universitario. Los años noventa fueron testigos de la puesta en marcha de diversos planes de calidad de las universidades (Consejo de Universidades 2001; 2002). Desde la aprobación de la Ley Orgánica 6/2001 de Universidades (modificada por la Ley Orgánica 4/2007), el objeto de las evaluaciones son tanto las enseñanzas, como las actividades docentes, investigadoras y de gestión del profesorado universitario y las actividades, programas, servicios y gestión de los centros e instituciones de educación superior. La Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA) y los organismos de evaluación de las comunidades autónomas son los encargados de cumplir todas estas funciones.

de elaboración y desarrollo de los programas hubieran dificultado notablemente la comparación entre casos.

- (iii) *Que las evaluaciones de los programas hubiesen sido financiadas por la administración educativa central*

El tercer y último de los criterios empleados fue la fuente de financiación de las evaluaciones. Puesto que resultaba prácticamente imposible conocer todas las evaluaciones que, cumpliendo con los anteriores requisitos, hubiesen contado con financiación pública, decidimos restringir el estudio a aquellas evaluaciones que hubieran sido financiadas por el Ministerio de Educación y Ciencia. Algunas de ellas lo fueron a través de las ayudas, los concursos y los premios a la investigación que convoca periódicamente el Centro Nacional de Investigación Educativa (CIDE)³⁰.

La aplicación de los tres criterios mencionados llevó a una primera selección de seis casos³¹. Pero razones prácticas de viabilidad del estudio hicieron aconsejable limitar aún más el número de casos que iban a ser sometidos a examen. Por ello, acudimos a las características distintivas de las evaluaciones como criterio *ad hoc*. Optamos, pues, por seguir esencialmente una estrategia de investigación intensiva con el fin de profundizar en mayor medida en cada una de las evaluaciones. Llegamos así a la selección de tres casos:

- (i) La evaluación externa de la reforma experimental de las Enseñanzas Medias (1984-1992).
- (ii) La evaluación del programa de integración escolar de alumnos con necesidades educativas especiales (1985-1989).

³⁰ Solo disponemos de datos exactos sobre el coste en el caso de la evaluación del proyecto *Atenea*. Contamos, no obstante, con las cifras del presupuesto asignado por el CIDE a las investigaciones educativas relacionadas en general con «evaluación y rendimiento» (y no solo de programas de reforma). La información puede encontrarse en GARCÍA SÁNCHEZ (2003: 143).

³¹ Nos estamos refiriendo a la evaluación externa de la reforma experimental de las Enseñanzas Medias (1984-1989), la evaluación del programa de integración escolar de alumnos con necesidades educativas especiales (1985-1989), las evaluaciones del Programa de Nuevas Tecnologías de la Información y la Comunicación (PNTIC) realizadas por la Universidad de Murcia (1987-1990) y por el Servicio de Inspección Técnica de Educación (a la que no logramos tener acceso), la evaluación de programas pertenecientes al plan experimental para la reforma de la educación infantil (1991), la evaluación de los Módulos Profesionales (1993) y la evaluación del programa «Prensa-Escuela» (1989). Todas ellas fueron financiadas por el CIDE bien a través de las ayudas, concursos y premios a la investigación que convoca periódicamente este Centro bien mediante contratación directa (caso de la evaluación del Programa «Prensa-Escuela»), a excepción de la evaluación del Programa de Nuevas Tecnologías de la Información y la Comunicación cuya financiación corrió a cargo de la Secretaría de Estado de Educación y de la evaluación del programa de integración, en la que el propio programa aportó también recursos. La relación completa de las investigaciones educativas en el ámbito de la evaluación de programas y centros financiadas por el CIDE entre 1982 y 1992 se encuentra en VÉLAZ DE MEDRANO *et al.* (1995).

- (iii) La evaluación del proyecto *Atenea* para la introducción de la informática en la escuela (1985-1990).

Las evaluaciones finalmente elegidas fueron, sin duda, las que gozaron de mayor predicamento en el ámbito evaluativo, quizá porque los programas que tenían por objeto tuvieron también una enorme relevancia. Tanto la reforma experimental de enseñanzas medias como el programa de integración pueden ser considerados, respectivamente, como el embrión de lo que fueron después la educación secundaria y la educación especial, y sus principios inspiraron en buena medida la reorganización del sistema educativo impulsada después por la Ley Orgánica para la Ordenación General del Sistema Educativo (LOGSE)³².

Más concretamente, gran parte del interés de la evaluación de la integración estriba en haber sido el fruto de la iniciativa personal de un cargo político (el responsable del programa), el mismo que se encargaría de poner en marcha más adelante el proceso de institucionalización de la evaluación y de coordinar los distintos proyectos de reforma para la reestructuración del sistema educativo.

La razón por la que se decidió considerar la evaluación del proyecto *Atenea* está relacionada con la importancia que adquirió este proyecto y la popularidad que alcanzó su evaluación entre los especialistas de la comunidad educativa. Pero, además, es relevante para nuestros propósitos por otros motivos. En primer lugar, es, junto con la relativa al programa «Prensa-Escuela», la única que fue totalmente externa³³, lo cual nos permitía analizar una cuestión de sumo interés, como es la relación cliente-evaluador. En segundo lugar, su carácter participativo nos brindaba la oportunidad de estudiar las relaciones entre actores y el modo en que ello afectaba a la viabilidad del proceso. Por último, es la única del listado inicial del que partíamos cuya financiación no provino del CIDE³⁴.

5. ELEMENTOS PARA LA APLICACIÓN DEL ESQUEMA A LOS ESTUDIOS DE CASO

5.1. Las fuentes de información y la selección de los informantes

Como suele ser habitual, la investigación recurrió tanto a fuentes primarias como a fuentes secundarias de información. La entrevista semiestructurada fue la fuente primaria por la que nos decantamos, no solo por los inconvenientes

³² Ley 1/1990, de 3 de octubre.

³³ Desde nuestro punto de vista, una evaluación es tanto más externa cuanto menor sea la responsabilidad del equipo evaluador en el diseño e implantación del programa que se somete a evaluación.

³⁴ La evaluación del proyecto *Atenea* pasa por ser un ejemplo de evaluación problemática. Fueron muchos los especialistas que nos desaconsejaron incorporar esta evaluación a nuestro estudio. A pesar de las numerosas dificultades que encontramos, y puesto que pudimos obtener aquellos datos de mayor relevancia, decidimos finalmente incluirla en la investigación.

que planteaba la utilización de otras técnicas³⁵ sino por su idoneidad para conocer las actitudes, intereses, valores, creencias y expectativas de los sujetos³⁶.

Uno de los problemas a los que tuvimos que hacer frente en las entrevistas fue el tiempo que había transcurrido desde que se realizaron las evaluaciones. Los entrevistados tuvieron que hacer un considerable esfuerzo para recordar los detalles, la *intrahistoria* de la evaluación y la imagen que de la misma tenían *en aquellos momentos*. Y es que no se trataba de conocer su opinión en el momento de la entrevista sino de averiguar cuál era su percepción cuando estaban implicados en el proceso evaluativo. El reto fue conseguir que los entrevistados intentasen discriminar sus opiniones actuales de las pasadas y que no nos transmitiesen la interpretación que, con el paso del tiempo, habían hecho de la evaluación.

La estrategia seguida para la selección de los sujetos que iban a ser entrevistados se definió en función de los objetivos de la investigación y del tipo de información que se deseaba obtener.

Se tomaron en cuenta únicamente los centros de la Comunidad de Madrid que tomaron parte en cada una de las evaluaciones, dada la imposibilidad de abarcar el elevado número de centros que participaron en los tres programas en sus respectivos ámbitos de aplicación. Para salvar, en la medida de lo posible, las deficiencias que esto podía acarrear se incluyeron en la selección centros de distinta naturaleza (públicos y privados) y de diferentes zonas de la Comunidad de Madrid. Si bien no es posible afirmar que las opiniones de los entrevistados representan (estadísticamente hablando) las de todo el colectivo (a nivel nacional) al que pertenecen, entendemos que son significativas desde el momento en que pueden apuntar o sugerir una tendencia más que probable.

Una vez identificados los centros participantes en los programas objeto de evaluación³⁷, procedimos a contactar con los profesores que habían tomado parte en las evaluaciones. En este punto encontramos algunas dificultades puesto que, en ocasiones, el profesorado que había estado implicado en el proceso de evaluación ya no trabajaba en el centro y, por tanto, su localización resultaba prácticamente imposible o, simplemente, no quiso colaborar en la investigación.

Por lo que se refiere a las evaluaciones del programa de integración y del proyecto *Atenea*, interesaba conocer también las opiniones y actitudes no solo

³⁵ Como los grupos de discusión. Véase GARCÍA SÁNCHEZ (2003).

³⁶ Una de las principales ventajas de esta técnica, como es bien sabido, es que en ella se parte de un guión con el que dirigir y orientar el discurso del entrevistado que puede ser modificado o ampliado en función de cómo se desarrolle la entrevista y de las características del entrevistado. Sobre la utilidad de las entrevistas semiestructuradas véanse, por ejemplo, FIELDING (1994), ALONSO (1994) y ORTÍ (1996).

³⁷ Nos servimos, para ello, de las órdenes ministeriales por las que se autorizaba a los centros su participación en cada uno de los programas o de los listados que nos facilitaron las unidades responsables de estos. Las referencias de ambos pueden encontrarse en GARCÍA SÁNCHEZ (2003).

de los responsables de los programas, del profesorado de los centros o de los miembros de los equipos de evaluación sino de aquellos otros actores que tuvieron un papel destacado en el proceso o que se encargaron de la recogida y/o del análisis de datos, como los *monitores* o los miembros de los *equipos multi-profesionales* (EE.MM.)³⁸.

El número total de entrevistas realizadas fue de cuarenta y dos, aunque sólo treinta y seis tuvieron relación directa con los estudios de caso (diez con la evaluación del programa de reforma de enseñanzas medias, once con la relativa al programa de integración y quince con la correspondiente al proyecto *Ate-nea*). Las seis entrevistas no asociadas a los casos tuvieron una finalidad eminentemente exploratoria: a través de ellas pudimos identificar a los actores más relevantes del proceso de institucionalización de la evaluación educativa y de las evaluaciones que pretendíamos analizar³⁹.

En cuanto a las fuentes secundarias hemos de mencionar, además de los informes de evaluación, la legislación educativa y la bibliografía especializada, un conjunto significativo de documentos, entre los cuales debemos destacar los informes del Consejo Escolar del Estado o las series estadísticas publicadas por el Ministerio de Educación y Ciencia. En ocasiones fue necesario acudir directamente a las unidades responsables de los programas a fin de completar la información que proporcionaban estas series estadísticas.

5.2. Las fases de los procesos evaluativos

La identificación de las fases que conforman la evaluación constituye un buen punto de partida para la aplicación del esquema de metaevaluación. Sin perder de vista el carácter de proceso de las evaluaciones, la visión de las fases permite reducir su complejidad lo que, a su vez, hace más sencillo su análisis y facilita la exposición y narración de los hechos.

³⁸ Para contactar con estos profesionales acudimos, respectivamente, a los Centros de Profesores (CEP) y a las sedes de los equipos multiprofesionales; en este último caso, nos servimos del listado que figura en EQUIPO DE EVALUACIÓN DE LA INTEGRACIÓN (1987).

³⁹ Téngase en cuenta que, en la investigación inicial, se analizó también el proceso de institucionalización de la evaluación, que no ha sido incluido en esta obra. Durante aquella tuvimos la oportunidad de contar con la colaboración de otras personas que habían estado relacionadas con el mundo de la educación en el período estudiado –nos referimos, concretamente, a M^a ÁNGELES QUIRALTE, JOSÉ MARÍA LUXÁN y FRANCISCO RAMOS FERNÁNDEZ-TORRECILLA. Las conversaciones que con ellos mantuvimos no corresponden *strictu sensu* al patrón de una entrevista, razón por la cual no han sido contabilizadas como tales.

La relación de las personas entrevistadas, las reseñas biográficas y las fechas en que se celebraron las entrevistas figuran en el anexo II. Con los profesores, los monitores y los miembros de EE.MM., se llegó al compromiso de que no figurase en la investigación ni su nombre ni el del centro al que estaban vinculados. Las reseñas biográficas de aquellos actores o personajes relevantes que no fueron entrevistados se recogen a pie de página en los capítulos correspondientes.

Las cuestiones sobre las que la metaevaluación pretende indagar se vinculan, en la mayoría de los casos, a una de las tres fases identificadas: diseño, desarrollo y conclusión y producción de informes. No obstante, algunas de ellas –por ejemplo, la imparcialidad del equipo evaluador con respecto al programa y a los clientes– atraviesan todo el proceso y encuentran difícil encaje en una etapa en concreto, a pensar de lo cual no han sido desestimadas del análisis.

Veamos qué cuestiones y subdimensiones se asocian a cada una de las fases de las evaluaciones:

(a) *Fase de diseño*

- Naturaleza de la iniciativa con la que se puso en marcha la evaluación.
- Selección (en su caso) del equipo evaluador y las razones de tal decisión; formación, experiencia y competencia técnica de los evaluadores.
- Contenido del contrato de evaluación –en su caso.
- Descripción y contextualización del programa que se realiza en el evaluación.
- Objetivos, tanto implícitos como explícitos: cuáles fueron, cómo y por quiénes fueron formulados y en qué medida respondían a las ideas y los intereses de ciertos actores.
- Enfoque metodológico y técnicas de recogida y análisis de datos, proceso decisorio que condujo a su elección y los criterios, intereses o preferencias a los que se vinculaban.
- Actores seleccionados y roles que les fueron asignados.
- Criterios de valor empleados para enjuiciar el programa o proyecto y proceso que se siguió en su formulación.
- Previsión de metaevaluación.

(b) *Fase de desarrollo*

- Rigor en el desarrollo del trabajo de campo.
- Fiabilidad y consistencia en el análisis de datos.
- Fluidez en los flujos de comunicación.
- Actores que intervinieron en el trabajo de campo y en el análisis de la información.
- Suficiencia y pertinencia de los recursos empleados.
- Posible veto u ocultamiento de información.

(c) *Fase final*

- Congruencia entre datos, análisis y conclusiones.
- Opinión de los distintos actores sobre las conclusiones de la evaluación.

- Existencia de presiones para cambiar el sentido del análisis y/o de las conclusiones.
- Existencia de recomendaciones y destinatarios de las mismas.
- Contenidos y estilo de los informes y actores que tomaron parte en su redacción.
- Respeto al principio de confidencialidad.
- Difusión y/o publicación de los resultados de la evaluación⁴⁰.
- Valoración de los destinatarios y el personal del programa acerca de la utilidad del proceso y de los resultados de la evaluación.
- Uso al que se destinaron los resultados de la evaluación (*instrumental*⁴¹ u orientado a la mejora del programa, *persuasivo*⁴² o encaminado a convencer a ciertos actores de que apoyen o defiendan una determinada posición o *conceptual*⁴³, dirigido a promover la reflexión y la producción de conocimiento sobre el programa).
- En su caso, relación entre la evaluación y la metaevaluación.
- Impacto de la evaluación e impacto de la metaevaluación.

Siguiendo la lógica expositiva que marcan estas tres fases procederemos, en la II parte de la obra, a exponer los estudios de caso realizados, no sin antes presentar el contexto –la política educativa– en el que se fraguaron los programas objeto de las evaluaciones aquí estudiadas.

⁴⁰ Desde nuestro punto de vista es este un aspecto especialmente importante pues creemos, con ALONSO HINOJAL, que «las investigaciones y sus resultados deben ser cualquier cosa menos confidenciales; máxime si (...) se financian con recursos de la colectividad (...) “el confidencialismo”, con frecuencia practicado, es la negación de la propia investigación social en una sociedad democrática» (1982: 10).

⁴¹ Lo que LEVITON y HUGHES (1981) definen como *uso instrumental* coincide con el *uso para la rendición de cuentas y para el desarrollo del programa* de los que habla ELEANOR CHELIMSKY (1997). Para esta autora, las evaluaciones que parten de una perspectiva de rendición de cuentas tienen mayores probabilidades de ser utilizadas porque, en ellas, los evaluadores responden a las cuestiones que son del interés de los usuarios y mantienen con estos un contacto permanente a lo largo de todo el proceso. Sin embargo, las evaluaciones que se orientan al desarrollo no son tan útiles para el debate político puesto que en ellas los evaluadores forman parte de la organización y/o del programa y su juicio puede verse influido por esta circunstancia. Cuanto más independiente sea un evaluador mayor será su capacidad de intervenir y de incorporar innovaciones en el debate sobre el programa. En opinión de WHOLEY y WHITE (1973) el *uso instrumental* es el principal valor de la evaluación.

⁴² LEVITON y HUGHES (1981).

⁴³ El término *uso conceptual*, acuñado por ROSSI y FREEMAN (1989), se aproxima bastante a los términos *uso para el conocimiento* propuesto por CHELIMSKY (1997) y *enlightenment* introducido por WEISS. Esta última autora define *enlightenment* como el impacto de los resultados de la evaluación sobre la comunidad «que afecta los términos del debate, el lenguaje que se emplea y las ideas que se consideran relevantes para su resolución» (1990: 176, citado por PATTON 1997: 72).

CAPÍTULO III

DE LA REFORMA EDUCATIVA A LA EVALUACIÓN DE PROGRAMAS: LOS ESCENARIOS DE LA POLÍTICA EDUCATIVA ESPAÑOLA ENTRE 1982 Y 1996

El delicado consenso alcanzado por el grupo parlamentario de Unión de Centro Democrático (UCD) y el grupo parlamentario socialista durante el proceso de elaboración del texto constitucional¹ en torno al que sería el artículo 27, aquel que consagraba el derecho a la educación, tuvo efectos aparentemente contradictorios sobre las posiciones de estos grandes grupos políticos en los años siguientes.

Por un lado, la necesidad de diferenciarse ante el electorado les llevó a polarizar sus posturas lo que, a su vez, provocó la parálisis de casi todas las iniciativas legislativas de UCD en materia de educación². De todas ellas, la única que llegó a prosperar, y no sin dificultades, fue la Ley Orgánica del Estatuto de Centros Escolares (LOECE), norma principal de desarrollo del artículo 27. Por supuesto, la situación de minoría en que se encontraba UCD en el Congreso de los Diputados³ y las disensiones internas que afectaban a este partido no dibujaban tampoco un escenario muy favorable para sacar adelante sus propuestas⁴.

¹ En él, UCD había capitalizado las propuestas clásicas de los partidos de derecha y de ciertos grupos de interés como la Conferencia Episcopal, la Confederación Católica Nacional de Padres de Alumnos (CONCAPA) y la Federación Española de Religiosos de la Enseñanza (FERE). El Partido Socialista, por su parte, se había hecho eco de las reivindicaciones tradicionales de las formaciones políticas de izquierda y de los movimientos sociales laicos (de entre los cuales destacaban los Movimientos de Renovación Pedagógica, MRP), los sindicatos y las asociaciones de padres que defendían la enseñanza pública. Sobre CONCAPA y CEAPA véase FEITO (1991). Una breve historia de los MRP puede encontrarse en DOMÈNECH y VIÑAS (1992).

² Así lo constata BONAL (1998).

³ La distribución de escaños entre partidos en las elecciones al Congreso de los Diputados de 1 de marzo de 1979 fue la siguiente: UCD 168; PSOE 121; Partido Comunista de España (PCE) 23; Coalición Democrática (CD) 10; Unión Nacional (UN) 1; Convergencia i Unió (CiU) 8; Partido Socialista Andaluz-Partido Andalucista (PSA-PA) 5; Partido Nacionalista Vasco (PNV) 7; Herri Batasuna (HB) 3; Esquerra Republicana de Catalunya-Front Nacional (ERFN) 1; Euzkadiko Ezquerria (EE) 1; Unión del Pueblo Canario (UPC) 1; Partido Aragonés Regionalista (PAR) 1. Datos procedentes de MINISTERIO DEL INTERIOR (2010).

⁴ Como es bien sabido, la Unión de Centro Democrático fue una coalición formada por distintas familias ideológicas, que no partidos, y articulada *ad hoc* para las elecciones del 15 de junio de 1977. Integraron esta coalición los democristianos Partido Demócrata Cristiano (PDC) y Par-

Por otro, el acuerdo en torno al artículo 27 sirvió para moderar las propuestas programáticas con las que el PSOE se presentaría a las decisivas elecciones de 1982 (Fernández Mellizo-Soto 2001). Si en los años previos a la aprobación de la Constitución las propuestas socialistas se orientaban a la «extinción progresiva de la enseñanza privada subvencionada y (la) absorción paulatina del alumnado por la escuela estatal o pública», a partir del 78 la pretensión esencial pasó a ser «la consecución de un sistema educativo integrado, aunque plural (...); concebido como servicio público, aunque no monopolizado por los poderes públicos; no discriminatorio, a la vez que respetuoso con las convicciones personales y las preferencias de los padres; racionalizado en su diseño, al tiempo que abierto a la participación y extremadamente democrático» (Maravall 1984: 29).

Sobre la base de la revisión de los presupuestos ideológicos de partida, el Partido Socialista planteó en su programa electoral de 1982 dos grandes objetivos: garantizar en la práctica el derecho a la educación –lo cual exigía corregir las desigualdades existentes e incrementar la oferta de puestos escolares, en el marco de una política de escolarización plena– y mejorar la calidad de la enseñanza.

1. DEL PROGRAMA ELECTORAL AL PROGRAMA DE GOBIERNO: LAS PRIMERAS ACTUACIONES DEL MINISTERIO MARAVALL⁵

El compromiso de hacer efectivo el derecho a la educación se convirtió en la principal seña de identidad de la política del primer ministerio socialista, fuertemente liderado por José María Maravall y en el que la presencia de miembros de MRP y de profesores –sobre todo de enseñanzas medias vinculados a organizaciones sindicales, especialmente a la Unión General de Trabajadores (UGT)– era muy destacada⁶. A tal fin, se entendió imprescindible:

- (i) Poner en marcha un conjunto de medidas de promoción educativa, tendentes a hacer efectivo el principio de igualdad en la educación. Estas medidas se concretaron en una serie de programas que comenzaron a implantarse paulatinamente, en algunos casos de forma experimental, a partir de 1983.

tido Popular (PP), el Partido Socialdemócrata (PSD), la Federación Socialdemócrata (FSD), el Partido Socialdemócrata Independiente (PSI), los liberales Federación de Partidos Demócratas y Liberales (FPDL) y Partido Demócrata Popular (PDP) y los regionalistas Partido Social Liberal Andalúz (PSLA), Partido Gallego Independiente (PGI), Acción Regional Extremeña (AREX), Unión Demócrata de Murcia (UDM) y Unión Canaria (UC). Sobre el proceso de formación de UCD véanse HUNEEUS (1985), CACIAGLI (1986) y DE ESTEBAN y LÓPEZ GUERRA (1982).

⁵ La obra de ZUFIAURRE (1994) analiza el proceso de reforma educativa desde la llegada al poder del PSOE hasta 1994.

⁶ En estos primeros meses, Maravall contó con la colaboración como asesor de Julio Carabaña, quien tendría luego un papel sumamente destacado en el desarrollo práctico y en la institucionalización de la evaluación. Véanse GARCÍA SÁNCHEZ (2003).

- (ii) Proceder al desarrollo legislativo del artículo 27 de la Constitución, en una línea que superase las deficiencias de la LOECE y fuera más acorde con lo defendido tradicionalmente por el Partido Socialista. Ello se traduciría en la elaboración de la futura Ley Orgánica del Derecho a la Educación (LODE).
- (iii) Acometer la reestructuración del sistema educativo en sus distintos ciclos, niveles y etapas, así como la redefinición de sus contenidos y metodologías. La Ley de Ordenación General del Sistema Educativo (LOGSE) y su desarrollo reglamentario constituyeron los instrumentos clave de esta reestructuración.

1.1. Las medidas de promoción educativa

Las medidas de promoción educativa se centraron en ámbitos de actuación muy diversos:

- (i) El programa de *educación compensatoria*⁷ surgió con el propósito de atender las necesidades educativas de aquellos grupos de población que se hallaban en condiciones económicas y sociales más desfavorables. Este programa se concretó en diversos proyectos, como el de «Escuelas Viajeras», el de creación de centros de recursos en zonas rurales, el de atención a minorías culturales o el de formación ocupacional.
- (ii) La *ampliación de la oferta de puestos escolares* –llevada a cabo especialmente en los niveles posobligatorios en los que la demanda era mucho mayor–, contribuyó a disminuir de modo sensible la ratio profesores/alumnos.

Así, la evolución del número de centros por nivel educativo fue, durante los años ochenta, claramente favorable a los centros públicos. El incremento en enseñanzas medias fue constante, mientras que en preescolar y en Educación General Básica (EGB) se produjeron fluctuaciones a lo largo de ese período. Entre los cursos 1982-83 y 1989-90 el número de centros públicos en BUP (Bachillerato Unificado Polivalente) y COU (Curso de Orientación Universitaria) se incrementó en un 55 por ciento. Mientras, el aumento en Formación Profesional (FP) fue de solo un 14,8 por ciento.

- (iii) En lo relativo a la educación de adultos, concebida hasta entonces como un simple paso hacia la obtención del título de Graduado Escolar, se empezó a impulsar una concepción integral, que contribuyera a «la inserción productiva del adulto en el mundo laboral y a su eventual

⁷ Aprobado por Real Decreto 1174/1983, de 27 de abril.

reciclaje; a su promoción cultural, a su maduración y autodeterminación personal» (Maravall 1984: 61 y 62).

Cuadro III.1
Evolución del número de centros por nivel educativo y titularidad
(cursos 1979-80 a 1989-90)

Curso	Preescolar		EGB		BUP y COU		FP	
	Púb.	Priv.	Púb.	Priv.	Púb.	Priv.	Púb.	Priv.
1982-83	1.178	1.930	14.779	5.218	1.033	1.462	994	1.372
1983-84	1.292	1.884	14.026	5.196	1.073	1.474	978	1.419
1984-85	1.386	1.865	14.785	5.162	1.110	1.485	1.043	1.291
1985-86	1.382	1.720	14.613	4.863	1.246	1.389	1.030	1.218
1986-87	1.291	1.465	14.577	4.716	1.302	1.357	1.073	1.121
1987-88	1.252	1.458	14.468	4.565	1.420	1.363	1.081	1.149
1988-89	1.036	1.051	15.003	4.774	1.526	1.133	1.157	1.096
1989-90	1.028	1.038	14.540	4.528	1.606	1.400	1.142	1.116

Fuente: Elaboración propia a partir de COBO SUERO (1995: 1118).

- (iv) En el ámbito de la *educación especial*, uno de los espacios de atención preferente en estos años, se apostó por la integración plena de los «alumnos con necesidades educativas especiales» en el sistema educativo ordinario y se abandonó la idea de su escolarización en centros especiales, con el fin de evitar situaciones de segregación social. El concepto de integración pasaría a convertirse en el eje de las actuaciones en materia de educación especial a raíz de la puesta en marcha del programa del mismo nombre⁸.
- (v) En los niveles posobligatorios, el principal instrumento fue el programa de becas⁹, que tuvo como consecuencia un incremento en el número de beneficiarios y de las dotaciones globales efectuadas. Entre los cursos 1982-83 y 1991-92, el número de becarios en enseñanzas medias y universitarias se incrementó en cerca de 600.000 personas. En ese mismo período, el coste de las convocatorias de becas del Ministerio de Educación y Ciencia sufrió un aumento de casi un 1.000 por cien.

⁸ De la ordenación de la educación especial se ocupa el Real Decreto 334/1985, de 6 de marzo. El programa experimental de integración escolar dio comienzo con la Orden ministerial de 20 de marzo de 1985. Para mayor detalle véase el capítulo V.

⁹ Puesto en marcha con el Real Decreto 2298/1983, de 28 de julio.

Cuadro III.2
**Evolución del número de becarios y de las inversiones en becas
en enseñanzas medias y universitarias (cursos 1982-83 a 1991-92)**

Curso	Número de becarios	Coste total (MM ptas.)
1982-83	162.269	6.879
1983-84	218.456	8.631
1984-85	251.595	11.844
1985-86	381.248	17.863
1986-87	473.487	29.586
1987-88	565.166	38.904
1988-89	626.738	49.687
1989-90	640.243	55.044
1990-91	708.037	64.989
1991-92	748.037	73.273

Fuente: COBO SUERO (1995: 1261).

1.2. La elaboración de la Ley Orgánica del Derecho a la Educación

Las innovaciones que la administración socialista pretendía introducir en el sistema educativo requerían de un marco legal más sustantivo que el diseñado por las medidas de promoción educativa a las que nos hemos referido en el apartado anterior. Se trataba, en definitiva, de disponer de una ley que, desde una perspectiva progresista, desarrollara los contenidos del artículo 27 de la Constitución.

A pesar de que el Partido Socialista disponía de mayoría absoluta en el Congreso¹⁰, el proyecto de ley orgánica del derecho a la educación encontró una fuerte oposición tanto dentro como fuera del Parlamento¹¹. Las asociaciones profesionales y de padres de alumnos, los sindicatos y la Iglesia Católica, hasta entonces representadas y parcialmente eclipsadas por el protagonismo de UCD y PSOE, comenzaron a cobrar fuerza y a participar activamente en el debate. El acercamiento de posturas entre los sectores conservadores y los progresistas, que había permitido el pacto constitucional, comenzó a quebrarse de nuevo (al

¹⁰ La distribución de escaños entre partidos en las elecciones al Congreso de los Diputados de 28 de octubre de 1982 fue la siguiente: *PSOE* 202; *UCD* 11; *Alianza Popular-Partido Demócrata Popular (AP-PDP)* 107; *PCE* 4; *Centro Democrático y Social (CDS)* 2; *CiU* 12; *PNV* 8; *HB* 2; *EE* 1 y *Esquerra Republicana de Cataluña (ERC)* 1. Véase MINISTERIO DEL INTERIOR (2010b).

¹¹ E, incluso, dentro del PSOE y del propio gobierno o en sectores tradicionalmente afines, como los MRP o la Federación Española de Trabajadores de la Enseñanza de UGT (FETE-UGT). Un análisis de los argumentos que se utilizaron para convencer a los críticos y a la opinión pública de las bondades de la Ley puede encontrarse en FERNÁNDEZ MELLIZO-SOTO (2001).

igual que había sucedido con ocasión de la LOECE) a raíz de la presentación, por parte del Ministerio de Educación y Ciencia, del anteproyecto de ley ante los medios de comunicación, en junio de 1983.

Los principales puntos de la polémica fueron la financiación de la enseñanza privada y el control del ideario de los centros por parte de los poderes públicos. Las críticas de las asociaciones vinculadas a la Iglesia Católica y de los sectores conservadores, como la Confederación Española de Centros de Enseñanza (CECE), la FERE o la CONCAPA, encontraron eco en las demandas y reivindicaciones del Grupo Popular y, en menor medida, del Grupo Vasco. Unos y otros entendían que la ley atentaba contra la libertad de enseñanza –que ellos consideraban básicamente como libertad de creación de centros– y que abría la puerta al intervencionismo estatal al autorizar el control, por parte de la administración, de los centros privados sostenidos con fondos públicos.

Por su parte, los sectores progresistas, con la CEAPA a la cabeza, hicieron frente común con el partido del gobierno por la «defensa de la escuela pública». Para ambos, el principio del derecho a la educación solo tenía sentido si se concebía como principio de igualdad y, a su vez, este solo podía hacerse efectivo si la enseñanza se entendía como servicio público.

La confrontación rebasó, como ya hemos dicho, el ámbito parlamentario. En diciembre de 1983, mientras se desarrollaba el debate sobre el anteproyecto en el Congreso de los Diputados, tuvo lugar la primera manifestación contra la LODE, encabezada por los principales representantes de la FERE, la CECE y la CONCAPA.

Mientras, en el Congreso, los Grupos Popular y Vasco presentaron enmiendas a la totalidad, se retiraron de la ponencia que estudiaba la ley y el primero llegó incluso a presentar recurso de inconstitucionalidad. A pesar de ello, el pleno del Congreso aprobó finalmente la Ley por 198 votos a favor (los del Grupo Socialista y Minoría Catalana¹²) y 112 en contra (los del Grupo Mixto, Grupo Popular y Minoría Vasca).

La divergencia ideológica persistió durante el proceso de tramitación parlamentaria de la ley en la Cámara Alta. De nuevo, el Grupo Popular monopolizó el enfrentamiento al presentar enmiendas a la ley y solicitar la nulidad del informe elaborado por la ponencia, solicitud que fue desestimada por la Mesa del Senado. La actitud de Minoría Vasca sirvió, en cierta medida, para neutralizar la estrategia del Grupo Popular: los vascos decidieron sumarse a socialistas y catalanes en un intento de vencer el obstruccionismo de los populares y de acelerar la aprobación de la LODE.

La disposición conciliadora de Minoría Vasca no duró demasiado. De hecho, la Ley no obtuvo el respaldo suficiente en el Senado y tuvo que ser remitida al

¹² Para BONAL (1998) este grupo parlamentario pretendía conseguir, a cambio de su apoyo, una interpretación más «autonomista» y menos «socialista» de la ley.

Congreso para su definitiva aprobación. En marzo de 1984, casi un año después de que fuera presentada por el Ministro, la LODE fue aprobada en la Cámara Baja, por 196 votos a favor (los del Grupo Socialista y Minoría Catalana), 96 en contra (Grupo Popular y Minoría Vasca) y 6 abstenciones (Grupo Mixto).

Varios meses después, en junio de 1985, el Tribunal Constitucional¹³ se pronunció a favor de la práctica totalidad de los artículos de la LODE, declarando inconstitucional solo un precepto, el que establecía el control del ideario de los centros privados sostenidos con fondos públicos por parte de la Administración. Esta sentencia consagró además la constitucionalidad del régimen de conciertos (y de la remisión reglamentaria para su desarrollo), de los criterios de admisión de alumnos en centros sostenidos con fondos públicos, el derecho del titular a fijar un ideario para el centro (siempre con el respeto a los derechos de los otros miembros de la comunidad escolar), de las competencias de los consejos escolares y de los límites a la libertad de creación de centros.

1.3. El desarrollo de la LODE

La LODE¹⁴ constituyó, así, la concreción legal de los principios inspiradores de la política socialista en materia de educación. Esta Ley descansaba en cinco grandes postulados:

- (i) La existencia de una red mixta de centros escolares, públicos y privados, como garantía de la libertad de elección de centro en condiciones de igualdad.
- (ii) El establecimiento de un régimen de conciertos al que podían acogerse los centros privados para obtener financiación pública, siempre que cumplieran determinados requisitos¹⁵.
- (iii) La importancia concedida a la programación general de la enseñanza como medio para la consecución de una efectiva igualdad en el derecho a la educación.
- (iv) La relevancia de la participación de la comunidad educativa en la programación general de la enseñanza, fundamentalmente a través de los Consejos Escolares (del Estado, de Comunidad Autónoma, de municipio y de centro).

¹³ Sentencia del Tribunal Constitucional de 27 de junio de 1985 en recurso de inconstitucionalidad núm. 180/1984.

¹⁴ Ley Orgánica 8/1985, de 3 de julio, reguladora del Derecho a la Educación.

¹⁵ Según el Real Decreto 2375/1985, de 18 de diciembre, los centros concertados, además de contar con unos criterios de admisión similares a los establecidos para centros públicos, habían de cumplir otros requisitos, como la gratuidad de la enseñanza, la participación de padres, profesores y alumnos en el control y en la gestión del centro a través del Consejo Escolar, el respeto a la libertad de conciencia y la voluntariedad de las prácticas confesionales.

- (v) La salvaguarda de la unidad del sistema educativo. La LODE incorporaba una de las disposiciones adicionales de la LOECE según la cual correspondían al Estado la ordenación general del sistema educativo, la fijación de las enseñanzas mínimas y la alta inspección del Estado.

No obstante, ni la sentencia del Constitucional ni la entrada en vigor de la Ley consiguieron el acercamiento de posturas entre el partido en el gobierno y las organizaciones de la enseñanza privada. La elaboración de los anteproyectos de reglamento de desarrollo de la LODE volvió a provocar nuevos enfrentamientos, pese a que el Ministerio había solicitado la colaboración de todos los sectores de la comunidad educativa.

La espinosa cuestión de los conciertos educativos llegó a dividir incluso a quienes antes habían hecho frente común contra la administración. De un lado, se encontraba la Iglesia, que sugería la firma de los conciertos y, de otro, la Coordinadora Pro Libertad de Enseñanza, que agrupaba a todas las asociaciones de la enseñanza privada y abogaba por la retirada del régimen de conciertos de todos los centros privados, fueran estos confesionales o seculares y, consiguientemente, por la financiación indiscriminada de todos ellos¹⁶. Los reglamentos sobre conciertos educativos, criterios de admisión de alumnos en centros sostenidos con fondos públicos, órganos colegiados de gobierno de centros públicos y Consejo Escolar del Estado fueron finalmente publicados en diciembre de 1985¹⁷.

Una vez cumplido el primer objetivo que se había marcado, el gobierno socialista se enfrentó a la tarea de proceder a la reestructuración y modernización del sistema educativo. A partir de este momento, la mejora de la calidad de la enseñanza se fue convirtiendo, paulatinamente, en el principal objetivo de la administración educativa¹⁸.

2. LOS ANTECEDENTES DE LA LOGSE

2.1. Los programas experimentales de reforma¹⁹

La política de promoción educativa y la elaboración de la LODE no fueron las únicas medidas que se pusieron en marcha durante la primera legislatura. En 1983, y en un intento consciente de huir del clásico modelo de «reforma por decreto», se había iniciado un proceso de reforma de carácter experimental que

¹⁶ PUELLES (1991).

¹⁷ Aprobados, respectivamente, en los Reales Decretos 2375, 2376, 2377 y 2378, todos ellos de 18 de diciembre de 1985.

¹⁸ Como muy acertadamente recuerdan BONAL (1998) y BALLART (1992), ello no fue exclusivo del caso español.

¹⁹ No daremos cuenta aquí de los programas experimentales de integración y de reforma de enseñanzas medias ni del proyecto *Atenea* cuyas evaluaciones constituyen nuestro objeto de estudio. Remitimos, para ello, al lector, a los capítulos correspondientes.

afectó a determinados niveles y ciclos educativos y que constituyó la base sobre la que se forjó la posterior reestructuración del sistema realizada por la LOGSE. Precisamente, en torno a este proceso experimental surgieron las primeras evaluaciones sistemáticas de programas educativos que se realizaron en España a instancias de la administración²⁰.

La implantación de programas experimentales tuvo dos objetivos prioritarios. En primer lugar, se trataba de comprobar en la práctica y en determinados centros la aplicabilidad de nuevos currícula, antes de proceder a su implantación generalizada; en segundo lugar, promover la implicación y la participación de los principales agentes educativos en la reforma del sistema.

Al menos en los primeros años del proceso, la participación de los centros en los diversos programas era voluntaria. Los centros interesados presentaban al Ministerio de Educación y Ciencia un proyecto en el que daban cuenta de cómo procederían a la implantación del programa experimental, cuyas directrices básicas habían sido establecidas por la administración. Aquellos centros cuyo proyecto era aprobado gozaban de una serie de ayudas (para adquisición de material, mejora de instalaciones y formación de profesorado, fundamentalmente) que el Ministerio de Educación y Ciencia convocaba periódicamente.

En 1983 se inició la reforma experimental de enseñanzas medias y, un año después, la del ciclo superior de EGB; las de educación infantil y educación especial dieron comienzo en 1985, mientras que la de la formación profesional hubo de esperar al curso 1987-88.

En efecto, en 1983, el Ministerio de Educación y Ciencia, procedió a la elaboración, junto con las Comunidades Autónomas que gozaban de plenas competencias²¹, del *Anteproyecto para la reforma de la segunda etapa de la Educación General Básica*, con objeto de disponer de un documento que sirviera de base al proceso de experimentación. En él, se realizaba un diagnóstico de la situación en la que se encontraba esta etapa y se señalaban sus principales deficiencias: la existencia, a su término, de una doble titulación (certificado de escolaridad y graduado escolar), su desconexión de la enseñanza secundaria y su inadecuación a la edad de los alumnos²².

Pero fue un año después, en junio de 1984, cuando dio comienzo formalmente la experimentación en centros públicos y privados. Una orden ministe-

²⁰ La única evaluación que el Ministerio de Educación y Ciencia se planteó inicialmente fue la del programa de reforma del ciclo superior de EGB aunque, finalmente, todos los programas fueron evaluados.

²¹ Entre 1980 y 1983 se había producido el traspaso de competencias en educación a las Comunidades Autónomas de Andalucía, Canarias, Cataluña, Galicia, País Vasco y Comunidad Valenciana. Véanse, respectivamente, los Reales Decretos 3936/1982, de 29 de diciembre; 2091/1983, de 28 de julio; 2809/1980, de 3 de octubre; 1763/1982, de 24 de julio; 2808/1980, de 26 de septiembre y 2093/1983, de 28 de julio.

²² Véanse MARAVALL (1984) y MINISTERIO DE EDUCACIÓN Y CIENCIA (1987).

rial²³ recogió lo esencial de los planteamientos formulados en el *Anteproyecto*, determinó los objetivos de la innovación y abrió el plazo para la presentación de solicitudes por parte de los centros.

El programa experimental de educación infantil, iniciado en 1985, pretendía ser una tentativa de reforma de la hasta entonces denominada «educación preescolar» que carecía de objetivos y contenidos pedagógicos específicos. Su propósito era conseguir que se superase la idea de que la educación de los menores de seis años era, básicamente, un problema asistencial. El programa partía de la consideración de que el aprendizaje en los primeros años de vida es esencial para el desarrollo del niño y para evitar problemas educativos posteriores²⁴.

La experiencia²⁵, desarrollada únicamente en el *territorio MEC*²⁶, tuvo como base tanto las iniciativas innovadoras que se estaban desarrollando ya en algunos centros de preescolar como los proyectos que se presentaron a la convocatoria del Ministerio. En la primera fase participaron en el proyecto 35 centros públicos de educación preescolar dependientes del Ministerio de Educación y Ciencia y 38 guarderías y escuelas infantiles de diversas Comunidades Autónomas sin competencias en educación. En el curso académico 1986-87 el número de centros se incrementó hasta alcanzar la cifra de 104 colegios públicos y 72 escuelas infantiles. También en esa etapa se elaboró el *Anteproyecto de Marco Curricular para la Educación Infantil*, a partir de los proyectos de innovación realizados por los centros.

La experimentación en el ámbito de la *formación profesional* no dio comienzo hasta el curso académico 1987-88, aunque la existencia de los denominados *módulos profesionales* estaba prevista con anterioridad²⁷. Los módulos profesionales se definían como «un período formativo de un año de duración orientado a completar la formación del alumno en aspectos relacionados de forma inmediata con el mundo del trabajo».

²³ Nos referimos a la Orden Ministerial de 13 de junio de 1984, por la que se abría el plazo para que los centros de Educación General Básica formularan solicitud de autorización para realizar la experimentación inicial de la reforma del ciclo superior de la Educación General Básica.

²⁴ MARAVALL (1984). Véase también el estudio de evaluación de los programas del plan realizado por AGUADO ODINA (1991).

²⁵ La reforma experimental de la educación infantil se inició oficialmente con la publicación de la Orden Ministerial de 26 de abril de 1985.

²⁶ El denominado *territorio MEC* estaba formado, en aquellos momentos, por las ciudades de Ceuta y Melilla así como por las Comunidades Autónomas que no habían asumido aún competencias en educación: Castilla-León, Castilla-La Mancha, Madrid, Cantabria, La Rioja, Asturias, Aragón, Baleares, Murcia, Extremadura y Navarra. Esta última comunidad recibió el traspaso de competencias en 1990 por lo que, a partir de ese momento, dejó de pertenecer al ámbito de gestión del MEC.

²⁷ En la Orden de 21 de octubre de 1986, que afectaba a todo el segundo ciclo de enseñanza secundaria. La regulación de los módulos profesionales se efectuó por Orden de 8 de febrero de 1988.

La implantación experimental de los módulos perseguía dos objetivos básicos: por un lado, actualizar los contenidos de la FP diseñada por la Ley General de Educación (en adelante, LGE) y adecuarlos a las enseñanzas establecidas en el marco de la Comunidad Europea y, por otro, promover la capacitación profesional del alumno, de cara a su inserción en el mercado laboral. En 1992 se dio por concluida la experimentación y la aprobación de nuevos módulos en el territorio MEC²⁸. Los ya existentes continuaron su vigencia hasta ser sustituidos por los ciclos formativos previstos en la LOGSE.

Por otra parte, el Gabinete Técnico de la Secretaría General de Educación²⁹ –que fue ocupada poco después de su creación por Alfredo Pérez Rubalcaba, quien sustituyó a Joaquín Arango³⁰– se responsabilizó de la implantación experimental de ciertos programas específicos, algunos de los cuales fueron después generalizados como materias transversales del currículum. Surgieron así el programa «Prensa-Escuela», el de «Escuelas Viajeras» o los Proyectos *Atenea* y *Mercurio*³¹.

En el transcurso de este proceso de experimentación se habían producido cambios en algunos de los puestos de mayor responsabilidad dentro del Ministerio. Las nuevas Direcciones Generales creadas en la reestructuración organizativa de noviembre de 1986³², fueron ocupadas casi en su totalidad por exper-

²⁸ En las Comunidades Autónomas de Galicia, Canarias, Navarra y País Vasco la aprobación de nuevos módulos continuó hasta el curso académico 1994-95.

²⁹ La Secretaría General de Educación se creó en 1985 (Real Decreto 504/1985, de 8 de abril), con el propósito de impulsar la reforma del sistema educativo. Con ello empezó a ser evidente la clara voluntad política que inspiraba todo el proceso. En esta Secretaría –que tenía rango de Subsecretaría– se encuadraron aquellas Direcciones Generales –de Promoción Educativa, Educación General Básica y Enseñanzas Medias– cuyas funciones se hallaban más estrechamente relacionadas con los ámbitos vinculados a la reforma.

³⁰ Doctor en Ciencias Químicas, ALFREDO PÉREZ RUBALCABA comenzó su carrera en el Ministerio de Educación y Ciencia en 1982 como Director del Gabinete del Secretario de Estado de Universidades e Investigación. Fue Director General de Enseñanza Superior entre 1985 y 1986, año en que pasó a ocupar la Secretaría General de Educación hasta 1992. Fue nombrado Ministro de Educación y Ciencia en julio de 1992, cargo en el que permaneció hasta julio de 1993. Entre 1993 y 1996 ocupó la cartera de Presidencia. Ha sido también Portavoz del Grupo Parlamentario Socialista (2004-2006). Desde abril de 2006 está al frente del Ministerio del Interior. En octubre de 2010 fue nombrado Vicepresidente Primero del Gobierno y Portavoz. Es miembro de la Comisión Ejecutiva Federal del PSOE.

³¹ El programa «Prensa-Escuela» surgió a raíz del convenio firmado, en julio de 1985, por el Ministerio de Educación y Ciencia y la Asociación de Editores de Diarios Españoles (AEDE) para fomentar el uso de la prensa escrita como recurso didáctico. El programa de «Escuelas Viajeras», puesto en marcha por la Orden de 24 de enero de 1985, se orientaba a promover el conocimiento directo de determinadas rutas geográficas, históricas y culturales por parte de los alumnos del segundo ciclo de EGB. Sobre los proyectos *Atenea* y *Mercurio* véase el capítulo VI.

³² Con el Real Decreto 2352/1986, de 7 de noviembre, se procedió a una nueva reestructuración de la Secretaría General de Educación, por la que se suprimieron las direcciones generales existentes y se crearon tres nuevas unidades: la Dirección General de Promoción Educativa, la Dirección General de Centros Escolares y la Dirección General de Renovación Pedagógica. Precisamente fue esta última la verdadera protagonista de la reforma del sistema educativo –junto

tos procedentes del ámbito universitario (especialmente del campo de la psicología) y cuyas ideas sobre la reforma diferían de las de quienes hasta entonces habían impulsado los programas experimentales³³. Ello se tradujo en una mayor «tecnocratización» y regulación del proceso de reforma, en un descenso de la participación de los agentes educativos y en un mayor interés por la eficiencia y la calidad³⁴. En cualquier caso, la futura LOGSE fue heredera tanto de aquellos primeros programas experimentales como de esa otra concepción más tecnocrática impulsada por quienes llegaron al Ministerio en 1986 y que se materializó en el *Proyecto para la Reforma de la Enseñanza*.

2.2. El debate público sobre la reforma: el Proyecto para la Reforma de la Enseñanza

Los resultados de todo este período de experimentación quedarían recogidos en un documento que vio la luz en junio de 1987, el Proyecto para la Reforma de la Enseñanza. La importancia de este Proyecto radica, como subraya Puelles (1991), en que por vez primera se advierte en el discurso oficial la necesidad de coordinar los programas experimentales existentes³⁵ y de emprender una reforma global del sistema educativo, que contribuyese a la mejora de la calidad de la enseñanza: «(...) La rápida ampliación de la oferta educativa experimentada por nuestro país durante la década transcurrida desde la recuperación de la democra-

con la Subdirección General de Programas Experimentales que dependía de ella— de la mano de quien fue su primer responsable, Álvaro Marchesi Ullastres. Meses antes de la reestructuración, en junio de 1986, se habían celebrado elecciones generales y el Partido Socialista había revalidado su mayoría. La distribución de escaños en el Congreso fue la siguiente: *PSOE* 184; *Alianza Popular-Partido Demócrata Popular –Coalición Popular– (AP-PDP-PL - CP)* 105; *Izquierda Unida (IU)* 7; *CDS* 19; *CiU* 18; *PNV* 6; *HB* 5; *EE* 2; *Coalición Galega (CG)* 1; *PAR* 1; *Agrupaciones Independientes de Canarias (AIC)* 1 y *Unio Valenciana* 1. Véase MINISTERIO DEL INTERIOR (2010c).

³³ A propósito de esta cuestión afirma José María Maravall «una cosa que me parecía importante era que quien había llevado a cabo la reforma experimental no se encargase de la reforma definitiva. A mí me parecía que quien había llevado a cabo un período experimental tenía demasiados sesgos y demasiadas inercias y que había que evaluar todo lo que se había hecho con ojos nuevos (...) Esto, a las personas que habían llevado a cabo la reforma experimental no les hizo mucha gracia, pero era bastante importante, porque significaba intentar evitar algunos de los problemas que habían surgido en la reforma y, sobre todo, porque lo que quería era abordar algunos temas que la reforma todavía no había tocado».

³⁴ Así lo señalan GIMENO SACRISTÁN (1993) y también algunos de los entrevistados, como Pilar Pérez Mas o José Segovia Pérez. Una postura aún más crítica es la de ZUFIAURRE, quien afirma que «La Reforma española ha recurrido a importar modelos, ha esbozado líneas de acción amplias pero difíciles de aplicar, ha dejado hacer o más bien se ha planteado esta posibilidad para que, al menos supuestamente, fuera la experimentación la que demostrase la validez de esta intencionalidad renovadora. Ahora bien, no se ha tenido en cuenta la situación de partida de la educación en España y, por tanto, se ha chocado con claras inercias institucionales, con lo que la empresa no ha sido tan ambiciosa como se pretendía» (1994: 32).

³⁵ Recuérdese que la coordinación de las reformas experimentales correspondía a la Dirección General de Renovación Pedagógica.

cia nos ha situado en unos niveles de escolarización comparables a los de la mayoría de las naciones europeas. Y, a la vez, ha puesto de manifiesto las graves insuficiencias que aquejan a nuestra institución escolar en lo referente a la calidad de los servicios que proporciona» (Ministerio de Educación y Ciencia 1987: 3).

En opinión de Puelles (1987), a esta constatación contribuyeron decisivamente las presiones de las asociaciones estudiantiles, que venían protagonizando huelgas y manifestaciones desde principios de 1987, el informe sobre la política educativa española que la OCDE realizó en 1986 y, lo que es más importante, las evaluaciones de los planes experimentales realizadas hasta el momento³⁶.

En el *Proyecto para la Reforma de la Enseñanza*, «a un tiempo un documento de definición (y) una invitación al debate»³⁷, se presentaron las propuestas del Ministerio de Educación y Ciencia acerca de la reestructuración de los diferentes niveles educativos, en lo que sería un avance de la LOGSE³⁸ y se inició formalmente el debate que el Ministerio de Educación y Ciencia pretendía entablar con el resto de la comunidad educativa³⁹. Las propuestas ministeriales descansaban en seis grandes principios:

- (i) Un nuevo modelo de currículum que vertebrase todos los niveles del sistema educativo (desde la educación infantil hasta la universitaria) con objeto de garantizar la continuidad entre las diferentes etapas, lo suficientemente abierto y flexible como para permitir la concurrencia de todas las administraciones educativas y de los propios centros escolares en su diseño e implantación. En definitiva, lo que las autoridades educativas daban en llamar la «descentralización de las responsabilidades educativas» (Marchesi y Fierro 1990: 156).
- (ii) Una nueva ordenación del sistema educativo, en la que la educación infantil llegara hasta los seis años de edad, la educación primaria cubriera el período que va desde los seis hasta los doce años, la educación secundaria obligatoria el tramo de edad doce-quince o doce-dieciséis y el bachillerato el tramo quince-dieciocho o dieciséis-dieciocho⁴⁰.

³⁶ Sin embargo, en la práctica, el impacto de las evaluaciones fue mínimo. Sobre esta cuestión véanse los capítulos IV, V y VI.

³⁷ MINISTERIO DE EDUCACIÓN Y CIENCIA (1987: 4).

³⁸ De hecho, en el *Proyecto para la Reforma de la Enseñanza*, se preveía la elaboración de un *Libro Blanco* como base de la futura ley de ordenación del sistema educativo.

³⁹ Curiosamente, la OCDE había señalado en su *Examen de la Política Educativa española* que no se tenía «la impresión (...) de que se (hubiera) desarrollado suficientemente, ni a nivel nacional ni a nivel local, un sistema regular de consulta». Para este organismo era necesario «establecer organismos de consulta ampliamente representativos a todos los niveles, que puedan servir de fuente de consejo (...)». En su informe se destacaba, con sorpresa, que el Consejo Escolar del Estado no cumpliera este tipo de funciones. Véase OCDE (1986: 57 y ss).

⁴⁰ En el *Proyecto* se proponían dos opciones en cuanto a las modalidades de bachillerato. La primera opción contemplaba las seis especialidades del bachillerato experimental: Ciencias Hu-

- (iii) La ampliación de la escolaridad obligatoria, bien hasta los quince bien hasta los dieciséis años.
- (iv) La conveniencia de actualizar los métodos didácticos de manera que el alumno dejara de ser mero receptor de conocimientos para convertirse en sujeto activo del proceso de aprendizaje.
- (v) La introducción del principio de optatividad, a través del establecimiento de diferentes modalidades en el bachillerato, con objeto de responder a las distintas necesidades del alumnado.
- (vi) La importancia de la evaluación –entendida no solo como verificación de conocimientos sino también como análisis crítico del funcionamiento de los programas, del rendimiento de los centros y de los procesos educativos–, como instrumento imprescindible para la mejora de la calidad del sistema educativo.

En los meses que siguieron a la presentación pública del *Proyecto* se sucedieron las reuniones en torno a las propuestas ministeriales tanto en foros académicos como profesionales⁴¹, aunque la continuidad del debate se vio con frecuencia comprometida por las huelgas que los profesores llevaron a cabo durante el curso 1987-88. El propio Ministerio emplazó a los distintos sectores educativos (asociaciones de estudiantes, profesores y padres, sindicatos, organizaciones confesionales y expertos) a que participasen y se pronunciasen sobre los aspectos fundamentales de la reforma. Las valoraciones de todos estos sectores quedaron recogidas en una publicación, los *Papeles para el Debate*, que el Ministerio de Educación y Ciencia editó en 1987⁴².

La Conferencia de Consejeros de Educación de las Comunidades Autónomas, celebrada a principios de 1989, respondió al mismo propósito. Se pronunció, en términos generales, a favor del proyecto de reforma presentado por el Ministerio, aunque el tema competencial despertó cierto recelo entre los gobiernos autonómicos.

En aquellos momentos, los principios básicos de la reforma estaban lo suficientemente perfilados como para que únicamente cupiera abundar en ellos. El

manas y Sociales, Lingüística-Literaria, Ciencias de la Naturaleza, Técnico-Industrial, Técnicos de Administración y Gestión y Artes Plásticas. La segunda opción se organizaba en torno a tres especialidades: Ciencias Humanas y Sociales, Ciencias de la Naturaleza y de la Salud y Bachillerato Técnico.

⁴¹ El Ministerio proporcionó incluso ayudas económicas para la celebración de reuniones de carácter científico sobre la reforma. Muchas de estas jornadas fueron organizadas por el propio Ministerio, como la *Reunión sobre la Renovación de la Educación Secundaria y Profesional* que tuvo lugar en Madrid en junio de 1987 y a la que asistieron más de trescientos especialistas en educación, españoles y extranjeros, así como representantes de las distintas administraciones, de los sindicatos, las asociaciones profesionales, las de padres y alumnos, las organizaciones patronales y los Movimientos de Renovación Pedagógica.

⁴² Véase MINISTERIO DE EDUCACIÓN Y CIENCIA (1987a). Puede consultarse también MARCHESI y MARTÍN (1998).

marcado interés político que presidía el proceso quedó patente, en el terreno organizativo, con la creación de la Secretaría de Estado de Educación⁴³. Tras la dimisión de José María Maravall en julio de 1988, Javier Solana⁴⁴ asumió la tarea de orientar el proceso de reforma hacia la elaboración de la ley de ordenación del sistema educativo, a través de la Secretaría de Estado, de la que se había hecho cargo Alfredo Pérez Rubalcaba.

2.3. El Libro Blanco para la Reforma del Sistema Educativo

Las principales conclusiones del debate entre el ministerio y la comunidad educativa dieron lugar a un nuevo documento, el *Libro Blanco para la Reforma del Sistema Educativo*⁴⁵, que fue presentado ante los medios de comunicación por el ministro Javier Solana en abril de 1989 y constituyó la base sobre la que se elaboraría la futura LOGSE.

El *Libro Blanco* se mantenía en la línea del *Proyecto para la Reforma de la Enseñanza* en lo relativo a lo que serían los objetivos de la reforma: la mejora de la calidad de la enseñanza, la ampliación de la educación obligatoria y la reestructuración del sistema educativo. En él se concretaban, además, algunas de las cuestiones que el documento precedente había dejado abiertas, como el número de especialidades del bachillerato –finalmente reducido a tres, Ciencias Humanas y Sociales, Ciencias de la Naturaleza y de la Salud y Bachillerato Técnico– o el tramo de edad que cubriría la enseñanza secundaria posobligatoria –de los dieciséis a los dieciocho años.

Coincidiendo con la aparición del *Libro Blanco*, se presentó el *Diseño Curricular Base* (DCB). En este documento el Ministerio daba cuenta de los contenidos curriculares mínimos que debían corresponder a cada una de las etapas de la educación obligatoria y que debían ser después concretados por las administraciones autonómicas y los centros educativos a través de los denominados *proyectos curriculares*. El *Diseño Curricular Base* confirmó los planteamien-

⁴³ Real Decreto 790/1988, de 20 de julio, por el que se modifica la estructura orgánica básica del Ministerio de Educación y Ciencia. A raíz de este Real Decreto la Secretaría de Estado pasó a estar integrada por las Direcciones Generales de Centros Escolares, Formación Profesional Reglada y Promoción Educativa, Renovación Pedagógica y Coordinación y Alta Inspección.

⁴⁴ JAVIER SOLANA MADARIAGA es Doctor en Ciencias Físicas. Fue Ministro de Cultura entre 1982 y 1988, Portavoz del Gobierno de 1985 a 1988, Ministro de Educación y Ciencia entre 1988 y 1992 y Ministro de Asuntos Exteriores de 1992 a 1995. En ese año fue nombrado Secretario General de la OTAN. En 1999 fue designado Alto Representante de la Unión Europea en el ámbito de la política exterior y de seguridad común. Fue además Secretario General de la Unión Europea Occidental (UEO). Actualmente es Presidente del Centro de Economía y Geopolítica Global de ESADE y profesor visitante senior de la London School of Economics, institución por la que fue investido Doctor Honoris Causa en 2010.

⁴⁵ Véase MINISTERIO DE EDUCACIÓN Y CIENCIA (1989).

tos subyacentes a la reforma: metodología didáctica activa y constructiva⁴⁶, op-tatividad y atención a la diversidad y comprensividad⁴⁷.

Sin duda, uno de los aspectos más innovadores del *Libro Blanco* era la im-portancia que en él adquirió la evaluación como instrumento al servicio de la calidad de la enseñanza. A partir de la publicación del *Libro Blanco* el binomio calidad-evaluación pasó a convertirse en uno de los elementos centrales de la política educativa de los siguientes gobiernos socialistas.

3. LA LEY DE ORDENACIÓN GENERAL DEL SISTEMA EDUCATIVO

El *Libro Blanco* y el *Diseño Curricular Base* fueron los referentes sobre los que se elaboró el anteproyecto de la Ley de Ordenación General del Sistema Educativo. El anteproyecto venía a confirmar cuál era el objetivo esencial del Ministerio: dotar al sistema educativo de un alto grado de integración y unidad y, al tiempo, de la suficiente flexibilidad como para adecuarse a las crecientes demandas sociales, a las nuevas condiciones económico-productivas y a las peculiaridades propias de los distintos niveles y áreas curriculares y de las dife-rentes Comunidades Autónomas.

A pesar de que tanto el Consejo General del Estado como el Consejo Gene-ral de la Formación Profesional emitieron informes globalmente favorables so-bre el anteproyecto este, al igual que la LODE, encontró una fuerte oposición por parte de distintos sectores de la comunidad educativa. En esta ocasión, el gobierno hubo de enfrentarse no solo a sus tradicionales adversarios sino tam-bién a los partidos y grupos con quienes antes había hecho frente común.

El desacuerdo de Comisiones Obreras (CC.OO.) y UGT con algunos de los puntos del anteproyecto intensificó su distanciamiento con respecto al PSOE⁴⁸. Ambos sindicatos presentaron, en marzo de 1990, una propuesta alternativa al proyecto de ley ante el Consejo Escolar del Estado. En ella daban cuenta de

⁴⁶ En el constructivismo, el aprendizaje se concibe como un proceso de construcción activa por parte de quien aprende. Sobre este enfoque véase COLL *et al.* (1990).

⁴⁷ Para una exposición más detallada de estos planteamientos véanse MARCHESI y FIERRO (1990) y FIERRO (1994).

⁴⁸ Este desacuerdo había quedado ya patente durante las manifestaciones que estos sindica-tos habían protagonizado, junto con CSIF (Central Sindical Independiente y de Funcionarios) y ANPE (Asociación Nacional de Profesionales de la Enseñanza), a partir de febrero de 1988, como protesta ante la negativa del Ministerio a equipar las retribuciones de los docentes a las del resto de cuerpos de la administración. Pero, como apunta ESTEVE (1987), las demandas no eran solo económicas. El llamado *malestar docente* tenía mucho que ver con las condiciones en las que el profesorado desarrollaba su trabajo: alto nivel de exigencia por parte de los padres, presión por parte de la administración para adaptarse a los cambios curriculares sin proporcionarles la formación necesaria y escaso reconocimiento social de su labor. Sin embargo, para FERNÁNDEZ MELLIZO-Soto, el descontento de los profesores tenía sobre todo que ver con cuestiones económi-cas (2001: 340).

sus principales reivindicaciones: la elaboración de una ley de financiación de la reforma⁴⁹, la creación de cuerpo único para todos los docentes⁵⁰, la mejora de la formación del profesorado y la creación de una comisión mixta de seguimiento de la ley.

Por su parte, la Conferencia Episcopal, la CONCAPA y la FERE se centraron en cuestiones como la libertad de enseñanza, la introducción de la ética en las aulas como alternativa a la religión o la gratuidad de la educación infantil en los centros privados, algunas de las cuales habían suscitado ya polémica durante el proceso de negociación de la LODE. Estas organizaciones convocaron varias manifestaciones hasta la entrada en vigor de la ley⁵¹.

El anteproyecto provocó también el malestar entre las organizaciones estudiantiles, que demandaron una ley más progresista y convocaron varias manifestaciones en el transcurso de la elaboración de la ley⁵². La primera tuvo lugar incluso antes de que el proyecto de ley fuera aprobado por el Consejo de Ministros y remitido al Congreso de los Diputados en marzo de 1990.

A raíz de las sugerencias que el Consejo Escolar del Estado, el Consejo General de la Formación Profesional y los distintos agentes de la comunidad educativa hicieron a la propuesta del Ministerio, se introdujeron en el anteproyecto de ley algunos cambios importantes. Sin duda, el más destacado fue la reordenación de la educación secundaria, que quedó desde ese momento integrada por la educación secundaria obligatoria, la posobligatoria y la formación profesional de grado medio⁵³.

El proceso de tramitación parlamentaria de la Ley de Ordenación General del Sistema Educativo fue extraordinariamente complejo. De entrada, la admisión a trámite del proyecto de ley se vio dificultada por la presentación por parte de PP, PSA, PAR y UV de cuatro enmiendas a la totalidad que fueron finalmente rechazadas por dos tercios del Congreso de los Diputados⁵⁴.

A mediados de junio, y durante los debates en la Comisión de Educación y Cultura del Congreso, los distintos grupos parlamentarios, a excepción del PP,

⁴⁹ El coste estimado de la reforma (en el ámbito de gestión del Ministerio) ascendía a 548.804.500 millones de pesetas. Véase MINISTERIO DE EDUCACIÓN Y CIENCIA (1990).

⁵⁰ En esta línea se sitúa también GIMENO SACRISTÁN (1987).

⁵¹ Por ejemplo, las del 3 de abril y el 9 de junio de 1990.

⁵² El 26 de abril y el 10 de mayo de 1990 tuvieron lugar otras manifestaciones estudiantiles (esta última convocada junto con el profesorado).

⁵³ Otras modificaciones importantes fueron la incorporación del doctorado a los estudios de música, danza y arte dramático, el compromiso de la Administración de potenciar la educación de adultos, las mejoras en la política de perfeccionamiento y promoción del profesorado y, finalmente, la introducción de una disposición adicional que permitía la jubilación anticipada del profesorado en determinadas condiciones.

⁵⁴ En aquellos momentos, la distribución de escaños en el Congreso de los Diputados era la siguiente: *PSOE* 175; *PP* 107; *CiU* 18; *IU* 17; *CDS* 14; *PNV* 5; *HB* 4; *PA* 2; *UV* 2; *EA* 2; *EE* 2; *PAR* 1 y *AIC* 1. Véase MINISTERIO DEL INTERIOR (2010d).

llegaron a un cierto consenso en torno a algunas cuestiones decisivas: se estableció que las enseñanzas mínimas no sobrepasaran el sesenta y cinco por ciento del horario escolar y se autorizó a las Comunidades Autónomas a la expedición de títulos académicos, de acuerdo con las normas que, a estos efectos, dictara el Estado. El Ministerio de Educación y Ciencia se comprometió, asimismo, a presentar un informe anual ante la Comisión de Educación y Cultura del Congreso para dar cuenta del desarrollo de la reforma, así como a incrementar los presupuestos en educación de forma que, al concluir la implantación de la ley, el gasto público en esta materia fuera equiparable al del resto de los países comunitarios.

Tras la aprobación, el 28 de junio, del proyecto de ley por el pleno del Congreso, por 198 votos a favor, 69 en contra (PP) y 3 abstenciones, se inició el trámite en la Cámara Alta. En el transcurso del mismo se introdujeron también algunas modificaciones, aunque de menor trascendencia que las anteriores. La Comisión de Educación del Senado rechazó la propuesta de veto al proyecto de ley que había presentado el Partido Popular y remitió la LOGSE a la Cámara Baja para su definitiva aprobación.

4. CALIDAD Y EVALUACIÓN EN LA AGENDA DE GOBIERNO: LA EDUCACIÓN EN ESPAÑA DESPUÉS DE LA LOGSE

La Ley Orgánica de Ordenación General del Sistema Educativo⁵⁵, aprobada finalmente en sesión conjunta de ambas Cámaras el 3 de octubre de 1990, vino a sustituir a la Ley General de Educación de 1970, la llamada «Ley Villar Palasí»⁵⁶, como norma fundamental de ordenación del sistema educativo. Culminó, así, un largo proceso de modernización y reestructuración del sistema educativo que había comenzado casi diez años antes con la puesta en marcha de los procesos experimentales de reforma.

Pero la implantación del nuevo sistema diseñado por la LOGSE⁵⁷ precisó del desarrollo normativo de alguno de sus preceptos. Para ello se procedió a determinar los contenidos curriculares mínimos de la educación primaria y de la educación secundaria obligatoria⁵⁸, parcialmente avanzados ya en los documentos de debate y en la propia LOGSE. Las propuestas curriculares para el bachillerato y la formación profesional fueron presentadas para su debate a la

⁵⁵ Ley Orgánica 1/1990, de 3 de octubre, de Ordenación General del Sistema Educativo (BOE del 4 de octubre).

⁵⁶ Ley 14/1970, de 4 de agosto, General de Educación y Financiamiento de la Reforma Educativa.

⁵⁷ El calendario de aplicación de la LOGSE quedó aprobado en el Real Decreto 986/1991, de 14 de junio.

⁵⁸ Los «mínimos» curriculares de la educación primaria y la educación secundaria obligatoria quedaron regulados, respectivamente, en los Reales Decretos 1006/1991 y 1007/1991, de 14 de junio.

comunidad educativa. Las críticas que hicieron los sectores conservadores denunciaban el marcado sesgo tecnológico del currículum y las carencias existentes en las áreas de humanidades, filosofía y cultura clásica. La estructura del bachillerato quedó finalmente establecida un año después de la aprobación de la LOGSE, mientras que la regulación de sus enseñanzas mínimas no tuvo lugar hasta 1992⁵⁹.

El reconocimiento que la LOGSE hizo de la importancia del binomio calidad-evaluación fue, sin duda, uno de sus aspectos más destacables e innovadores. La evaluación dejó de ser un elemento periférico para convertirse, junto con la calidad, en uno de los principales objetivos de los siguientes ministros: Alfredo Pérez Rubalcaba, Gustavo Suárez Pertierra y Jerónimo Saavedra Acevedo⁶⁰.

Las actuaciones más relevantes en este terreno fueron la creación, en 1993, del Instituto Nacional de Calidad y Evaluación (INCE), la puesta en marcha, con carácter experimental, del Plan EVA de Evaluación de Centros Docentes en el curso 1991-92, y la elaboración de la Ley Orgánica de Participación, Evaluación y Gobierno de los Centros (LOPEG), en noviembre de 1995. Esta última constituyó la concreción legal de las medidas de mejora de la calidad educativa que habían sido recogidas, un año antes, en el documento *Centros educativos y calidad de la enseñanza*⁶¹.

Evidentemente, el interés por la calidad y la evaluación no fue algo exclusivo de la administración educativa. Se inscribía en la estrategia de modernización de las administraciones públicas que venía desarrollándose en España desde comienzos de la década de los noventa y tuvo su primera y más importante plasmación en 1992, con la publicación del *Plan de Modernización de la Administración del Estado*. En este documento se entendía que los mecanismos de seguimiento y evaluación son imprescindibles para «(...) responder adecuadamente a los niveles de calidad y efectividad de los servicios que exigen los ciudadanos»⁶².

⁵⁹ Reales Decretos 1700/1991, de 29 de noviembre, por el que se establece la estructura del bachillerato y 1178/1992, de 2 de octubre, por el que se establecen sus enseñanzas mínimas.

⁶⁰ Gustavo Suárez Pertierra estuvo al frente del Ministerio de Educación y Ciencia desde julio de 1993 a junio de 1995. Le sucedió Jerónimo Saavedra Acevedo, quien permaneció en el cargo desde junio de 1995 a mayo de 1996, cuando José María Aznar formó gobierno tras la victoria del Partido Popular en las elecciones generales del 3 de marzo de ese mismo año.

⁶¹ Remitimos, de nuevo, en este punto a GARCÍA SÁNCHEZ (2003; 2004).

⁶² MINISTERIO PARA LAS ADMINISTRACIONES PÚBLICAS (1992: 22). Casi nadie hubiera aventurado en aquel momento que solo catorce años después se crearía la AEVAL, la Agencia Estatal de Evaluación de las Políticas Públicas y la Calidad de los Servicios.

SEGUNDA PARTE

CAPÍTULO IV

LA EVALUACIÓN DE LA REFORMA EXPERIMENTAL DE ENSEÑANZAS MEDIAS

1. EL PROGRAMA DE REFORMA EXPERIMENTAL DE ENSEÑANZAS MEDIAS

1.1. **La situación de partida: las enseñanzas medias a comienzos de los años ochenta**

Las enseñanzas medias constituían, a principios de los años ochenta, una de las áreas del sistema educativo más necesitadas de reforma. Los cambios políticos, sociales y económicos que se habían ido produciendo en la sociedad española en la década de los setenta pusieron de relieve las deficiencias del modelo establecido por la LGE¹:

- (i) Existía, en primer lugar, un desfase entre la edad a la que concluía la escolaridad obligatoria y gratuita (catorce años)² y la edad legal de acceso al mercado de trabajo (dieciséis años).
- (ii) En ese nivel de enseñanza se tenían que conjugar elementos contradictorios. Por un lado, el currículum debía tener un carácter comprensivo y, al tiempo, responder eficazmente a la diversidad de intereses y motivaciones del alumno. Por otro, las enseñanzas medias debían tener un valor en sí mismas, con independencia de las etapas siguientes aunque, simultáneamente, tenían que servir para preparar al alumno para los niveles universitarios³.
- (iii) Se registraban unos preocupantes niveles de abandono y de fracaso escolar, sin duda debidos a la prematura elección de la opción académi-

¹ Sobre la situación de las enseñanzas medias antes de la reforma experimental pueden consultarse, entre otros, PUELLES (1987), CENTRO DE INVESTIGACIÓN, DOCUMENTACIÓN Y EVALUACIÓN (1992) y FERNÁNDEZ ENGUIA (1987).

² Recuérdese que la extensión de la escolaridad obligatoria y gratuita fue uno de los logros más destacados de la Ley Villar Palasí.

³ COLL (1988) y ZUFIAURRE (1994) han destacado especialmente estos aspectos.

ca y profesional por parte de los alumnos al finalizar la EGB, a la naturaleza excesivamente teórica de las asignaturas y al elevado número de asignaturas y horas lectivas.

Cuadro IV.1
Evolución del alumnado matriculado en BUP y COU y del alumnado que terminó COU (cursos 1983-84 a 1991-92)

	Alumnos matriculados	Alumnos que terminaron COU
1983-84	1.142.308	167.688
1984-85	1.182.154	170.929
1985-86	1.230.029	169.029
1986-87	1.265.894	171.995
1987-88	1.355.278	181.295
1988-89	1.425.777	196.570
1989-90	1.470.816	209.680
1990-91	1.500.141	215.817
1991-92	1.507.203	217.940

Fuente: Para los datos de los cursos académicos 1983-84 a 1986-87, CENTRO DE INVESTIGACIÓN, DOCUMENTACIÓN Y EVALUACIÓN (1995: 45). Para los restantes, MINISTERIO DE EDUCACIÓN Y CIENCIA (1991), (1992), (1993), (1994) y (1994a).

- (iv) El Curso de Orientación Universitaria (COU), por su marcado academicismo, quedaba convertido en la práctica en una mera prolongación del bachillerato y carecía además de un carácter verdaderamente orientador para el alumno.
- (v) La deficiente estructuración de la Formación Profesional convertía a este nivel en una enseñanza de inferior calidad a la que se veían abocados aquellos estudiantes que no podían superar el bachillerato.

1.2. El proceso de reforma

En el programa electoral con el que el PSOE se presentó a las elecciones de octubre de 1982 había quedado patente la importancia que este partido concedía a los temas educativos y, muy especialmente, a las enseñanzas medias. Las propuestas incluidas en él constituyeron, junto con las experiencias de innovación desarrolladas por diferentes centros y colectivos relacionados con la enseñanza, la base sobre la que se forjó la *reforma experimental de la educación secundaria* emprendida por el primer gobierno socialista. El ministro de Educación y Ciencia, José María Maravall, encomendó al director general de Enseñanzas Medias, José Segovia Pérez, quien había tomado parte en la elaboración del programa electoral del PSOE en materia de enseñanzas medias y con quien

tenía «una afinidad ideológica absoluta»⁴, la tarea de llevar a la práctica el proyecto de reforma. Dicho proyecto habría de convertirse en el auténtico «laboratorio de experimentación de la reforma» del sistema educativo y en uno de los elementos clave de la política del ministro Maravall.

1.2.1. *Los principios y los objetivos de la experimentación*

Los objetivos del programa de reforma quedaron recogidos en un documento –*Educación para la vida: presupuestos políticos, sociales y escolares de la REM*– que José Segovia presentó al ministro Maravall en la primavera de 1983 y en el que se contenían los principios programáticos que se entendían irrenunciables. Meses más tarde, el Ministerio haría público el documento oficial: *Hacia la reforma*⁵. El proyecto perseguía cuatro grandes propósitos:

- (i) Contribuir a la formación integral del alumno, mediante la utilización de métodos de enseñanza participativos y no memorísticos que estimularan la capacidad de reflexión crítica.
- (ii) Ampliar la escolaridad obligatoria hasta los dieciséis años y establecer una formación básica común de los catorce a los dieciséis años, incorporando al bachillerato enseñanzas de carácter profesional.
- (iii) Adaptar el sistema educativo al nuevo contexto de descentralización política y administrativa, teniendo en cuenta la responsabilidad de las Comunidades Autónomas en la ordenación y la gestión de la enseñanza.
- (iv) Promover la innovación en los centros y los equipos docentes para, de esta forma, lograr una mejor adecuación de los proyectos educativos a la realidad del alumnado y a su entorno.

En el diagnóstico realizado por el Ministerio, quedaba claro que la solución a los problemas de las enseñanzas medias pasaba no solo por su reestructuración (incluido su engarce con la EGB y con la enseñanza superior) sino también por la renovación de los enfoques didácticos tradicionales. La reforma propuesta abogaba por la utilización de aquellos métodos pedagógicos (la realización de prácticas o el trabajo en equipo) que mejor pudieran contribuir a la formación integral del alumno, al desarrollo de su creatividad y su espíritu crítico. Se partía, pues, de la idea de que el alumno debía participar de manera activa en el proceso de enseñanza-aprendizaje. En consonancia con estas premisas, el currículum era concebido como un elemento dinámico y flexible, que debía adaptarse tanto a las demandas personales y profesionales de los estudiantes como a las del entorno social y económico.

⁴ Así, por lo menos, lo recordaba el propio José Segovia en la entrevista realizada.

⁵ Véase MINISTERIO DE EDUCACIÓN Y CIENCIA (1983).

La reforma de enseñanzas medias fue la primera de carácter experimental que se puso en marcha en el ámbito educativo⁶. La experimentación –otro de los principios irrenunciables, como recuerda José Segovia– constituía para la administración una oportunidad de contrastar la teoría y la práctica y de disponer de datos fiables acerca de la viabilidad de los proyectos antes de su definitiva generalización.

Además, la voluntad del equipo ministerial y, especialmente del Director General de Enseñanzas Medias, era que el proceso contara con la participación de todos los sectores de la comunidad educativa. Por esta razón, se ofreció a los centros la posibilidad de desarrollar y concretar las directrices generales establecidas por el Ministerio mediante la implantación de sus propios proyectos de innovación educativa. A este respecto recuerda José Segovia: «No queríamos una reforma “erudita” o de “especialistas”⁷ (...) Por eso, intentamos que la reforma fuera protagonizada de verdad por el profesorado, que era quien luego la tenía que poner en marcha y mantener, y queríamos que el aula no fuera un mero lugar para la reproducción sin sentido de una cultura vacía, sino el lugar habitual de la innovación y la creación» (2010: 52).

El escenario en el que se llevó a cabo la reforma, aunque marcado por la ilusión, presentaba algunos rasgos que hacían probable su fracaso: la escasez de presupuesto⁸ (algo lógico si se tiene en cuenta la crisis económica por la que atravesaba España en aquellos momentos), las presiones y las resistencias corporativas y, cómo no, la manifiesta hostilidad de los sectores más conservadores de la enseñanza.

1.2.2. *La reforma del primer ciclo de enseñanzas medias*

En septiembre de 1983, se inició formalmente el proceso experimental en el primer ciclo de enseñanzas medias⁹ en el denominado *territorio*

⁶ El Real Decreto 2326/1983, de 13 de julio, sobre regulación de centros piloto y de experiencias en centros docentes ordinarios, sin suprimir las facultades de los Institutos de Ciencias de la Educación (ICE) y de los centros piloto, había atribuido a las direcciones generales del Ministerio la posibilidad de impulsar proyectos de innovación y permitía que los centros docentes los pusieran en marcha. Las direcciones generales asumían, además, la responsabilidad de apoyar y de evaluar las experiencias.

⁷ Que es la que, a juicio de Segovia y de otros protagonistas del proceso, se implantó tras la reorganización ministerial de 1986. Véase MENOR y MOREIRO (2010).

⁸ El propio SEGOVIA afirma a este respecto: «Cuando pienso que la estimación del coste de la Reforma de las enseñanzas medias que hicimos enseguida se elevaba a medio billón de pesetas, supimos cuál era su destino. Ninguna ley de educación en España, comenzando por la de 1970, ninguna, repito, ha llevado en su bolsillo el dinero necesario» (2010: 50).

⁹ A raíz de la Orden Ministerial de 30 de septiembre de 1983 por la que se autorizaba la experimentación de nuevos planes y programas en centros ordinarios de enseñanzas medias. Algunas de las comunidades autónomas que gozaban de competencias en educación pusieron también en marcha planes experimentales en este nivel educativo. En el País Vasco, la experimentación

*MEC*¹⁰, con la autorización ministerial de los planes experimentales presentados por los centros de bachillerato, Formación Profesional y enseñanzas integradas. El primer ciclo, que constaba de dos cursos académicos, se concebía como un tronco común para todos los alumnos que hubiesen superado la EGB. Las experiencias se referían a varias áreas temáticas: Instrumental, Ciencias de la Materia y del Hombre, Práctica y Tecnología y Artística.

En la selección de los centros participantes en la experiencia se tuvieron en cuenta criterios como su localización geográfica o la extracción social de su alumnado (con el fin de garantizar la representatividad de la muestra), el grado de adecuación de sus propuestas a los objetivos generales de la reforma y el grado de disponibilidad del profesorado. La participación en el proyecto era también voluntaria para los alumnos puesto que, en cada centro, solo seguían el plan experimental aquellos cuyos padres así lo habían decidido. En la primera convocatoria tomaron parte 24 centros, de los cuales 17 eran de titularidad pública¹¹.

La formación del profesorado tuvo un lugar muy destacado en estos primeros años de experimentación. Inicialmente, esta tarea fue asumida por la propia Dirección General de Enseñanzas Medias (DGEM), con la colaboración de la Subdirección General de Formación y Perfeccionamiento del Profesorado. Por iniciativa del Ministerio, se elaboraron y editaron documentos y materiales de apoyo, se organizaron seminarios y cursos de actualización para perfeccionar el diseño curricular y jornadas de debate sobre los resultados del proceso. La creación de los Centros de Profesores (CEP)¹² permitió descentralizar las funciones de formación de docentes. A pesar del esfuerzo realizado tanto por parte de la administración educativa como del profesorado, la formación fue siempre una cuestión problemática y llegó incluso a comprometer la generalización de la reforma.

La previsión de incorporar mecanismos de seguimiento y evaluación fue otro de los aspectos novedosos de la experimentación. Fue la propia Dirección General –a través de la Subdirección General de Ordenación Académica y de los Servicios de Inspección Educativa– la responsable de organizar la evaluación interna. El proceso fue extraordinariamente complejo. Se recabaron datos de alumnos, profesores y padres. Para la recogida de datos se recurrió, además

había comenzado en el curso 1983-84, después de haberse desarrollado algunos proyectos previos. En Cataluña, el proceso experimental inició su andadura en 1985. Sobre este y otros procesos experimentales véanse ZUFIAURRE (1994) y CENTRO DE INVESTIGACIÓN Y DOCUMENTACIÓN EDUCATIVA (1995).

¹⁰ Véase nota a pie núm. 26 del capítulo III.

¹¹ En 1989 (Orden de 16 de marzo), el Ministerio autorizó a dos nuevos centros la puesta en marcha de proyectos de innovación, de acuerdo con las directrices establecidas en la Orden de 30 de septiembre de 1983.

¹² Real Decreto 2112/1984, de 14 de noviembre. Ello fue posible gracias a la iniciativa de la entonces subdirectora de Formación y Perfeccionamiento del Profesorado, Pilar Pérez Mas.

de a las pruebas estandarizadas en el caso de los alumnos y docentes, a la organización de visitas, jornadas y reuniones de trabajo a las que asistían, además del profesorado implicado en el proceso experimental, los coordinadores provinciales de la reforma y los asesores de materia –que estaban vinculados a la Subdirección General antes citada–. Este proceso de evaluación interna –que, en puridad, no merecería este calificativo puesto que ni siquiera era una investigación– tuvo, al contrario que la evaluación externa, una enorme repercusión sobre el desarrollo de la reforma¹³.

En un primer momento se pretendió que, a la hora de introducir en los proyectos experimentales las mejoras aconsejadas por la evaluación interna, la administración educativa interviniera solo en último extremo y que las innovaciones se difundieran de un centro a otro. Sin embargo, la insuficiente formación del profesorado¹⁴ para llevar adelante una tarea de esta envergadura y la propia voluntad de liderazgo de la administración, hicieron que los equipos de experimentación dependieran cada vez en mayor medida de esta y que el proceso fuese cada vez más centralista (González Muñoz 1987).

Por otra parte, y debido precisamente a los resultados de la evaluación interna, se puso en evidencia la necesidad de introducir algunas modificaciones con respecto a lo establecido en un comienzo. Así, dos años después de iniciarse la experimentación, se revisaron parcialmente los objetivos generales que debía alcanzar el alumno al finalizar el primer ciclo, se fijó el número máximo de horas lectivas a la semana y se establecieron las cuatro grandes áreas del primer ciclo de enseñanzas medias: Instrumental, Ciencias de la Materia y del Hombre, Tecnológica y Artística¹⁵.

1.2.3. *La reforma del segundo ciclo*

En 1985 se procedió a la aprobación del segundo ciclo de enseñanzas medias experimentales¹⁶. Este ciclo, también de dos años de duración, constaba de

¹³ Cfr. *infra.*, apartado 2. Además de estos dos procesos evaluativos, se llevó a cabo un tercer estudio de evaluación desde una perspectiva sociológica y bastante más crítica que la adoptada en la evaluación externa. La investigación analizaba el proceso de implantación de la reforma en diez centros de la Comunidad de Madrid (ocho experimentales y dos no experimentales), siguiendo una metodología esencialmente cualitativa. Véase FERNÁNDEZ ENGUIITA *et al.* (1985).

¹⁴ Aspetto este en el que coinciden, además de la autora citada, todos los profesores de enseñanzas medias entrevistados.

¹⁵ Orden de 7 de noviembre de 1985. En ella se afirmaba que «El seguimiento y la evaluación de las experiencias autorizadas por Orden de 30 de septiembre de 1983 (...) han puesto de manifiesto la conveniencia de modificar la hipótesis sometida a experimentación». La evaluación a la que se refiere esta Orden es, lógicamente, la interna.

¹⁶ Orden de 19 de noviembre de 1985 por la que se autoriza la realización de las experiencias que se definen en centros ordinarios de enseñanzas medias. En esta misma Orden se preveía la existencia de un módulo profesional de un año de duración. Los módulos profesionales quedarían finalmente regulados en la Orden de 8 de febrero de 1988.

una parte común para todos los alumnos, de una parte específica y obligatoria para cada una de las modalidades de bachillerato («Lingüístico», «Ciencias Sociales y Humanas», «Ciencias de la Naturaleza», «Técnico-administrativo» y «Técnico-industrial») y de un bloque de asignaturas optativas. En el curso 1986-87, primero en el que se impartió, participaron 21 centros (de los cuales 20 ya habían tomado parte en la reforma experimental del primer ciclo), lo que suponía cerca de 2.900 alumnos.

En 1986, se produjeron ciertos cambios en la estructuración del bachillerato. Uno de los objetivos perseguidos era dotar al modelo de la flexibilidad necesaria para que las comunidades autónomas pudieran adaptarlo en función de sus necesidades y particularidades. Se introdujo, además, una nueva modalidad de bachillerato –el «Bachillerato Artístico»– y se precisaron las asignaturas de cada una de las seis modalidades, tanto las de la parte común como las de la parte específica¹⁷.

La estructura del bachillerato fue modificada de nuevo en los años posteriores. La división de la comunidad educativa en torno a las modalidades que debía tener el bachillerato, la supresión de la Dirección General de Enseñanzas Medias en 1986¹⁸ y la salida del Ministerio de quienes habían estado hasta la fecha al frente de la reforma hicieron que el proyecto inicial distase bastante del que acabaría siendo establecido en la LOGSE. El modelo que salió adelante, según José Segovia, fue el de la «reforma de gabinete», el que propugnaban los psicólogos y pedagogos procedentes de la universidad que se hicieron cargo de la reforma desde mediados de la década de los ochenta y que obviaba lo realizado durante la etapa experimental: un modelo, en definitiva, más preocupado por la calidad y la eficiencia pero menos participativo que el experimental¹⁹. La reorganización ministerial y la sustitución del equipo inicial significaron la desaparición de «la idea de la reforma como un proceso de innovación vinculado directamente a la práctica cotidiana en el aula (...) se inició un proceso de alejamiento de la administración de esa práctica cotidiana en los centros y del protagonismo de los profesores en el proceso de la propia reforma» (Segovia

¹⁷ Orden de 21 de octubre de 1986, por la que se define y aprueba la experiencia relativa al segundo ciclo de enseñanza secundaria. Esta Orden derogó la anterior de 19 de noviembre de 1985. Meses antes, el Real Decreto 942/1986 había unificado los principios que regían los planes de experimentación desarrollados por los centros educativos, independientemente de la Comunidad Autónoma en que se ubicaran, del nivel de enseñanza al que se refirieran y de que dichos centros fueran ordinarios o experimentales.

¹⁸ La supresión de esa Dirección General se produjo en noviembre de 1986 (Real Decreto 2352/1986), coincidiendo con la creación (en ese mismo Real Decreto) de la Dirección General de Renovación Pedagógica, que se ocupó de impulsar definitivamente la reestructuración y la modernización del sistema educativo. Para mayor detalle véase el apartado 2.1 del capítulo III.

¹⁹ Para José Segovia, «a partir de la llegada de Álvaro Marchesi la reforma experimental no existe». Con esta apreciación coincide también Pilar Pérez Mas, para quien los responsables de la reforma a partir de 1986, «hicieron como si partieran de la nada». Segovia atribuye en parte el fracaso de la ESO a que la reforma, desde 1986, fuera impuesta, hecha «desde arriba» y sin la participación del profesorado.

2010: 67). Para Patricio de Blas, responsable de la Subdirección General de Ordenación Académica, desde 1986, «la reforma experimental de las EE.MM., como las del Ciclo Superior de EGB y de Infantil, quedaron absorbidas en un proyecto de reforma global. Ese hecho, positivo sin duda, no se vio acompañado de una estrategia adecuada de convergencia de los procesos en marcha. Los centros experimentales continuaron, cierto, su andadura, para que los alumnos pudieran concluir sus estudios por esta vía, pero ya sin el horizonte de estar contribuyendo a la conformación de un proyecto por vía experimental. Las decisiones sobre el currículum se adoptarían, en lo sucesivo, sin ese referente» (De Blas 2010: 102). Sin embargo, como apuntamos en el capítulo sobre política educativa, a juicio de Maravall, los cambios eran poco menos que imprescindibles si se querían evitar algunos de los problemas e inercias de la etapa anterior²⁰.

* * *

Los cuadros que figuran a continuación reflejan la evolución del número de alumnos que participaron en la reforma experimental de enseñanzas medias que, en el *territorio MEC*, concluyó en el curso 1992-93, al suprimirse la matrícula en los primeros cursos del primer y segundo ciclo.

Entre los cursos académicos 1985-86 y 1991-92 participaron en la reforma cerca de 440.000 alumnos. Ahora bien, el porcentaje de alumnos que cursaron el plan experimental sobre el total de alumnos matriculados en centros de educación secundaria no llegó a superar, en ningún momento, el 5 por ciento, aunque se incrementó en 4 puntos porcentuales entre 1985-86 y 1991-92. El aumento fue más acusado en el caso de los centros públicos (casi 5 puntos porcentuales) que en el de los centros privados (tan solo 2 puntos en ese mismo período).

Cuadro IV. 2
Evolución del número de alumnos matriculados en el plan experimental
(cursos 1984-85 a 1991-92)

	1984-85	1985-86	1986-87	1987-88	1988-89	1989-90	1990-91	1991-92	TOTAL
Centros públicos	–	13.121	20.669	29.088	37.068	49.387	70.241	104.264	
Centros privados	–	7.815	12.783	14.682	16.588	18.542	21.948	21.965	
Total	–	20.936	33.452	43.770	53.656	67.929	92.189	126.229	438.161

Fuente: INSTITUTO NACIONAL DE ESTADÍSTICA (2011).

²⁰ Véase nota a pie núm. 33, capítulo III.

Cuadro IV.3
Distribución porcentual del alumnado matriculado en el bachillerato experimental, por titularidad de centro (cursos 1984-85 a 1991-92)

	1984-85	1985-86	1986-87	1987-88	1988-89	1989-90	1990-91	1991-92
Centros públicos	–	62,7	61,2	66,5	69,1	72,7	76,2	82,6
Centros privados	–	37,3	38,8	33,5	30,9	27,3	23,8	17,4
Total	–	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Fuente: MINISTERIO DE EDUCACIÓN Y CIENCIA (1988), (1989a), (1991), (1992), (1993), (1994) y (1994a).

Cuadro IV.4
Distribución porcentual del alumnado matriculado en todos los centros de educación secundaria y FP, por enseñanza (cursos 1984-85 a 1991-92)

	1984-85	1985-86	1986-87	1987-88	1988-89	1989-90	1990-91	1991-92
2º ciclo Educación Secundaria Obligatoria (ESO)	–	–	–	–	–	–	–	–
BUP y COU	61,9	62,9	62,8	62,8	63,0	62,4	61,3	59,7
Bachillerato LOGSE	–	–	–	–	–	–	–	–
Bachillerato experimental¹		1,0	1,6	2,0	2,4	2,9	3,8	5,0
FP I y FP II	38,1	37,1	36,7	35,2	34,6	34,7	34,6	34,7
Módulos formativos II (grado medio)	–	–	–	–	–	–	0,1	0,2
Módulos formativos II (grado superior)	–	–	–	–	–	–	0,2	0,3
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

¹ No disponemos de datos sobre el número de alumnos matriculados en el curso 1984-85, por lo que los porcentajes se han calculado sin tener en cuenta esa cifra.

Fuente: MINISTERIO DE EDUCACIÓN Y CIENCIA (1988), (1989a), (1991), (1992), (1993), (1994) y (1994a).

Cuadro IV.5

Distribución porcentual del alumnado matriculado en los centros públicos de educación secundaria y FP, por enseñanza (cursos 1984-85 a 1991-92)

	1984-85	1985-86	1986-87	1987-88	1988-89	1989-90	1990-91	1991-92
2º ciclo ESO	–	–	–	–	–	–	–	–
BUP y COU	65,6	65,2	63,9	65,0	64,5	63,5	61,8	59,7
Bach. LOGSE	–	–	–	–	–	–	–	–
Bach. experimental ¹		1,0	1,5	2,0	2,4	3,0	4,1	5,7
FP I y FP II	34,4	33,8	34,6	33,0	33,1	33,5	33,8	33,9
Módulos Formativos II (medio)	–	–	–	–	–	–	0,2	0,3
Módulos Formativos II (superior)	–	–	–	–	–	–	0,2	0,4
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

¹ No disponemos de datos sobre el número de alumnos matriculados en el curso 1984-85, por lo que los porcentajes se han calculado sin tener en cuenta esta cifra.

Fuente: MINISTERIO DE EDUCACIÓN Y CIENCIA (1988), (1989a), (1991), (1992), (1993), (1994) y (1994a).

Cuadro IV.6

Distribución porcentual del alumnado matriculado en los centros privados de educación secundaria y FP, por enseñanza (cursos 1984-85 a 1991-92)

	1984-85	1985-86	1986-87	1987-88	1988-89	1989-90	1990-91	1991-92
2º ciclo ESO	–	–	–	–	–	–	–	–
BUP y COU	55,7	55,8	57,3	58,0	59,8	59,9	60,0	59,9
Bach. LOGSE	–	–	–	–	–	–	–	–
Bach. experimental ¹		1,1	1,8	2,1	2,4	2,6	3,1	3,1
FP I y FP II	44,3	43,1	41,0	40,0	37,8	37,5	36,8	36,8
Módulos Formativos II (medio)	–	–	–	–	–	–	0,0	0,1
Módulos Formativos II (superior)	–	–	–	–	–	–	0,1	0,1
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

¹ No disponemos de datos sobre el número de alumnos matriculados en el curso 1984-85, por lo que los porcentajes se han calculado sin tener en cuenta esta cifra.

Fuente: MINISTERIO DE EDUCACIÓN Y CIENCIA (1988), (1989a), (1991), (1992) y (1993), (1994) y (1994a).

1.2.4. Balance del proceso experimental

No cabe duda de que el proceso de reforma experimental de enseñanzas medias tuvo una gran trascendencia en la reciente historia de nuestro sistema

educativo. En primer lugar, por la importancia misma del nivel al que se refería y porque esta reforma constituyó el embrión de la reestructuración de la educación secundaria efectuada en la LOGSE y, en segundo lugar, por la significación política que tuvo dentro y fuera del Ministerio²¹. Ahora bien, ni el interés ni la relevancia de esta experiencia deben ocultar la complejidad y las dificultades que hubieron de salvarse durante su desarrollo. El proceso de reforma de las enseñanzas medias se vio dificultado, principalmente, por dos factores: su carácter experimental y el hecho de que cada centro pudiera poner en marcha sus propios proyectos de innovación, dentro de las coordenadas establecidas por el ministerio.

La práctica de la experimentación no contaba aún con antecedentes en España, lo cual exigió un esfuerzo considerable a todos cuantos se vieron implicados en la reforma. Incluso la dimensión participativa que se pretendió caracterizara al proceso tuvo una influencia negativa, al exigir la puesta en marcha de un «costoso mecanismo de reuniones, jornadas y comisiones de trabajo» (González Muñoz 1987: 17).

La existencia de distintos proyectos de innovación que discurrían paralelamente provocó también considerables problemas, sobre todo por el número tan reducido de efectivos que la administración puso a disposición de la reforma y por la ausencia de materiales que pudieran servir de guía al profesorado. Todos los profesores de enseñanzas medias a los que entrevistamos coinciden en afirmar que si el proceso salió finalmente adelante fue gracias al voluntarismo de docentes y coordinadores²². Estos últimos asumieron en muchas ocasiones, como recuerda uno de los entrevistados, el papel de «pantalla del Ministerio. Eran los que daban la cara ante las bases y luego muchas veces se encontraban sin el respaldo presupuestario e institucional necesario». Como suele suceder en los procesos experimentales, la desmotivación de los participantes fue creciendo a medida que iban disminuyendo los recursos y la reforma se iba generalizando.

2. LA EVALUACIÓN EXTERNA DE LA REFORMA EXPERIMENTAL

La evaluación externa de la reforma experimental de enseñanzas medias (1984-1992) estuvo integrada por tres estudios parciales²³. Aunque en estos se partía de idénticos objetivos y de un único diseño metodológico, existían entre los mismos algunas diferencias. La más notable, sin duda, era que cada uno de

²¹ Así lo creemos, en contra de la opinión de José Segovia, quien sostiene que para el Ministerio de Educación y Ciencia «la reforma de medias no era un tema prioritario, eran prioritarias la LRU, la LODE o la ley de la ciencia». A su juicio, esta reforma solo se convirtió en prioritaria «cuando la asumieron los psicólogos y los pedagogos».

²² Algo que también destacó José María Maravall en la entrevista realizada.

²³ Véanse ÁLVARO PAGE *et al.* (1988), (1990) y (1992).

ellos se centraba en distintas generaciones de alumnos y/o en distintos ciclos de la enseñanza experimental:

- (i) En el primero, las pruebas fueron aplicadas a la *segunda generación de alumnos que cursó el primer ciclo del plan experimental*, es decir, a aquellos que iniciaron la experimentación en el curso académico 1984-85.
- (ii) El segundo hacía referencia a los resultados que la *tercera generación de alumnos* (aquellos que comenzaron en el curso académico 1985-86) obtuvo en el *primer ciclo de enseñanzas medias*.
- (iii) El tercer y último estudio analizaba los resultados obtenidos por la *tercera generación de alumnos* (1985-86) en el *segundo ciclo de enseñanzas medias experimentales*.

Sin embargo, estas diferencias no impiden que los tres estudios puedan ser considerados, para nuestros propósitos, como una única evaluación.

2.1. El diseño de la evaluación

2.1.1. De la iniciativa a la formulación de objetivos

En 1984, un año después de que comenzara la reforma de las enseñanzas medias, desde el Centro de Investigación y Documentación Educativa (CIDE) comenzó a vislumbrarse la posibilidad de realizar una evaluación externa de ese proceso experimental. Concretamente, fue Mariano Álvaro Page, un investigador del Centro, quien presentó la propuesta a Joaquín Arango, Secretario General Técnico del Ministerio²⁴. La iniciativa de esta evaluación partió, así pues, de un actor técnico. El nuevo director del CIDE, Julio Carabaña, consciente «de las necesidades y de la urgencia política del asunto», decidió apoyar en lo esencial la propuesta de Álvaro Page convirtiéndose, de esta forma, en su principal valedor político²⁵.

Las decisiones adoptadas por Mariano Álvaro Page y Julio Carabaña se explican, en gran medida, a partir de su *posición*, sus *habilidades* –su excelente

²⁴ Recuérdese que, en esa época, este Centro dependía de la Secretaría General Técnica del Ministerio de Educación y Ciencia. MARIANO ÁLVARO recuerda así la iniciativa: «Así pues, ni corto ni perezoso, tuve la osadía de escribir una carta a Joaquín Arango (...). En dicha carta adjuntaba un diseño para la evaluación del funcionamiento de la REM, que podría catalogarse de *cuasi-experimental* (...)» (2010: 202).

²⁵ La llegada de Carabaña a la dirección del CIDE, en septiembre de 1983, para sustituir a Juan Delval, fue una feliz casualidad para la evaluación: «Para mi sorpresa (...) después de la remisión de mi misiva a Joaquín Arango, fue nombrado director del CIDE Julio Carabaña, quien me encomendó, como tarea importante a desarrollar en el futuro inmediato, la evaluación externa de la reforma de las enseñanzas medias. De inmediato, me puse manos a la obra, con gran dedicación y entusiasmo, contando con la inestimable e imprescindible ayuda del propio Julio Carabaña» (ÁLVARO PAGE 2010: 202).

formación, su sólida trayectoria profesional y su dilatada experiencia como investigadores, en especial en evaluación del rendimiento, y su conocimiento de las evaluaciones de programas educativos realizadas en otros países— y sus *sistemas de creencias* —sus ideas acerca de la evaluación como instrumento de producción de conocimiento científico al servicio de las intervenciones públicas y su coincidencia con los objetivos y medidas de la reforma.

Los *intereses* de ambos actores tuvieron también un peso decisivo. Para Julio Carabaña y para Mariano Álvaro, el proyecto tenía un evidente *interés* intelectual y profesional. Representaba una magnífica oportunidad de poner en marcha un tipo de investigación aún novedoso en España, la evaluación de programas, y de hacerlo sobre una de las reformas emblemáticas emprendidas desde el Ministerio: la reforma de las enseñanzas medias. Detrás de la iniciativa había, pues, un evidente interés científico pero no solo eso²⁶. Un estudio que se presumía de tal envergadura podía ser muy beneficioso para ambos, al permitirles afianzar su posición en el Centro y en el Ministerio y, a su vez, otorgar al CIDE el prestigio necesario para asegurar su consolidación. La evaluación resultaba, además, oportuna tanto desde el punto de vista técnico como desde el político y de ello eran plenamente conscientes todos los actores implicados.

Como es obvio, el aval de otros actores políticos que ocupaban posiciones formales estratégicas, como José Segovia y Joaquín Arango, fue requisito imprescindible para que la propuesta pudiera prosperar. Uno y otro tuvieron la *habilidad* de advertir la importancia que podría llegar a tener una evaluación para sus *intereses* políticos y para los del Ministerio. La evaluación les brindaba la posibilidad de obtener información útil para, si era preciso, reorientar el programa y legitimarlo ante quienes participaban en él y ante el resto de la comunidad educativa.

El *entorno de oportunidad* no podía ser tampoco más favorable: el clima de renovación pedagógica que se vivía en aquellos momentos redundaba positivamente en la motivación de todos cuantos trabajaban en el Ministerio²⁷ y de buena parte del profesorado y propiciaba la puesta en marcha de cualquier iniciativa innovadora (y nada más novedoso que una evaluación de un programa experimental). José Segovia expresa con claridad meridiana el espíritu y el ambiente del momento: «Nuestra llegada al Ministerio estaba impregnada de expectación: la novedad, nuestra inexperiencia, la poca trayectoria política del PSOE, que sólo llevaba seis años de vida legal... Había poco dinero, pero mucha ilusión» (2010: 41).

²⁶ Las palabras de MARIANO ÁLVARO al describir los momentos iniciales del proceso son tremendamente reveladoras: «Pensé: la reforma propuesta es plausible, desde distintos puntos de vista, pero antes de generalizarla sería conveniente comprobar, de manera empírica, si los resultados obtenidos con su puesta en marcha mejoraban los de la educación vigente en aquellos momentos» (ÁLVARO PAGE 2010: 201).

²⁷ Afirma JOSÉ SEGOVIA: «Creíamos también con una ingenuidad rayana en la estupidez (...) que un cambio en la educación acarrearía un cambio en la sociedad» (2010: 40).

Además, como ya hemos apuntado, aunque Mariano Álvaro no desempeñaba ningún cargo de confianza política, sintonizaba claramente con los principios y con el planteamiento de la reforma²⁸ y contaba con el respaldo de Carabaña (con quien había trabajado en el INCIE) quien, a su vez, se encontraba muy próximo a José María Maravall. Todas estas circunstancias alejaban la posibilidad de que el uso que se hiciera de la información obtenida resultara perjudicial para los altos cargos ministeriales. Así lo recuerda el propio Carabaña: «todo el mundo confiaba en nosotros, confiaba en que nosotros no íbamos a dedicarnos a jugar malas pasadas con los datos».

2.1.2. *La apuesta por un enfoque excluyente de evaluación por objetivos*

El objetivo formal y explícito de la evaluación de la reforma experimental de enseñanzas medias, tal y como fue definido por quien sería su director, Mariano Álvaro, y por Julio Carabaña, era comparar los resultados obtenidos, en determinadas variables, por los alumnos que seguían el plan experimental y por los que cursaban el plan establecido en la Ley General de Educación²⁹. En esta decisión, como en las restantes *decisiones configuradoras* de la evaluación, no tomaron parte ni José Segovia, responsable del programa, ni ningún otro miembro del mismo. Se trató, pues, de una decisión no negociada, sino adoptada unilateralmente por Mariano Álvaro y Julio Carabaña. Ambos actores se convirtieron, desde un primer momento, en *actores centrales* del proceso evaluativo. Si bien se puede hablar de la existencia, en el plano analítico, del *equipo de evaluación* como un actor diferenciado de naturaleza técnica (y en el que podemos incluir a los diseñadores de las pruebas psicométricas, a los aplicadores y al personal encargado del análisis de los datos), lo cierto es que dicho equipo actuó siempre siguiendo las directrices que fijaban Álvaro Page y Julio Carabaña.

La evaluación de la reforma experimental de enseñanzas medias pretendía, en definitiva, evaluar los resultados alcanzados por los alumnos en el cumplimiento de los objetivos académicos marcados por la administración educativa. Más concretamente, la evaluación externa trataba de comprobar la incidencia que el tipo de enseñanza (variable independiente o explicativa) tenía sobre el rendimiento del alumnado (variable dependiente), medido a través de pruebas objetivas y teniendo en cuenta la influencia moduladora de otras variables, algunas de las cuales podían actuar, al mismo tiempo, como variables dependientes.

Las variables sobre las cuales se efectuaron las comparaciones fueron idénticas en los tres estudios:

²⁸ En opinión de Modesto Escobar, uno de los técnicos que formaron parte del equipo de evaluación, la valoración que Álvaro Page hacía de la reforma era aún más positiva que la del propio Carabaña.

²⁹ Véase ÁLVARO PAGE *et al.* (1988).

Cuadro IV.7
**VARIABLES CONSIDERADAS EN LA EVALUACIÓN DE LA REFORMA EXPERIMENTAL
 DE ENSEÑANZAS MEDIAS**

Independiente	Moduladoras	Dependiente
Tipo de enseñanza	Rendimiento escolar (calificaciones)	Rendimiento
	Rendimiento objetivo (resultados en las pruebas aplicadas)	
	Aptitudes	
	Reacciones emocionales ante el estudio	
	Actitudes cívico-sociales	
	Personales	
	Familiares	
	Contextuales	
	Expectativas académico-profesionales	
	Didácticas	

Fuente: ÁLVARO PAGE *et al.* (1990: 29).

En línea con sus preferencias metodológicas, Mariano Álvaro y Julio Carabaña optaron por emplear un diseño de corte cuasi-experimental. Esta decisión estuvo motivada por la imposibilidad de aplicar un diseño experimental puro, ya que la participación de los centros educativos y de los alumnos en la experimentación era voluntaria³⁰.

Sin duda, una evaluación referida al diseño y la implantación hubiera proporcionado una información más rica y exhaustiva acerca del programa –y, por ello, también más comprometida– que la que se obtenía de una evaluación que atendía únicamente a los resultados de los alumnos. Es muy probable que quienes estaban más directamente implicados en el proyecto de reforma, el profesorado y la propia Dirección General de Enseñanzas Medias (a la que pertenecían no solo actores individuales de carácter político sino también técnico-burocrático), hubieran considerado una evaluación de ese tipo como una clara agresión

³⁰ Ello impedía utilizar procedimientos de aleatorización a la hora de asignar los sujetos a cada uno de los grupos. Recuérdese que, en los diseños de naturaleza experimental, el impacto del programa se determina a partir de la comparación de los resultados obtenidos por dos grupos, uno experimental y otro de control, a los cuales se asignan, respectivamente y de forma aleatoria, los individuos que son objeto del programa y aquellos que no reciben ningún tratamiento o reciben uno alternativo. El control de las variables perturbadoras o contaminantes resulta esencial para determinar cuáles de los efectos producidos son atribuibles al programa y cuáles no lo son. Cuando resulta imposible aplicar diseños de esta naturaleza –lo que sucede en la mayor parte de las investigaciones sociales– suele recurrirse a los diseños cuasi-experimentales en los que no es necesario distribuir aleatoriamente a los individuos entre los grupos. Tanto los diseños experimentales como los cuasi-experimentales gozan de gran predicamento entre los cuantitativistas por las ventajas que presentan para el control de las amenazas a la validez interna. Para mayor abundamiento remitimos a la clásica obra de CAMPBELL y STANLEY, publicada en 1973.

a sus intereses. Carabaña era plenamente consciente de ello: «Creíamos que eso (una evaluación del diseño y la implantación) lo tenían que hacer ellos, la parte interna. Primero porque (nosotros) no teníamos medios (para hacerlo) y si hubiéramos tenido los medios no lo hubiéramos hecho porque hubiera sido visto como una injerencia intolerable en la marcha de la reforma». El tipo de evaluación elegido no era quizá el más adecuado (por más que el diseño técnico fuera impecable) pero, sin duda, era el menos arriesgado desde el punto de vista político.

Ni Carabaña ni Álvaro Page se plantearon en ningún momento la posibilidad de optar por otro tipo de evaluación, y no solo por la razón antes mencionada sino porque, por su formación, eran unos claros partidarios del enfoque de evaluación por objetivos. Muy en especial para Mariano Álvaro, el éxito de un programa e, incluso, el del propio sistema educativo, han de juzgarse básicamente a través de los resultados obtenidos por los alumnos, ya sea en las pruebas académicas habituales o en los tests psicométricos aplicados de manera ocasional. La evaluación educativa se equipara, así, a la medición del rendimiento, para lo que se hace imprescindible recurrir a enfoques como los propuestos por Tyler, Schuman o Stufflebeam³¹ y a técnicas cuantitativas de recogida y análisis de datos. Dicho en otros términos, en opinión de Mariano Álvaro la evaluación de un programa y la evaluación del alumno eran, en lo esencial, intercambiables.

Obviamente, los alumnos son un elemento central en cualquier programa educativo, pero no el único. Toda evaluación que pretenda ofrecer una visión panorámica de un programa debe examinar y enjuiciar no solo el cumplimiento de ciertos objetivos por parte de los destinatarios sino también la actuación de aquellos actores encargados del diseño o implantación de dicho programa³². A

³¹ Autores que el propio Mariano Álvaro mencionó durante la entrevista. Recordemos que, en el modelo de evaluación de RALPH TYLER (1950), el principal criterio de valor de una evaluación es, precisamente, el cumplimiento de los propios objetivos del programa. El éxito de este se mide a través de los resultados y la conducta de los alumnos y en él se prima la utilización de tests y de pruebas psicométricas. Los problemas que presenta este modelo, predominante en el panorama de la evaluación durante décadas (dificultad de la identificación de objetivos del programa, consideración solo de los objetivos explícitos y formales y excesiva atención a los resultados sin tener en cuenta el proceso de implantación), llevaron a los especialistas a la búsqueda de nuevos modelos más comprensivos y cualitativos. Para SCHUMAN (1967), la evaluación ha de considerarse como una investigación aplicada acerca de la eficiencia de un programa. Según este autor, la evaluación debe entrar a conocer no solo los resultados de la intervención sino también las premisas en las que esta se sustenta, las razones de su éxito o fracaso o los medios que serían necesarios para la consecución de los objetivos. En opinión de STUFFLEBEAM (1971) la evaluación debe incorporar los cuatro elementos o dimensiones constitutivas de un programa: su contexto (C), los inputs (I), el proceso (P) y los productos (P). En este enfoque, conocido como CIPP, se insiste en que la evaluación debe analizar el programa de la forma más completa y sistemática posible y no solo a través de sus resultados. La finalidad última de la evaluación es la de proporcionar información útil para la adopción de decisiones.

³² Esta es una de las críticas a los enfoques cientifistas más extendidas entre los pluralistas. A este respecto, puede resultar de interés consultar, entre otras, las obras de HOUSE (1980), STAKE (1986), GUBA y LINCOLN (1981; 1987).

nuestro juicio, una de las principales debilidades de este proceso de evaluación radica no tanto en su excesivo cuantitativismo cuanto en la identificación del éxito del programa con un elevado rendimiento de los alumnos.

La evaluación que ahora analizamos no es, en sentido estricto, una evaluación. Ya en el II informe se reconocía que la denominada «evaluación de la reforma experimental de enseñanzas medias» era, en realidad, un análisis de *assessment*³³ puesto que en ella no se realizaban juicios de valor sino que únicamente se presentaban datos sobre la consecución de unos objetivos³⁴. Sin embargo, esos datos no permitían hacerse, como se afirma también en ese II informe, «una clara composición de lugar sobre cómo está funcionando la reforma»³⁵ sino tan solo sobre los resultados obtenidos por sus alumnos.

Una de las principales consecuencias derivadas de la elección de un enfoque de evaluación por objetivos fue, desde nuestra perspectiva, la exclusión de ciertos actores, como José Segovia, la DGEM, los centros escolares, los alumnos o las familias, de los procesos decisorios de la evaluación. La participación de José Segovia se limitó al aval político que hubo de otorgar para que la evaluación se pudiese en marcha, aunque Joaquín Arango hubiese dado ya su visto bueno y hubiera impulsado el proyecto. Ambos desempeñaron, así, un *papel activo* en el proceso de evaluación pero en modo alguno *central*. El director de la evaluación y el director del CIDE reconocen que no contemplaron, en ningún momento, la posibilidad de requerir su participación como decisores ni siquiera como informantes. No parece, sin embargo, que esta exclusión fuera una decisión consciente y premeditada con el fin de evitar la presencia de perspectivas potencialmente críticas sino que respondió únicamente a las preferencias metodológicas de los responsables de la evaluación³⁶.

El enfoque de la evaluación de la reforma de enseñanzas medias mereció una valoración bastante negativa por parte del profesorado. En su opinión, ese enfoque llevó a la obtención de unas estadísticas «excesivamente generales», que «profundizaban poco», «se quedaban en el puro dato» y «en las que resultaba difícil reconocerse». De hecho, la mayoría insistió en que hubiera preferido una metodología más cualitativa y participativa que hubiese permitido conocer con mayor profundidad el programa.

³³ El término *assessment* hace referencia, según la definición proporcionada por el *Joint Committee on Standards for Educational Evaluation* (1994), a aquellos estudios que sirven para determinar el nivel o el valor de un objeto en una determinada variable aunque, más específicamente, se emplea para hablar de la evaluación de alumnos. El término *appraisal*, por su parte, incorpora también, como la anterior, el matiz de medición pero, en esta ocasión, referido al rendimiento del trabajador. En castellano disponemos de un único vocablo, *evaluación*, para referirnos a los tres significados.

³⁴ A pesar de ello, la palabra *evaluación* se mantuvo en los tres informes.

³⁵ ÁLVARO PAGE *et al.* (1990: 21).

³⁶ Como ya se indicó en el segundo apartado del capítulo I, los enfoques cuantitativistas no conceden importancia alguna al pluralismo de los procesos evaluativos.

En el seno de equipo de evaluación hubo también quienes criticaron la pertinencia del enfoque, aunque esta circunstancia no provocara, en ningún momento, una fractura en el mismo. En opinión de Martínez Arias, el estudio era más una investigación que una evaluación y «estaba demasiado centrado en el rendimiento académico de los alumnos. En las evaluaciones no puedes fijarte únicamente en ese producto bruto, tan burdo. Los métodos cuantitativos son tremendamente útiles pero se quedan “cortos” cuando se trata de responder a preguntas de evaluación».

Otros miembros del equipo evaluador tenían una imagen bastante más positiva del enfoque a pesar de que algunos de ellos se mostrasen muy críticos con el excesivo volumen de datos recabados. Para Susana Marcos, el diseño era «impecable desde el punto de vista técnico»; según Modesto Escobar era, «dentro de los estudios cuantitativos, el mejor que se podía haber hecho».

La evaluación de la reforma experimental de enseñanzas medias mereció también críticas de expertos ajenos al proceso. A juicio de Alvira, este estudio constituye «el típico ejemplo en el que una evaluación de la implementación era necesaria. Independientemente de que la no existencia de un grupo de control equivalente plantea problemas de validez interna, las diferencias existentes entre grupos ¿deben achacarse a la reforma?, ¿cómo hacerlo si no se sabe cómo ni en qué grado se ha implementado esta en los diferentes centros experimentales? Estudiar las diferencias de resultados en una serie de tests en distintos grupos, experimentales y de comparación, es importante pero al no haber realizado una mínima evaluación de la implementación resulta difícil interpretar las diferencias y desde luego las no diferencias» (1991: 53). Angulo Rasco, por su parte, considera que esta evaluación fue «muy cuestionable», una buena muestra de las ideas imperantes en nuestra administración, según las cuales «evaluar suele limitarse a evaluar los puntos terminales del (sistema): centros de enseñanza, docentes y alumnos» (1993: 9).

2.1.3. *La aparente ausencia de criterios de valor*

La formulación de los criterios que van a ser utilizados para enjuiciar un programa es, sin duda, uno de los aspectos más importantes en cualquier proceso de evaluación. De hecho, puede afirmarse que no existe evaluación en tanto no se incorpore a la investigación algún componente de carácter valorativo.

En la evaluación de la reforma experimental de enseñanzas medias no se formularon, de forma explícita, criterios de valor. No obstante, lo cierto es que, en la práctica, el éxito de la reforma se medía a través de los resultados de los alumnos. Los evaluadores manejaban, aunque no fueran conscientes de ello y no lo hicieran explícito, un criterio de valor fundamental. La reforma era tanto mejor cuanto mejores fueran los resultados de los alumnos. De esta forma, se inducía a que los juicios derivados de las conclusiones de la evaluación se basaran tan solo en los logros del alumnado.

2.1.4. *¿Una evaluación externa?*

La evaluación de la reforma de enseñanzas medias fue formalmente definida como externa. Para Álvaro Page, «evaluar los resultados producidos por la reforma que pudiera ponerse en marcha era una función que no debería recaer en quienes fueran a responsabilizarse de su implementación, ya que serían parte interesada y, en consecuencia, el grado de objetividad que pudieran poner en la tarea sería más que discutible» (2010: 201). El hecho de que el programa y su evaluación no recayeran en la misma unidad organizativa aseguraba, a su juicio, el carácter externo de esta última. Los responsables del CIDE, sin embargo, advirtieron las críticas que esta decisión podía suscitar:

«Así, desde una cierta perspectiva, cabe objetar que el ahora denominado Centro de Investigación, Documentación y Evaluación no deja de ser parte de la administración educativa, lo que puede poner en duda su carácter externo. A esta observación se le pueden hacer dos objeciones. La primera es la ausencia de agencias de evaluación con la necesaria experiencia y tradición, a diferencia de lo que sucede en otros países. La relativa novedad de estas tareas en el nuestro explica la escasez de equipos e instituciones capaces de llevarlas a cabo con plenas garantías (aunque esta es una deficiencia que el tiempo y las nuevas circunstancias se encargarán de solucionar). La segunda es que el CIDE, siendo parte efectiva de la propia administración educativa, ha gozado de amplia autonomía para desarrollar sus tareas de evaluación. Ni el diseño de estas últimas, ni su organización, ni sus conclusiones han sufrido deformación, veto u ocultamiento por parte de los diversos órganos ministeriales. Y ello constituye garantía suficiente del carácter externo, si no de la institución, sí al menos de sus trabajos de evaluación»³⁷.

La naturaleza externa de la evaluación no es, sin embargo y en nuestra opinión, un asunto pacífico. Ciertamente no existía una vinculación directa entre los evaluadores y la unidad responsable del programa, pero no lo es menos que ambos trabajaban conforme a unas mismas directrices políticas y que el director de la evaluación era un claro partidario de la reforma. De nuevo, la mayor parte del profesorado se pronuncia en este sentido. Las palabras de uno de ellos resultan enormemente reveladoras: «Si son ellos los que hacen la reforma, ¿qué van a decir de la reforma? Yo por lo menos no tengo la sensación de que haya una independencia entre el CIDE y el Ministerio, por lo tanto se sospecha que lo que dice el CIDE va a refrendar lo que dice el equipo ministerial».

2.1.5. *El procedimiento de muestreo*

En cada uno de los tres estudios de la evaluación³⁸ se constituyeron tres grupos, uno *experimental* y dos de *control* (uno interno y otro externo). Los

³⁷ ÁLVARO PAGE *et al.* (1990: 17).

³⁸ Recuérdese que, en el primero de los estudios, se consideraron los alumnos de la segunda generación de enseñanzas medias experimentales, mientras que en los dos estudios siguientes los alumnos pertenecían a la tercera generación.

grupos experimentales estaban formados, como es obvio, por los alumnos que seguían el plan experimental. Los grupos de *control interno* se formaron con aquellos alumnos de centros experimentales que cursaban el plan antiguo, mientras que los de *control externo* se constituyeron a partir de una muestra representativa, a nivel estatal, de sujetos que seguían el plan de la Ley General del Educación de 1970 en centros no experimentales.

La inclusión de los grupos de control externo en el diseño quedaba justificada por el hecho de que en algunos centros experimentales no existía grupo de control (es decir, todos sus alumnos seguían la enseñanza experimental) y porque, aunque no hubiera sido así, el clima de experimentación que se vivía en los centros hubiera podido afectar los resultados de la comparación.

Puesto que para evaluar las dos generaciones se iban a considerar los mismos centros, la determinación de las poblaciones se pudo realizar al comienzo de la evaluación. Las poblaciones evaluadas fueron, así, idénticas en los tres estudios³⁹. Con objeto de reforzar la validez interna de la investigación, se aplicaron tanto medidas *pretest* como medidas *posttest* a los distintos grupos⁴⁰, si bien el número de mediciones varió de un estudio a otro:

Cuadro IV. 8
Mediciones *pretest* y *posttest* realizadas a las generaciones estudiadas

	Ciclos	Mediciones		
		Pretest	1 ^{er} Posttest	2 ^o Posttest
II generación (1984-85)	1 ^o	Inicio de 1 ^{er} curso (octubre-noviembre 84)	Al término de 1 ^{er} curso (octubre-noviembre 85)	Al término de 2 ^o curso (junio 86)
III generación (1985-86)	1 ^o	Inicio de 1 ^{er} curso (octubre-noviembre 85)	Al término de 2 ^o curso (mayo-junio 87)	—
	2 ^o	Inicio de 1 ^{er} curso (octubre-noviembre 85)	Al término de 2 ^o curso (mayo-junio 87)	Al término del 2 ^o curso (mayo-junio 89)

Fuente: Elaboración propia a partir de ÁLVARO PAGE, *et al.* (1988), (1990) y (1992).

³⁹ Alumnos de primer curso de enseñanzas medias experimentales de todo el territorio español (a excepción de los alumnos del País Vasco, que renunció a participar en el estudio), alumnos de primer curso del plan vigente de enseñanzas medias en centros experimentales y alumnos de primer curso del plan vigente de enseñanzas medias en centros no experimentales.

⁴⁰ Precisamente, uno de los principales problemas que se deben afrontar en una evaluación de resultados es el de garantizar su validez interna. Esto se logra siempre que sea posible determinar con precisión cuáles de los efectos producidos en la población objeto son atribuibles al programa y cuáles no lo son. La clave reside, pues, en controlar todas las explicaciones alternativas (o variables perturbadoras) cuyos efectos se pueden confundir con los producidos por el programa. La aplicación de mediciones antes y después de la intervención contribuye también a incrementar el grado de validez interna. Sobre los problemas de validez interna en los diseños de evaluación pueden consultarse, entre otros, CAMPBELL y STANLEY (1973) y ALVIRA (1991).

El diseño de la evaluación tuvo, así, carácter *longitudinal*⁴¹. Ello permitió que se efectuaran comparaciones (i) entre los tres grupos en cada una de las mediciones, (ii) entre las distintas medidas obtenidas por un mismo grupo y (iii) entre las evoluciones de los tres grupos a lo largo del tiempo.

Debido al elevado número de centros y alumnos participantes en el programa, se decidió que la aplicación de las pruebas se realizase, en algunos casos, sobre una muestra de cada uno de los tres grupos. Del diseño muestral se hicieron cargo unos expertos, a partir de las consignas establecidas por Mariano Álvaro, que fueron contratados específicamente para esa tarea. Estos expertos no conformaron, en ningún momento, un actor diferenciado sino que quedaron encuadrados en el equipo evaluador. Esta circunstancia hizo que su entrada en escena no alterase la distribución de poder en la evaluación.

El proceso de muestreo se llevó a cabo en dos fases. En una primera se procedía a la selección de los centros mientras que, en la segunda, se seleccionaban los alumnos que integrarían las muestras.

De nuevo, existieron ciertas diferencias entre los tres estudios que integraban la evaluación, que mostramos en el siguiente cuadro:

Cuadro IV.9
Tipos de muestreo empleados en cada uno de los estudios

1 ^{er} Estudio		
CC. experimentales		CC. control externo
Sin muestreo		Polietápico estratificado con afijación proporcional
AA. experimentales	AA. control interno	AA. control externo
Aleatorio sistemático		Aleatorio sistemático

2 ^o y 3 ^{er} Estudio		
CC. experimentales		CC. control externo
Sin muestreo		Estratificado por conglomerados con afijación proporcional
AA. experimentales	AA. control interno	AA. control externo
Aleatorio sistemático		Aleatorio sistemático

Fuente: Elaboración propia a partir de ÁLVARO PAGE *et al.* (1988), (1990) y (1992).

⁴¹ Precisamente, en los diseños longitudinales se realizan sucesivas mediciones a lo largo de la investigación con objeto de conocer la evolución de los sujetos de un mismo grupo o de establecer comparaciones entre la evolución de distintos grupos.

Los cuadros IV.10 y IV.11 reflejan el número de centros y de alumnos de cada una de las muestras:

Cuadro IV.10
Número de centros experimentales y no experimentales considerados en cada estudio

	CC. experimentales	CC. no experimentales (submuestra)	Total
1^{er} estudio	109	150	259
2^o estudio	109	131	240
3^{er} estudio	109	131	240

Fuente: Elaboración propia a partir de ÁLVARO PAGE *et al.* (1988), (1990) y (1992).

Cuadro IV.11
Número de alumnos considerados en cada estudio

	1 ^{er} estudio ¹		2 ^o y 3 ^{er} estudio		3 ^{er} estudio
	Muestra inicial	Muestra final	Pretest	1 ^{er} postest	2 ^o postest ²
Experimentales	3.304	2.153	2.426	1.481	604
Control interno	2.337	1.198	2.400	1.258	510
Control externo	8.424	4.256	7.018	4.064	1.520
Total	14.065	7.607³	11.844	6.803	2.634

¹ En el primer estudio se realizaron tres mediciones –una pretest y dos postests–. La muestra inicial fue la que participó en la primera medición o pretest. La muestra final es aquella que participó en las tres mediciones. Desconocemos el número de alumnos a los que se les aplicó el primer postest.

² Entre el primer y el segundo postest se produjo una reducción muy significativa del número de sujetos de la muestra. Entre las causas de mortandad pueden citarse la repetición de curso por el alumno, el abandono de los estudios y el cambio de lugar de residencia o de centro escolar. No obstante, parece que las diferencias entre las dos muestras empleadas no afectaron al poder de generalización de los resultados, tal y como se señala en el III informe de evaluación.

³ La cifra que aparece en el I informe de evaluación es de 7.818 alumnos.

Fuente: Elaboración propia a partir de ÁLVARO PAGE *et al.* (1988), (1990) y (1992).

El procedimiento de selección de las muestras se ajustó a las características de las poblaciones sobre las cuales se efectuaba el estudio. Sin embargo, el número de centros y alumnos considerados fue, como se aprecia en los cuadros anteriores, muy elevado. Ello repercutió negativamente en el desarrollo de la evaluación (al incrementarse de modo considerable la cantidad de información que se había de recabar y analizar) y, consiguientemente, en el coste de la misma⁴². En este senti-

⁴² No disponemos de datos exactos sobre este punto. Uno de los entrevistados estima que la evaluación tuvo un coste anual aproximado de 50 millones de pesetas.

do, Modesto Escobar afirma: «Quizás el estudio era demasiado ambicioso, yo creo que con una cosa menos ambiciosa se hubiese podido hacer... ahí el Ministerio se podía haber ahorrado la mitad del dinero, se podía haber gastado la mitad». La opinión de Rosario Martínez Arias es aún más crítica: «la cantidad ingente de datos que se recabaron al final terminó perjudicando a la propia información y al procedimiento de recogida. Se podían haber obtenido resultados similares acudiendo a muestras más pequeñas y mejor controladas».

2.1.6. *La elección de las técnicas de recogida y análisis de datos*

En coherencia con la perspectiva cuantitavista que dominaba el estudio, Mariano Álvaro y Julio Carabaña optaron por los tests y los cuestionarios como instrumentos para la recogida de datos.

La elección de las pruebas que iban a ser utilizadas no fue tarea fácil. Las pruebas debían servir no sólo para medir la consecución de los objetivos de la reforma sino también para comparar los resultados de los alumnos del plan experimental con los de aquellos que cursaban BUP/COU o FP.

La medición de la mayoría de las variables no planteaba grandes dificultades puesto que existían ya pruebas suficientemente contrastadas. En el caso de las variables para las que no se disponía de un test apropiado fue necesario diseñar la prueba *ex novo*, tarea para la que se contó con la colaboración de Rosario Martínez Arias. De la redacción de los ítems se hicieron cargo profesores de diferentes asignaturas, tanto de enseñanza experimental como no experimental.

Los dos cuadros siguientes muestran cuáles fueron los principales instrumentos empleados y las variables y dimensiones analizadas con cada uno de ellos⁴³:

⁴³ Puede encontrarse una descripción más detallada del contenido de las diferentes pruebas en ÁLVARO PAGE *et al.* (1988: 35 y ss.).

Cuadro IV.12
Instrumentos empleados para la recogida de datos en el 1^{er} y 2^o estudio de la evaluación

Cuestionario CIDE	Variables personales, familiares, contextuales y didácticas, calificaciones escolares, satisfacción con los estudios y actitudes cívico-sociales.
Cuestionario de estudios	Variables actitudinales, a excepción de la satisfacción con los estudios.
Pruebas para la medición de las aptitudes	TEA 3, para las aptitudes numérica, verbal y de razonamiento abstracto. DATA MR, para la capacidad de razonamiento mecánico. <i>Rotación de figuras macizas</i> , para la aptitud espacial.
Pruebas de rendimiento basadas en el currículum de EGB	Capacidad de comprensión lectora, rendimiento en ortografía, cálculo matemático y aplicación de cálculo aplicado a la resolución de problemas.
Pruebas de rendimiento basadas en los objetivos generales de la reforma de enseñanzas medias	Cuestiones generales como las técnicas de estudio, la organización del trabajo o el manejo de información.

Fuente: Elaboración propia a partir de ÁLVARO PAGE *et al.* (1988) y (1990).

Cuadro IV.13
Instrumentos empleados para la recogida de datos en el 3^{er} estudio de la evaluación

Cuestionario CIDE	Además de las variables consideradas en el 2 ^o estudio incluyó otras como el contexto socioescolar, la trayectoria académica o las aspiraciones profesionales y educativas.
Cuestionario M.O.S.A.S.	Escalas para la medición de la motivación del logro, esperanza de éxito, perseverancia, autoestima y satisfacción con los estudios.
Cuestionario de intereses profesionales	Estructura diferencial y jerarquización de los intereses profesionales de los alumnos e incidencia de los aspectos socioeconómicos en la elección.
Pruebas de aptitudes	Aptitudes verbal, numérica, comprensión mecánica, capacidad espacial, razonamiento abstracto y rapidez y precisión perceptiva.
Pruebas de rendimiento académico	Materias comunes de los planes de estudio evaluados: Lengua española, Filosofía, Matemáticas, Física, Química, Historia, Tecnología mecánica, Tecnología electrónica y Administración y Gestión.

Fuente: Elaboración propia a partir de ÁLVARO PAGE *et al.* (1992).

2.2. El desarrollo de la evaluación: la obtención de datos, el tratamiento estadístico y el análisis de la información

La dimensión y la envergadura de la evaluación experimental de la reforma de enseñanzas medias hicieron que el CIDE se viera obligado a recurrir a empresas especializadas a la hora de desarrollar el trabajo de campo, ya que este centro no disponía ni del personal ni de los medios materiales necesarios para ello⁴⁴. Los responsables de la evaluación (esencialmente Mariano Álvaro) se ocuparon de fijar las condiciones en las que habían de realizarse las pruebas. Puesto que el objetivo era establecer comparaciones entre dos tipos distintos de enseñanza, no solo era imprescindible que se aplicasen las mismas pruebas en todos los casos sino también que todas las aplicaciones se desarrollasen bajo condiciones si no idénticas sí, al menos, muy similares⁴⁵. Tampoco las empresas encargadas de la aplicación de las pruebas conformaron una entidad diferenciada al margen del equipo evaluador, razón por la cual su intervención no afectó las relaciones de poder.

Una vez concluida la realización de las pruebas, se procedía al procesamiento informático de los datos de cara a facilitar y simplificar su análisis estadístico⁴⁶. Antes de la creación de los ficheros que iban a ser utilizados los datos fueron sometidos a un proceso de depuración por parte del equipo del CIDE.

Por iniciativa del Director del CIDE, la explotación y el análisis de los datos obtenidos tras la aplicación de las pruebas correspondió a diferentes especialistas, que fueron contratados con este fin previa convocatoria de concurso público. Se pretendía evitar, de esta forma, que la interpretación de los datos se basara en una única perspectiva y paliar, en lo posible, el hecho de que solo se hubieran utilizado técnicas cuantitativas. De nuevo, la incorporación de estos expertos no sirvió para alterar el equilibrio de fuerzas en el seno de la evaluación: Mariano Álvaro y Julio Carabaña siguieron manteniendo el control de las principales decisiones.

El tratamiento estadístico de los datos obtenidos se centró en dos procedimientos generales de control de las variables contaminantes⁴⁷:

⁴⁴ En concreto, las aplicaciones de octubre de 1984 fueron realizadas por EOS y EMOPUBLICA, S.A. Las restantes fueron encomendadas a la empresa IMOP, S.A. El responsable del estudio en esta última empresa se había encargado anteriormente de dirigir el trabajo de campo en EMOPUBLICA, S.A.

⁴⁵ Sobre la organización y el desarrollo de la aplicación de las pruebas, véase GARCÍA SÁNCHEZ (2003).

⁴⁶ La grabación de los datos y la creación de los ficheros matrices fueron encargados a la empresa PROCESO DE DATOS A-2. Aunque el proceso de tratamiento informático de los datos solo se describe en el informe del segundo estudio entendemos que se procedió de manera similar en los dos estudios restantes.

⁴⁷ Nuevamente existieron ciertas diferencias entre los tres estudios. Para mayor detalle, véase GARCÍA SÁNCHEZ (2003).

- Por un lado, la constitución de los grupos experimental y de control interno no se había realizado de forma aleatoria; además, los alumnos de este último grupo sabían que estaban siendo objeto de una comparación y ello podía introducir distorsiones. Para salvar esta deficiencia se seleccionó una muestra representativa de la población escolar que cursaba el plan antiguo en centros no experimentales y de ella solo se escogió, en el diseño definitivo, a aquellos individuos que estaban más próximos en su origen, en todas las variables utilizadas, al grupo experimental.
- Por otro, ya que los grupos experimental y de control interno no se habían equiparado y que la equiparación entre el experimental y el de control externo no garantizaba la equivalencia entre ellos, se optó por *incluir como predictor y, en consecuencia, como covariable, cuando era necesario, los resultados en el pretest al estudiar los resultados de los postests*. De esta forma se conseguía equiparar los grupos comparados en su línea base.

Los centros fueron, junto con el director del programa, José Segovia, los grandes ausentes del proceso de evaluación. La participación de los primeros se limitó a su actuación como informantes, esto es, como meros *actores pasivos*. La recogida de datos les ocasionó, además, un sinnúmero de problemas. La realización de las pruebas distraía la atención de los alumnos y restaba tiempo a las actividades habituales del aula; estas circunstancias, junto con el marcado carácter cuantitativo del estudio, hicieron que su imagen de la evaluación fuera tan negativa.

2.3. Los resultados

2.3.1. *Las principales conclusiones y el estilo de los informes*⁴⁸

Para el equipo de evaluación, los resultados que arrojó la comparación entre alumnos de reforma experimental y alumnos que cursaban el plan establecido por la Ley General de Educación en centros de BUP, eran claramente favorables a la reforma:

⁴⁸ Las conclusiones generales de la investigación se recogen en el tercer y último informe de la evaluación externa. Aunque en el mismo se incluyen también los resultados obtenidos para cada una de las variables estudiadas, nos referiremos aquí únicamente a las generales porque consideramos que ello no altera de modo significativo nuestro análisis.

Cuadro IV.14

Resumen de los resultados encontrados al comparar los alumnos experimentales y de control, según tipo de centro (BUP o FP) y ciclo (primero o segundo)

	Centros de BUP		Centros de FP	
	1 ^{er} ciclo	2 ^o ciclo	1 ^{er} ciclo	2 ^o ciclo
Rendimiento académico	Mejor, experimentales	Mejor, experimentales	Igual	Mejor, control
Aptitudes	Mejor, experimentales	Mejor, experimentales	Mejor, control	Igual
Aspiraciones profesionales y educativas	Más altas, experimentales	Más altas, experimentales	Igual	Igual
Personalidad	No se midió	Igual	No se midió	Igual
Actitudes cívico-sociales	Mayor desarrollo, experimentales	Igual	Mayor desarrollo, experimentales	Igual
Metodología del profesor	Más participativa, en reforma	Más participativa, en reforma	Más participativa, en reforma	Más participativa, en reforma
Satisfacción con los estudios	Igual	Igual	Mayor, experimentales	Mayor, experimentales
Satisfacción con los profesores	Mayor, experimentales	Igual	Mayor, experimentales	Igual

Fuente: ÁLVARO PAGE *et al.* (1992: 92).

Como puede apreciarse, en los centros de BUP y para el primer ciclo de la reforma experimental, en seis de las siete dimensiones evaluadas (ya que no se hizo ninguna medición de la variable «personalidad»), las valoraciones más altas fueron las obtenidas por los grupos experimentales. En el caso del segundo ciclo, la reforma experimental superaba al plan vigente en cuatro de las ocho variables.

Por lo que se refiere a los resultados del primer ciclo en los centros que impartían Formación Profesional, se observa que en la variable «aptitudes» los resultados de los alumnos de los grupos de control, esto es, de quienes seguían el plan de la Ley del 70, eran superiores a los obtenidos por aquellos que cursaban los módulos profesionales experimentales, si bien estos obtenían puntuaciones más altas en otras cuatro variables. De nuevo, es en el segundo ciclo donde la reforma experimental conseguía los peores resultados.

En los tres informes que conforman la evaluación externa se afirma que, aunque las diferencias entre los resultados obtenidos por los dos tipos de alumnos considerados no son muy grandes, sí pueden calificarse de significativas. Sin embargo, esos resultados no eran tan concluyentes como se afirmaba, si tenemos en cuenta los propios informes y los comentarios de algunos de los

entrevistados. En este sentido, Modesto Escobar sostiene que «los resultados del estudio es que daba lo mismo con reforma que sin reforma, había unas diferencias muy, muy pequeñas entre las puntuaciones de uno y otro». Muy probablemente, la positiva valoración de los resultados está relacionada con las actitudes que quienes diseñaron la evaluación tenían con respecto al programa.

El profesorado, por su parte, señala que los buenos resultados obtenidos por los alumnos experimentales se debieron no tanto a la calidad de los nuevos *curricula* sino, fundamentalmente, a que estos alumnos recibían una mayor atención y un trato más personalizado, al hecho de que los grupos fueran más reducidos y a que la actitud de los docentes fuera claramente favorable a la reforma.

La estructura de los informes de evaluación está estrechamente relacionada con los contenidos que en ellos se abordan. En cada uno de ellos se dedica un capítulo íntegro a la descripción de la metodología empleada, así como a la exposición de los datos resultantes de la comparación entre los grupos estudiados. En los informes II y III se incluye, además, un apartado en el que se recogen las características de los sujetos estudiados. En todos los casos se incorporan un amplísimo número de tablas, cuadros y gráficos para dar cuenta de los datos más sobresalientes. Es significativo que en ninguno de los informes se describieran ni el programa de reforma ni su contexto ni tampoco el de la evaluación.

En la estructura influyó también el propio enfoque de la evaluación. Nos encontramos aquí con unos informes prototípicos de las evaluaciones de corte cuantitativista. El estilo de los informes (a excepción de los apartados dedicados a exponer las conclusiones generales) es eminentemente técnico y su lectura y comprensión requieren de ciertos conocimientos estadísticos. No se trata de informes accesibles, pensados y dirigidos al gran público o a la comunidad educativa, que cumplan con el criterio de «claridad e imparcialidad» establecido por el *Joint Committee*⁴⁹, lo que no deja de ser una prueba más del carácter que tuvo la evaluación.

2.3.2. *La utilidad y el impacto de la evaluación*

Más allá de la utilidad de sus resultados, lo cierto es que el propio proceso evaluativo suele desencadenar una serie de efectos o consecuencias que pueden ser de gran interés para los distintos actores implicados⁵⁰.

La valoración que el responsable del programa y los profesores participantes en el mismo hacen acerca de la utilidad de la evaluación de la reforma de enseñanzas medias no puede ser más negativa.

⁴⁹ Según dicho criterio, las evaluaciones han de ser comprensibles e imparciales a la hora de tratar las necesidades e intenciones de los *stakeholders*. Véase YARBROUGH *et al.* (2011)

⁵⁰ Sobre los distintos usos de la evaluación véanse las notas a pie núms. 41 y 43 del capítulo II.

Para José Segovia, uno de los grandes ausentes del proceso, la evaluación externa fue una experiencia carente de interés. Si recibió su visto bueno fue porque no interfería en el normal desarrollo del programa y porque los datos que iba a producir no podían constituir amenaza alguna. A su juicio, la evaluación interna le ganó la batalla a la evaluación externa. Las sesiones que reunían a profesores, coordinadores y Ministerio ofrecían, a juicio del Director General de Enseñanzas Medias, una información (cualitativa) más apegada a la realidad que las innumerables y frías estadísticas que proporcionaba la evaluación realizada por el equipo del CIDE. En opinión de Patricio de Blas, la evaluación externa terminó por interferir en los procesos de seguimiento y evaluación interna establecidos desde la Dirección General de Enseñanzas Medias: «la abundancia de pruebas a los alumnos fue relegando a un segundo plano las otras fuentes de información que componían, con ellas, el proceso de seguimiento (...) Poco a poco, nos invadió a todos (profesores, equipos de apoyo, responsables en la DGEM) la preocupación por comprobar si los alumnos de la reforma obtenían mejores que los de FP y BUP. Y esa obsesión nos hizo menospreciar, a mi juicio, la riquísima y positiva información que nos llegaba por los otros conductos» (De Blas 2010: 88).

Para el profesorado, la principal deficiencia que se le puede atribuir a la evaluación es la de no haber entrado a analizar el proceso de implantación. Un análisis de esta naturaleza hubiera permitido, en su opinión, obtener información «más aplicable», «que hubiera profundizado más en lo que la gente piensa y no en el puro dato». El escaso interés que, inicialmente, pudo albergar este colectivo fue disminuyendo a medida que comprobaba la poca utilidad de los resultados de la evaluación.

Pese a esta circunstancia, el profesorado colaboró sin mayores problemas en la investigación. Para ellos, la evaluación no pasaba de ser una cuestión secundaria y los problemas que les ocasionaba no tenían la suficiente entidad como para enfrentarles a la administración. Seguramente, su actitud hubiese sido mucho más combativa si hubiesen sido «objeto de evaluación» y si esta hubiera entrado en conflicto con sus intereses profesionales.

Por el contrario, el impacto que la evaluación tuvo sobre el CIDE y sobre su personal no pudo ser más beneficioso. La realización de la evaluación externa proporcionó al Centro no solo recursos económicos sino también prestigio y una inmejorable oportunidad para su consolidación como unidad especialista en el ámbito de la evaluación educativa dentro y fuera del Ministerio. La creación, en 1986, del Servicio de Evaluación y el nombramiento de Mariano Álvaro como responsable del mismo supuso el reconocimiento a la labor que este y el resto del personal del CIDE habían realizado en la evaluación externa⁵¹.

⁵¹ Es esta también la opinión de MARIANO ÁLVARO, quien afirma «no podemos dejar de reconocer la clara visión de futuro que tuvieron los directores del CIDE durante los años en que duró la evaluación de la Reforma de enseñanzas medias (1983-1992) y se pusieron las primeras pie-

Ni el director del CIDE ni Mariano Álvaro ponían, entonces, el énfasis en la utilización de la evaluación más allá de la mera «producción de conocimiento». Efectivamente, desde una perspectiva cuantitativista y de evaluación por objetivos, la utilidad de los resultados de la evaluación no se considera un asunto prioritario. Casi diez años después de concluir la evaluación externa, la opinión de Mariano Álvaro a este respecto es bien distinta. Al interrogante que él mismo se plantea —«¿Sirvió la evaluación externa llevada a cabo para mejorar o reconducir, en algún caso, la puesta en marcha de la REM?»—, su respuesta «es que no, de manera clara, por varias razones». Los argumentos que aduce son tan ilustrativos de cuanto hemos venido afirmando que, pese a su extensión, merecen ser aquí expuestos: el primero, «achacable al equipo responsable de esta evaluación externa. Probablemente (...) llegábamos tarde a las necesidades puntuales y concretas (...) de las personas responsables de la implementación de la reforma. Acompasar ambos ritmos hubiera sido fructífero (...) Lo óptimo puede ser enemigo de lo bueno y llegar a tiempo con lo bueno, en ocasiones, puede ser más importante que llegar tarde con lo óptimo» (...) Otra importante razón (...) tuvo que ver con la poca e insuficiente comunicación entre los evaluadores y los responsables de diseñar (la reforma) e, incluso, de ejecutarla. Deberíamos haber trabajado con una mayor interacción entre unos y otros, con el fin de hacer fluir la información, intensamente, desde ambas orillas de un mismo río. Desde la nuestra, no se conocía a fondo la manera de implementar la REM y, desde la suya, no se tenían conocimientos suficientes de los resultados que iban apareciendo con la evaluación externa, lo que les hubiera podido ayudar a modificar el énfasis de sus intervenciones (...)»⁵².

dras de la evaluación del sistema educativo en España» (2010: 203-204). En GARCÍA SANCHEZ (2003; 2004) puede encontrarse información adicional sobre el proceso de institucionalización de la evaluación educativa.

⁵² Véase ÁLVARO PAGE (2010: 219).

CAPÍTULO V

LA EVALUACIÓN DEL PROGRAMA DE INTEGRACIÓN ESCOLAR DE ALUMNOS CON NECESIDADES EDUCATIVAS ESPECIALES

1. EL PROGRAMA DE INTEGRACIÓN ESCOLAR DE ALUMNOS CON NECESIDADES EDUCATIVAS ESPECIALES

1.1. De la segregación a la integración¹

Hasta los años sesenta, era opinión predominante entre psicólogos y pedagogos que las personas aquejadas de alguna minusvalía requerían de un tipo de educación diferenciado dentro del sistema educativo (Echeita 1991). Este enfoque –que algunos han denominado tradicional–, tenía como base los principios de *categorización-segregación-especialización-cuidado*: se consideraba (i) que se podían diferenciar al menos dos tipos o categorías de alumnos, los «normales» y los «deficientes»; (ii) que las particulares condiciones de estos últimos hacían necesario que su educación se llevase a cabo en centros especializados, lo que no era sino una forma de segregación en los centros ordinarios y (iii-iv) que en estos centros especializados, la principal actividad de profesores y especialistas debía ser el cuidado de los alumnos «deficientes» y no su atención educativa.

Los cambios producidos en el ámbito educativo a partir de la década de los setenta tuvieron también su reflejo en el terreno de la educación especial. Precisamente, una de las ideas propugnadas desde los enfoques pedagógicos más innovadores era que la escuela debía responder a las necesidades educativas de todos los individuos, independientemente de cuáles fueran sus condiciones de partida. Los «nuevos enfoques» en educación especial se apoyaban en cuatro ideas fundamentales: *necesidades, integración, currículum y colaboración*². Desde su perspectiva, la función educativa no puede ni debe centrarse en las

¹ Sobre los cambios operados en el ámbito de la educación especial véanse COLL, PALACIOS y MARCHESI (1990) y ECHEITA (1991).

² La descripción de estos nuevos enfoques puede encontrarse en ECHEITA (1991) quien, a su vez, se basa en la obra de AINSCOW y MUNCEY (1989).

discapacidades o características limitadoras del alumno sino en sus *necesidades educativas*: se entiende, en definitiva, que «todos los alumnos son iguales en su diferencia y todos tienen el derecho a ser educados en función de esa diferencia» (Echeita 1991: 72). No se trata de diseñar un sistema educativo específico para estos alumnos sino de *integrarlos* plenamente en los centros ordinarios, mediante la adaptación del *currículum* a sus necesidades educativas, la coordinación de los recursos necesarios y la *colaboración* de distintos especialistas.

La integración se convirtió así en la piedra angular de los nuevos enfoques en educación especial; su finalidad, que todos los niños pudieran desarrollar al máximo sus capacidades potenciales.

1.2. La regulación de la educación especial en España: la asunción del principio de integración

Tanto la Ley General de Educación de 1970 como la normativa más específica sobre educación especial³, habían incorporado ya referencias a la idea de integración. Pero, al margen de ciertas experiencias innovadoras desarrolladas por algunos centros educativos, en la práctica la educación especial seguía funcionando de acuerdo con el enfoque más tradicional.

De hecho, la asunción definitiva del principio de integración por parte de la administración educativa no se produjo hasta marzo de 1985, cuando el Ministerio de Educación y Ciencia aprobó la nueva ordenación de la educación especial⁴. En ella, se entendía que la integración solo es posible si se parte de una previa aceptación de la diversidad: «solamente en una escuela distinta, que huya de la homogeneidad, que se centre en las diferencias, que acepte las diferencias como criterio fundamental de funcionamiento, se podrá hablar de integración» (Marchesi 1986: 6).

El modelo de integración por el que se apostaba en la nueva ordenación es el denominado *funcional*. En él, la escolarización de los alumnos con deficiencias se lleva a cabo en centros ordinarios (siempre que la gravedad de su minusvalía no haga imposible su integración). Los alumnos participan de las actividades de las aulas ordinarias y se integran totalmente en la dinámica del centro. No se trata, pues, de que existan en la escuela aulas especiales para los alumnos integrados (integración *física*) ni de que estos compartan espacios o participen en actividades conjuntas con el resto del alumnado (integración *social*), sino de llegar al estadio más avanzado de la integración⁵.

³ Nos referimos a la Ley 13/1982, de 7 de abril, de Integración Social de los Minusválidos y al Real Decreto 2639/1982, de 15 de octubre, de Ordenación de la Educación Especial.

⁴ Por Real Decreto 334/1985, de 6 de marzo.

⁵ EQUIPO DE EVALUACIÓN DE LA INTEGRACIÓN (1987). Puede encontrarse información más exhaustiva acerca de los tres tipos de integración mencionados en el llamado *Informe Warnock*, que

1.3. El programa experimental de integración

1.3.1. *Los principios y las características del programa*

El Real Decreto de Ordenación de la Educación Especial no hizo sino establecer el marco general en el cual se habría de hacer efectivo el principio de integración. Esto no sucedió hasta la aprobación, en marzo de 1985, de una Orden ministerial⁶ en la que se procedía a la planificación de actuaciones en educación especial y se autorizaba la realización de experiencias en el denominado *territorio MEC*⁷. Álvaro Marchesi, subdirector general de Educación Especial⁸, se hizo cargo del diseño y puesta en marcha del *programa de integración de alumnos con necesidades educativas especiales*. El éxito con el que Marchesi llevó adelante esta tarea, junto con el impulso que otorgó a la evaluación educativa⁹, fueron la base sobre la que se forjó su carrera política en el Ministerio.

El denominado *Programa de Integración Escolar de Alumnos con Deficiencias* era, en realidad, el conjunto de experiencias en integración desarrolladas por aquellos centros educativos que decidieron sumarse a la propuesta del Ministerio de Educación y Ciencia¹⁰. Como otros tantos programas implantados durante la primera legislatura socialista, el de integración tuvo también naturaleza experimental durante sus tres primeros años de existencia. Fue precisamente esta fase experimental la que fue sometida a evaluación.

Al igual que en el resto de los programas experimentales, el Ministerio procedía periódicamente a la convocatoria de ayudas para la realización de proyectos de integración¹¹. Los centros interesados en tomar parte en dicha convocatoria debían presentar un plan educativo en el que dieran cuenta de las actividades previstas, del personal y de los recursos materiales de los que disponían, del estado de las instalaciones y del número y características de los alumnos de integración.

fue presentado al Parlamento británico en 1973 por un Comité de estudio de la educación especial y publicado en 1981.

⁶ Orden de 20 de marzo de 1985, sobre planificación de la educación especial y experimentación de la integración en el curso 1985-1986.

⁷ Véase nota a pie núm. 26 del capítulo III.

⁸ La Subdirección General de Educación Especial estaba adscrita en aquel momento a la Dirección General de Educación Básica (Real Decreto 504/1985, de 8 de abril). Cuando esta Dirección General fue suprimida en 1986 (Real Decreto 2352/1986, de 7 de noviembre), la Subdirección General de Educación Especial pasó a depender de la Dirección General de Renovación Pedagógica.

⁹ Sobre el papel desempeñado por Álvaro Marchesi en el desarrollo e institucionalización de la evaluación educativa véase GARCÍA SÁNCHEZ (2003).

¹⁰ Algunas comunidades autónomas se adelantaron al programa de integración del Ministerio de Educación y Ciencia. En efecto, en Cataluña las experiencias en integración se habían iniciado en 1984. En el País Vasco había dado comienzo en 1982. Sobre los proyectos experimentales de integración en las comunidades autónomas véase ZUFIAURRE (1994).

¹¹ Las órdenes ministeriales de convocatoria pueden consultarse en GARCÍA SÁNCHEZ (2003).

Las direcciones provinciales del Ministerio de Educación y Ciencia fueron las encargadas de la planificación y la implantación de la experiencia en su ámbito territorial. El Ministerio de Educación y Ciencia estableció una serie de directrices a las que aquéllas debían atenerse con el fin de homogeneizar, en lo posible, los diferentes proyectos de integración. Así, debía darse prioridad a la integración de alumnos en edad preescolar y de primer curso de Educación General Básica y debía garantizarse la continuidad en el tiempo de la experiencia y de los centros participantes. Además, tenía que preverse la existencia de, cuando menos, un centro de integración por cada tramo de población de 100.000 a 150.000 habitantes. Las barreras arquitectónicas, en caso de existir, debían ser fácilmente suprimibles.

La administración educativa se comprometía, por su parte, a asumir el coste de las obras que se tuvieran que acometer en los centros, a proporcionar al equipo docente la formación y el asesoramiento que precisaran y, en el caso de centros públicos, a dotarles de profesores de apoyo y a garantizar la estabilidad del profesorado implicado en el centro durante al menos tres años. En contrapartida, las autorizaciones podían ser revocadas en caso de incumplimiento o insuficiente desarrollo de la experiencia por parte de los centros. Al término de los tres años de experimentación, los centros debían pronunciarse sobre su intención de continuar en el programa.

La idea de evaluar el programa, aunque no quedaba explícitamente recogida en la normativa, estuvo siempre presente en los planes de su máximo responsable, Álvaro Marchesi.

1.3.2. *El proceso de implantación*

El programa de integración ocupó un lugar muy destacado en la política del Ministerio en la segunda mitad de los años ochenta y tuvo siempre una gran repercusión en la comunidad educativa. Para las familias y los centros, el programa ofrecía la posibilidad de que la formación de los niños con alguna discapacidad o con problemas de conducta se desarrollara en entornos normalizados. Para las autoridades ministeriales, el programa representaba la oportunidad de reordenar y modernizar de acuerdo con enfoques educativos más innovadores y progresistas un ámbito, como el de la educación especial, que hasta el momento había estado estructurado sobre la base de perspectivas ciertamente conservadoras.

Desde sus inicios, el programa de integración contó con un fuerte respaldo de todos los titulares del Ministerio, especialmente de José María Maravall, quien fue su principal impulsor. El elevado número de centros participantes en la experiencia y la cuantía de las inversiones realizadas fueron dos de las manifestaciones más claras de la importancia que las familias, los centros escolares y el Ministerio de Educación y Ciencia otorgaron al programa.

Como muestra el siguiente cuadro, durante la fase experimental participaron en el programa 443 centros, 2.561 profesores (de aula y de apoyo) y casi 7.500 alumnos.

Cuadro V.1
Evolución del número de centros, profesores y alumnos del programa de integración¹ (cursos 1985-86 a 1989-90)

Curso	Centros			Profesores			Alumnos		
	Púb.	Priv.	Total	Aulas	Apoyo	Total	Púb.	Priv.	Total
1985-86	152	29	181	362	240	602	1.824	348	2.172
1986-87	146	24	170	521	420	941	2.360	404	2.764
1987-88	87	5	92	354	664	1.018	2.352	156	2.508
Total fase exp.	385	58	443	1.237	1.324	2.561	6.536	908	7.444
1988-89	106	10	116	324	813	1.137	3.024	140	3.164
1989-90	158	36	194	504	926	1.530	4.092	472	4.564

¹ Las cifras que figuran en el cuadro, relativas al número de centros participantes, se refieren solo a aquellos a los que se concedía nueva autorización para cada uno de los cursos académicos.

Fuente: CONSEJO ESCOLAR DEL ESTADO (1991: 69).

El coste del programa fue, como queda reflejado en el siguiente cuadro, ciertamente elevado; sólo en la fase experimental la inversión realizada ascendió a 2.429.129.021 millones de pesetas. Sin embargo, las inversiones no siguieron siempre una tendencia al alza. En 1987 y en 1989 se produjeron reducciones significativas con respecto a los ejercicios económicos previos, más acusadas en el primer caso debido, probablemente, a la disminución en el número de centros participantes.

Cuadro V.2
Ejecución presupuestaria del programa de integración entre 1986 y 1990
(en millones de pesetas)

1986	1987	1988	1989	1990	Total
429.639.495 ¹	350.594.102 ²	1.648.895.424	1.606.304.085	1.672.823.091	5.708.256.197

¹ No se dispone de datos relativos a obras y equipamientos.

² No se dispone de datos relativos a obras y equipamientos.

Fuente: Datos proporcionados por la SUBDIRECCIÓN GENERAL DE EDUCACIÓN ESPECIAL Y ATENCIÓN A LA DIVERSIDAD (1997).

Por otra parte, en el transcurso de la fase experimental tuvieron lugar en el contexto y en el seno del programa algunos cambios que incidieron notablemente en su proceso de implantación.

El primero de esos cambios se produjo cuando el Ministerio de Educación y Ciencia decidió conceder el carácter de *centros permanentes de integración* a aquellos que habían iniciado la experiencia en el curso 1985-86 y que manifestaron su voluntad de continuar en el programa¹². Esta autorización parecía confirmar el éxito de la experimentación y la conveniencia de proceder a la generalización y consolidación del programa.

Otra de las decisiones que mayor repercusión tuvo sobre el desarrollo del programa fue la creación y/o regulación de los *Equipos Multiprofesionales* (EE.MM.), de la figura del *profesor de apoyo a la integración*, de los *Equipos de Atención Temprana* (EAT) y del *Centro Nacional de Recursos para la Educación Especial* (CNREE). Todos ellos desempeñaron un importante papel en el proceso de implantación del programa y algunos también durante su evaluación.

Los *Equipos Multiprofesionales* habían nacido en 1982¹³ como unidades dependientes del Instituto Nacional de Educación Especial¹⁴, aunque su funcionamiento no quedó planamente regulado hasta la publicación del Real Decreto de Ordenación de la Educación Especial¹⁵. Los EE.MM. estaban constituidos por psicólogos, pedagogos y trabajadores sociales o expertos en audición o logopedia y centraban su actividad en la evaluación y la atención pedagógica a los alumnos con necesidades educativas especiales, así como en el asesoramiento a los centros educativos. La mayoría de los EE.MM. eran generalistas, aunque algunos se especializaron en el análisis y tratamiento de ciertas minusvalías (los denominados *Equipos Específicos*). Los EE.MM. serían además los encargados de recoger en los centros los datos necesarios para la evaluación del programa de integración.

La figura del *profesor de apoyo a la integración*¹⁶ constituía el nexo de unión entre los EE.MM. y los profesores responsables de las aulas de integración. Los *profesores de apoyo* eran profesionales especializados en educación especial y eran los encargados de facilitar a los alumnos, a sus familias y a los profesores de las aulas de integración el acceso al tipo de asesoramiento que precisasen. Asumían, además, el seguimiento del programa de desarrollo individual de cada niño y la orientación de los padres de los alumnos integrados.

¹² Órdenes de 14 de junio de 1988 y de 23 de junio de 1989. Los centros habían manifestado su deseo de continuar en el programa a través de una encuesta que el Ministerio de Educación y Ciencia realizó a los centros participantes en 1988. Un año después se realizó una encuesta similar a los miembros de los EE.MM. Véase el apartado 2.3 de este mismo capítulo.

¹³ Concretamente, con la Ley de Integración Social de los Minusválidos (Ley 13/1982, de 7 de abril). La composición y funciones de los EE.MM. quedaron reguladas en la Orden de 9 de septiembre de 1982.

¹⁴ Este Instituto se creó por Decreto 1151/1975, de 23 de mayo. Fue suprimido por Real Decreto 504/1985, de 8 de abril.

¹⁵ Real Decreto 334/1985, de 6 de marzo.

¹⁶ Las funciones de los profesores de apoyo a la integración quedan recogidas en el artículo 13 del Real Decreto 334/1985, de 6 de marzo, de Ordenación de la Educación Especial.

Los *Equipos de Atención Temprana*¹⁷, por su parte, surgieron con el objetivo de prestar a los centros y a las familias el apoyo necesario en la detección y la prevención de las discapacidades y las inadaptaciones en los primeros años de vida escolar del niño.

El *Centro Nacional de Recursos para la Educación Especial* se creó en 1986¹⁸ con el objetivo de dar respuesta a las crecientes demandas y necesidades surgidas a raíz de la implantación del programa. Este centro, que dependió inicialmente de la Subdirección General de Educación Especial y, después, de la Dirección General de Renovación Pedagógica, pretendía constituirse en el principal foro de recursos para la investigación, la formación y la información en el ámbito de la educación especial.

Según el testimonio de los profesores entrevistados, las buenas relaciones que el Ministerio y los centros mantuvieron al inicio del programa comenzaron a deteriorarse cuando concluyó la fase experimental. A partir de este momento, la administración comenzó a desatender sus compromisos –la formación del profesorado, la supresión de las barreras arquitectónicas, el reconocimiento de la participación en el programa como mérito docente– o a demorarse en su cumplimiento y decidió, en muchos casos, imponer a los centros su participación. La adscripción obligatoria al programa parecía ser conveniente desde el punto de vista político en una etapa en la que comenzaba a disminuir el número de solicitudes por parte de los centros. Además, en muchas ocasiones, los centros se vieron obligados a admitir a algunos alumnos que presentaban un tipo de discapacidad en el que no estaban especializados o a otros que, en realidad, precisaban asistir a un centro de educación especial.

Todas estas circunstancias hicieron que, a medida que el programa se generalizaba, decrecieran la ilusión y el interés iniciales de profesores y equipos multiprofesionales. Como veremos en las páginas que siguen, la evaluación del programa no hizo sino agravar estos problemas.

2. LA EVALUACIÓN DEL PROGRAMA DE INTEGRACIÓN

2.1. Los comienzos

La evaluación del programa de integración de alumnos con necesidades educativas especiales dio comienzo en septiembre de 1985. El subdirector general de Educación Especial y director del programa, Álvaro Marchesi –un ac-

¹⁷ Estos Equipos se crearon en el Real Decreto 334/1985, de 6 de marzo, de Ordenación de la Educación Especial. Tras la aprobación de la LOGSE, los EE.MM. y los Equipos de Atención Temprana constituyeron, en el territorio MEC, los denominados *Equipos de Orientación Educativa y Psicopedagógica*. Su estructura y funciones quedaron reguladas en la Orden de 9 de diciembre de 1992.

¹⁸ Real Decreto 969/86, de 11 de abril.

tor político-administrativo—, tenía ya en mente la idea de evaluarlo cuando se inició el proceso de experimentación.

Marchesi supo valerse de sus *habilidades* para aprovechar las ventajas que le brindaban su *posición formal* y el *entorno de oportunidad*. Tenía en su haber una sólida formación y una más que contrastada experiencia en el campo de la educación especial y, lo que es aún más importante, la visión política necesaria para llevar adelante el proyecto de evaluación de manera que tanto sus propios *intereses* como los del Ministerio quedasen salvaguardados.

Las decisiones adoptadas por Marchesi encontraron un *entorno* verdaderamente favorable. Para este, como para José María Maravall, la evaluación tenía un doble *interés*. En primer lugar, ambos entendían que la evaluación resultaba imprescindible para disponer, en tiempo real, de una información fiable sobre los problemas surgidos durante la implantación del programa y los posibles efectos no deseados del mismo. En segundo lugar, la evaluación representaba, en especial para Marchesi, una oportunidad única para rentabilizar y legitimar el esfuerzo y el trabajo realizado en la reforma de la educación especial. La legitimación resultaba ser clave para la continuidad del programa en un momento en el que algunos centros y EE.MM. comenzaban a hacer llegar sus críticas sobre la implantación. El nombramiento de Álvaro Marchesi, en 1986, como Director General de Renovación Pedagógica no hizo sino reforzar su posición en el escenario de actores de la evaluación y otorgarle un mayor margen de maniobra.

Además del apoyo que le brindó Maravall, Marchesi precisó de la intervención de otros dos actores político-administrativos: su inmediato superior político, Jaime Naranjo, en aquellos momentos Director General de Educación Básica¹⁹, y el entonces director del CIDE, Julio Carabaña, responsable del centro que iba a asumir la financiación del estudio.

La evaluación del programa de integración fue, por tanto, un proceso formativo e interno²⁰. Álvaro Marchesi pudo haber optado por contratar a evaluadores externos o por recurrir a los servicios del CIDE pero decidió que no fuera así y ello por varias razones. En primer lugar porque, a su juicio, era el *equipo de integración* —esto es, las personas que habían colaborado con él en el diseño y puesta en marcha del programa— el que estaba en mejores condiciones de acometer la evaluación, por su conocimiento y por su implicación en la interven-

¹⁹ Recuérdese que la Subdirección General de Educación Especial dependía en esos momentos de la Dirección General de Educación Básica (Real Decreto 1266/1983, de 27 de abril). Jaime Naranjo González estuvo al frente de esta Dirección General entre marzo de 1984 y noviembre de 1986.

²⁰ Con respecto a la dependencia de los organismos de evaluación de la administración educativa, Marchesi, aun considerando razonables las críticas que ello podía suscitar, afirmaba que lo que nadie podía negar era «que el Ministerio estaba luchando por generar una cultura de la evaluación».

ción. En segundo lugar, porque tanto Marchesi como los integrantes de su equipo²¹ discrepaban «del modelo de enseñanzas medias (...) los problemas de los alumnos solo se entienden en relación... con las condiciones en las que estudian y aprenden y, por tanto, del contexto educativo en el que están, de los centros, del propio sistema, mientras que la de enseñanzas medias fue una evaluación muy clásica, de resultados de los alumnos y nosotros vimos que había otro enfoque, más cualitativo en gran medida, más de intervención-acción». Así, al menos en teoría, eran partidarios de enfoques evaluativos más pluralistas, alejados del cuantitativismo que venía caracterizando los trabajos de evaluación del CIDE. Por último, porque la legitimación del programa podía verse comprometida si un equipo externo se hacía con el control de la evaluación. Fueran cuales fueran los motivos, lo cierto es que el hecho de que la evaluación fuera interna simplificó las relaciones entre actores y constituyó así un factor decisivo para la viabilidad del proceso²².

2.2. El diseño de la evaluación

2.2.1. *La definición de los objetivos*

La definición de los objetivos de la evaluación, como el resto de las tareas de diseño técnico y planificación, corrieron a cargo de un grupo de trabajo seleccionado y dirigido por Álvaro Marchesi, el denominado *equipo de diseño de la evaluación*. José María Maravall no intervino de manera directa en la formulación de los objetivos, a pesar de lo cual sus ideas quedaron perfectamente reflejadas en los mismos. La participación de ambos actores en esta decisión les situó en una *posición central* en el proceso evaluativo.

El equipo de diseño de la evaluación estaba formado, en su mayor parte, por expertos procedentes del mundo de la psicología²³ que habían colaborado anteriormente con Marchesi en la puesta en marcha del proceso experimental. El consenso entre sus miembros fue la nota dominante: no en vano, todos ellos tenían *ideas* muy similares acerca del programa y de los principios sobre los que debía descansar la evaluación. Como recuerda Rosario Martínez Arias, la proximidad de perspectivas en el seno del equipo evaluador se quebraba tan solo cuanto se debatían cuestiones metodológicas de carácter menor. Precisamente, la coincidencia de *visiones, valores e intereses* entre sus miembros y su capacidad «real» de actuar como una unidad frente al resto son los elementos

²¹ En ese sentido se pronuncian también Rosario Martínez Arias y Ángel Rivière.

²² En ello coinciden tanto Marchesi y Martínez Arias como los dos miembros de los EE.MM. entrevistados.

²³ Estos expertos fueron Amalio Blanco Abarca, César Coll Salvador, Alfredo Fierro Bardají, Rosario Martínez Arias (quien, como se recordará, tomó también parte en la evaluación de la reforma experimental de enseñanzas medias), Jesús Palacios González, Ángel Rivière Gómez (Director del CIDE entre 1986 y 1989) y Alberto Rosa Rivero.

que nos permiten considerar a este grupo como un actor colectivo compuesto y no meramente nominal.

La concepción que el equipo de diseño tenía del programa condicionó la formulación de los objetivos de la evaluación. Quienes lo integraban entendían que el programa se caracterizaba, básicamente, por (i) su naturaleza abierta y flexible: el Ministerio establecía unos mínimos o condiciones generales a partir de los cuales los centros participantes ponían en marcha su propia experiencia de integración; así pues, en la práctica, había tantos programas de integración como centros; (ii) la diversidad de formas en que se concretaban los apoyos proporcionados por la administración educativa y (iii) la inexistencia de directrices claras sobre cómo había de adecuarse el *currículum* de la educación ordinaria a las necesidades de los alumnos integrados (Equipo de Evaluación de la Integración 1987).

En función de estas premisas, se entendió que la evaluación debía perseguir dos grandes objetivos: (i) emitir un juicio valorativo general sobre el programa de integración y (ii) determinar en qué medida dicho programa cumplía sus dos propósitos, promover el desarrollo de los alumnos integrados e introducir mejoras en el sistema educativo.

De cara a su operacionalización, los objetivos generales fueron, a su vez, desagregados en los siguientes objetivos específicos²⁴:

- (i) Analizar las diferentes aplicaciones en las que se materializaba el programa, así como el modo en que cada una de ellas había evolucionado.
- (ii) Proponer las estrategias de intervención necesarias para la implantación de las modalidades de integración más positivas para todos los implicados.
- (iii) Valorar el impacto del programa sobre los niños integrados, los profesores de apoyo y de aula y los procesos de enseñanza-aprendizaje.

2.2.2. *Un enfoque metodológico supuestamente iluminativo*

La determinación del enfoque metodológico de la evaluación, al igual que la formulación de los objetivos, correspondió única y exclusivamente al equipo de diseño de la evaluación. Ciertamente es que en dicho equipo tomaron parte algunos investigadores que no tenían vinculación oficial con el programa, pero sus posturas acerca de la integración eran bastante similares a las mantenidas por el Ministerio. El equipo de diseño se aseguró, así, una *posición central* en el proceso decisorio.

La determinación del enfoque metodológico por parte del equipo de diseño estuvo marcada no solo por los objetivos de la investigación y la naturaleza del programa sino también por los *intereses* de sus miembros.

²⁴ Véanse AGUILERA *et al.* (1990) y EQUIPO DE EVALUACIÓN DE LA INTEGRACIÓN (1987).

En los primeros momentos, el equipo de diseño no tenía una idea clara acerca de cuál debía ser el enfoque de la evaluación. La única cuestión decidida desde el principio era que la evaluación debía tener carácter formativo, esto es, debía desarrollarse de forma paralela al programa y servir, de este modo, para su revisión y perfeccionamiento. Ello parecía lógico no solo desde el punto de vista estrictamente metodológico sino también desde la óptica de las necesidades estratégicas de los actores implicados en el diseño. Una evaluación de este tipo permitiría ir introduciendo en el programa los ajustes necesarios a medida que se advertían las dificultades y los problemas.

La decisión sobre si el enfoque de evaluación debía ser cuantitativo o cualitativo no resultó fácil. El equipo de diseño advirtió que la complejidad del programa desaconsejaba la utilización de un diseño de corte experimental, en el que resulta imprescindible poder distinguir las variables dependientes de las independientes. Finalmente, decidió apostar por un enfoque mixto e iluminativo, en definitiva, en palabras de Marchesi, por un enfoque más cualitativo y menos «psicologicista». Sin embargo, lo cierto es que, en realidad, la del programa de integración distó de ser un ejemplo de evaluación iluminativa.

En primer lugar porque, en la práctica, se acudió en ella a un modelo de evaluación por objetivos –y es bien sabido que los principios en los que este y el enfoque iluminativo descansan son, en gran medida, contrapuestos²⁵.

En segundo lugar, porque quienes debieran haber sido sus auténticos protagonistas –las familias, los alumnos, los profesores de aula y de apoyo y los centros– quedaron relegados a una posición pasiva que les permitió participar pero, exclusivamente, en calidad de informantes. Y es que la demostración de las bondades del programa a través de su evaluación pasaba por la exclusión del proceso de todo aquel que tuviera o pudiera tener una actitud negativa o crítica hacia el programa. Los miembros del equipo de diseño conocían la opinión que en los centros y en los EE.MM. se tenía sobre la integración a través de las distintas reuniones informativas que habían mantenido con ellos. La valoración de estos actores, aunque favorable a los principios esenciales del programa, no era especialmente positiva en lo referente a su implantación.

²⁵ Como se recordará, el modelo de evaluación por objetivos, desarrollado por RALPH TYLER (1950), es un enfoque típicamente acrítico, en el que se entiende que entre el programa y los resultados producidos en la población objeto existe una relación de causa-efecto unidireccional y en el cual se presupone la bondad de los objetivos tal y como son formulados por los responsables del programa. La importancia de los objetivos es tal que su cumplimiento se considera el principal criterio de valor de la evaluación. Por el contrario, el enfoque iluminativo desarrollado por PARLETT y HAMILTON (1977), considera que la implantación de un programa educativo desencadena una serie de procesos que repercuten sobre él y pueden llegar a modificarlo sustancialmente. Se considera, pues, que entre el programa y el contexto en el que este se aplica existe una relación bidireccional. Desde este enfoque se apuesta por la utilización de técnicas cualitativas y por la implicación activa de los distintos *stakeholders* y de sus diferentes perspectivas en el proceso de evaluación.

El equipo de diseño no ofreció a los centros la posibilidad de participar en la elección del enfoque metodológico y, según afirman todos los profesores de aula entrevistados, tampoco fueron receptivos a sus sugerencias y críticas. Para la mayoría de los docentes su exclusión, y la de los centros, del proceso decisorio fue uno de los aspectos más negativos de la evaluación.

Es cierto, sin embargo, que los centros tampoco presionaron al equipo de diseño para poder intervenir en la evaluación. En primer lugar, porque su *posición formal* no se lo permitía y, en segundo lugar, porque ello les podía llevar a un enfrentamiento con el Ministerio que perjudicara claramente sus *intereses*. La integración representaba la oportunidad de contar con el apoyo ministerial a la hora de hacer frente a las necesidades de los alumnos discapacitados que, en muchos casos, habían acudido al centro incluso antes de que el programa diese comienzo. La evaluación, aunque podía proporcionarles datos de enorme interés, no dejaba de ser una cuestión secundaria que, en la mayoría de las ocasiones, solo servía para incrementar la carga de trabajo de los profesores y para alterar el ritmo de las clases.

Los miembros de los EE.MM. con los que tuvimos ocasión de intercambiar impresiones se muestran bastante menos críticos en este sentido. Creemos que ello está relacionado con dos cuestiones. En primer lugar, los EE.MM. tenían aún mayor *interés* que los centros en la continuidad del programa: no en vano la consolidación de su *posición* dependía de la reforma de la educación especial. En segundo lugar, su participación como *actores activos*, aunque *periféricos*, les otorgó desde el principio un mayor protagonismo en el escenario de la evaluación.

2.2.2.1. *El modelo teórico*

El modelo teórico de la evaluación partía de la premisa de que los cambios que el programa pretendía inducir –promover el desarrollo de los alumnos integrados e introducir innovaciones y mejoras en el sistema educativo– eran el resultado de la interacción de cinco sistemas²⁶: el centro escolar, el «ambiente inmediato de aprendizaje»²⁷, el profesor, el alumno y el sistema familiar. Cada uno de dichos sistemas quedaba, a su vez, caracterizado en función de una serie de variables que se consideraba podían influir, *a priori*, en la implantación y en los resultados del programa. Las variables empleadas aparecen recogidas en el cuadro V.3.

²⁶ Desde nuestro punto de vista, ni el «profesor» ni el «alumno» pueden considerarse como tales. Según la teoría general de sistemas, un sistema es un conjunto de partes o elementos interrelacionados que tiene, a su vez, relación con el entorno. La aplicación de esta teoría al ámbito de la política fue realizada, como es bien sabido, por EASTON (1969/2006).

²⁷ PARLETT y HAMILTON (1977) definen este concepto como el «entorno socio-psicológico y material donde los estudiantes y los docentes trabajan juntos»; citado en EQUIPO DE EVALUACIÓN DE LA INTEGRACIÓN (1987: 14).

Cuadro V.3
Variables analizadas en la evaluación (por sistema)

Centro	<p>Características en los aspectos rural-urbano, público-privado y número de unidades</p> <p>Proyecto educativo</p> <p>Estructura organizativa y funcionamiento</p> <p>Experiencias previas en educación especial</p> <p>Grado de acuerdo previo para participar en el programa</p> <p>Actitudes y expectativas ante el programa</p> <p>Utilización de los recursos disponibles</p>
Ambiente inmediato de aprendizaje	<p>Características físico-ambientales del aula</p> <p>Ratio profesor-alumnos y número de alumnos integrados</p> <p>Tipos, naturaleza y organización de las actividades de aprendizaje</p> <p>Programación de objetivos, contenidos y actividades</p> <p>Materiales didácticos utilizados</p> <p>Seguimiento y evaluación de los alumnos</p> <p>Organización concreta del trabajo con niños integrados: modalidad de integración</p> <p>Utilización efectiva de los apoyos proporcionados por el programa</p>
Profesor	<p>Formación profesional previa y permanente</p> <p>Experiencia docente y con alumnos con necesidades especiales</p> <p>Experiencia previa en integración</p> <p>Actitudes y expectativas ante el programa y los apoyos que proporciona</p> <p>Estilo de enseñanza</p>
Alumnos	<p>Integrados</p> <p>Tipo, naturaleza y grado de déficit</p> <p>Experiencias educativas y tratamientos personalizados previos a la participación en el programa</p> <p>Nivel intelectual</p> <p>Nivel de lenguaje</p> <p>Nivel de desarrollo de la identidad personal</p> <p>Naturaleza de las relaciones con los compañeros y tipo de interacciones que establece con ellos</p> <p>Nivel de integración en el grupo y grado de aceptación por los compañeros</p> <p>Tipo de relación con el profesor</p> <p>Nivel de aprendizaje en las diferentes áreas del currículum</p> <p>Pares¹</p> <p>Nivel de desarrollo de la identidad personal</p> <p>Tipo de interacciones que establece con los compañeros</p> <p>Nivel de aprendizaje en las diferentes áreas del currículum</p>
Familia	<p>Edad y nivel de estudios</p> <p>Actitudes y expectativas ante el programa</p>

¹ Los alumnos pares eran aquellos con los cuales se efectuaban las comparaciones.

Fuente: Elaboración propia a partir de AGUILERA *et al.* (1990: 22) y EQUIPO DE EVALUACIÓN DE LA INTEGRACIÓN (1987: 32).

La complejidad y la diversidad de los procesos de integración hicieron que el equipo de diseño decidiera estructurar el proceso de evaluación en tres fases; no eran etapas diferenciadas en el tiempo sino distintas estrategias de investigación, en las que se recurría a diferentes técnicas de obtención y análisis de datos pero cuyos resultados habían de interpretarse de manera conjunta. En definitiva, se apostó por una estrategia de validación convergente²⁸, con el objeto de superar las limitaciones de cada uno de los métodos y de obtener una visión global y panorámica del programa. De la coordinación técnica de las tres fases se hizo cargo un grupo de expertos, el llamado *equipo de coordinación*²⁹ que, junto con el de diseño, conformaba el *equipo de evaluación de la integración*.

2.2.2.2. *Las fases de la evaluación y las técnicas de recogida y análisis de datos*

- (i) La denominada fase cuantitativa se desarrolló en el transcurso de la fase experimental, esto es, durante los tres primeros años de implantación del programa. En ella se procedió a la medición periódica de las variables definitorias del programa con objeto de obtener una descripción exhaustiva del mismo. El diseño tuvo, pues, carácter longitudinal. A partir de esta información se pretendía identificar las condiciones que garantizaban la eficacia del programa, definir las modalidades o tipos posibles de integración, analizar los procesos de cambio generados por la implantación y estudiar las relaciones entre los diferentes sistemas.

Las aplicaciones fueron realizadas en 60 de los centros participantes en la experiencia: en ellos cursaban estudios 379 alumnos discapacitados. Los centros se seleccionaron mediante muestreo aleatorio sistemático a partir del total de centros que habían manifestado su interés por tomar parte en la evaluación³⁰. Sólo 52 de los 60 continuaron hasta 1989, fecha en que concluyó el estudio.

²⁸ La estrategia de validación convergente, también conocida como «triangulación», tal y como fue definida por CAMPBELL (1956), hace referencia al empleo de distintas técnicas de recogida de datos para medir un mismo concepto. DENZIN (1978) amplió el significado de este último término para significar el uso no solo de diferentes técnicas de recogida de datos sino también de distintas teorías, distintos investigadores, distintas metodologías (o bien la combinación de todos estos) con el fin de evitar o paliar las diferencias que presenta cada uno. La obra de BERG (1998) ofrece una interesante descripción de esta estrategia de investigación.

²⁹ Este equipo estaba integrado por Elena Martín (quien posteriormente se haría cargo de la dirección del gabinete de Álvaro Marchesi cuando este fue nombrado Secretario de Estado de Educación, en junio de 1992), Katia Álvarez, Mercedes Babío, Gerardo Echeita y Marisa Galán. En algunos momentos participaron también María Jesús Aguilera y Mercedes Rey.

³⁰ A pesar de ello, tres de los cinco profesores de aula entrevistados consideran bastante poco probable que esto fuera cierto. En su opinión, el equipo de diseño solo seleccionó a los centros más motivados e innovadores, aunque no fueran los de mejores resultados. Sostienen, incluso, que en cada centro solo se contactaba con aquellos padres que estaban más implicados y que mostraban una actitud más abierta y menos crítica con el programa.

Los siguientes cuadros ofrecen una visión detallada acerca del carácter de los centros y de los alumnos participantes en la fase cuantitativa:

Cuadro V.4
**Características de la muestra seleccionada
 para la fase cuantitativa (centros)**

Tipo de centros	Número de centros
Públicos	48
Privados	12
Rurales	13
Urbanos	47
Nueva creación	6
Ya existentes	54
Experiencia previa en integración	28
Sin experiencia previa en integración	32
0-12 unidades	20
13-23 unidades	25
Más de 24 unidades	15

Fuente: EQUIPO DE EVALUACIÓN DE LA INTEGRACIÓN (1987: 41).

Cuadro V.5
**Características de la muestra seleccionada
 para la fase cuantitativa (alumnos)**

Deficiencia mental	226
Inadaptación social	12
Alteraciones de la conducta	16
Deficiencia auditiva	28
Deficiencia motórica	56
Deficiencia visual	12
Otros	25
Total	375¹

¹ El número de alumnos de la muestra, tal y como se recoge en el informe de evaluación, fue de 379.

Fuente: EQUIPO DE EVALUACIÓN DE LA INTEGRACIÓN (1987: 41).

Como muestra el siguiente cuadro, para la recogida de datos se emplearon diferentes tipos de tests y cuestionarios:

Cuadro V.6
**Cuadro-resumen de los instrumentos de recogida de datos empleados
 en la fase cuantitativa**

Instrumento	Información que proporcionaba
Hoja descriptiva de las características físico-ambientales del centro	Instalaciones y características de las aulas de integración (ubicación del centro, número de m ² , existencia de barreras arquitectónicas, ratio alumnos/profesor, número de alumnos integrados por aula)
Cuestionario para el análisis del funcionamiento del centro educativo (Q.U.A.F.E.)	Información general sobre los puntos fuertes y débiles del funcionamiento del centro. Pretendía convertirse en una herramienta para la autoevaluación formativa del centro
Informe cualitativo sobre el centro	
Hoja de seguimiento del alumno	Evolución del alumno integrado en las áreas más importantes de la actividad escolar a lo largo del curso académico
Escala de observación de la interacción social	Tipo de conducta que establece el alumno integrado con su entorno (pasiva, solitaria o relacional)
Escala de actitudes de los padres	Expectativas de éxito o fracaso de la experiencia y sobre la incidencia del programa en el alumno. Actitudes generales ante la integración
Escala de actitudes de los profesores ante la integración	Actitudes de los tutores de aulas de integración ante el programa
Escala de inteligencia de Weschler	Aspectos cuantitativos y cualitativos de la inteligencia
Test de matrices progresivas de Raven	Aptitud general para la deducción de relaciones
Test de vocabulario de Peabody	Recepción por el alumno de un vocabulario de palabras simples pronunciadas por el aplicador
Pruebas pedagógicas de Terrassa	Aptitud matemática (cálculo, lógica, geometría) y lingüística (lenguaje oral y lenguaje escrito) y habilidades perceptivo-motrices

Fuente: Elaboración propia a partir de EQUIPO DE EVALUACIÓN DE LA INTEGRACIÓN (1987: 34-39) y AGUILERA *et al.* (1990: 21).

El elevado número de pruebas aplicadas durante la fase cuantitativa y la diversidad de niveles de medida de las variables hicieron aconsejable acudir a técnicas de análisis de datos de muy distinta naturaleza. De la selección de dichas técnicas se ocupó el equipo de diseño de la evaluación³¹.

³¹ Entre las empleadas destacaron los análisis descriptivos de las variables, las técnicas de reducción de la dimensionalidad, el análisis de regresión múltiple (para el estudio del rendimiento académico de los alumnos integrados y pares), los análisis de varianza uni y bifactoriales (para

- (ii) La segunda fase se orientó al análisis de las diferencias entre las distintas modalidades de aplicación del programa y se llevó a cabo a finales del curso 1986-87. Para ello, se efectuaron mediciones de las principales variables utilizadas en la fase anterior pero sobre un número más reducido de centros.
- (iii) La denominada fase cualitativa³² se desarrolló a lo largo del curso 1987-88, último de la fase experimental, y en ella se acudió a dos estrategias de investigación: un *análisis en profundidad de un reducido número de centros*, siguiendo la metodología propia del estudio de caso y un *proceso de intervención*, cuyo objetivo era determinar si los cambios inducidos por dicha intervención se ajustaban a lo previsto por los evaluadores en el modelo teórico. Dos integrantes del *equipo de coordinación* eran los encargados del estudio de caso; para el proceso de intervención, se les unía un tercer especialista: estos tres asesores conformaban el denominado *equipo de intervención*. En ambos casos, la metodología fue cualitativa.

Para el *estudio en profundidad* se procedió a la selección de tres casos a partir de los tres tipos resultantes del análisis de clusters efectuado en la fase cuantitativa: aquellos centros cuyos valores en las variables del sistema «centro» se consideraban más favorables para la ejecución del programa, aquellos cuyos valores eran menos favorables y aquellos que presentaban valores intermedios. De cada uno de los tres tipos se eligió un centro que pudiera ser considerado paradigmático. Puesto que mediante el análisis de clusters se había determinado que cada centro era representativo de su tipo, resultaba posible generalizar los resultados obtenidos a los centros de características semejantes. Posteriormente se incluyó un nuevo centro del tipo 1 (valores más favorables), esta vez de carácter privado, con el fin de conocer la influencia que ello podía tener en la organización y funcionamiento del centro. Para la recogida de datos se realizaron entrevistas semiestructuradas a los padres, los profesores y los alumnos, en las que se obtuvo información sobre todas las variables y dimensiones de la investigación. A través del análisis de categorías se extrajeron y agruparon los datos más relevantes.

Asimismo, se desarrolló un proceso de observación participante en distintas situaciones de la vida del centro, especialmente en el aula, en el recreo y en las reuniones del profesorado.

el estudio de las relaciones entre los distintos sistemas) y el análisis de varianza no paramétrico de Friedman, el análisis multivariante o el análisis de tablas de contingencia (según la naturaleza de las variables) para el análisis longitudinal. Una breve descripción de estas técnicas puede encontrarse en GARCÍA SÁNCHEZ (2003).

³² Del diseño de esta fase se hizo cargo César Coll.

Se analizó también la documentación de los centros (proyectos educativos, planes de estudio, memorias, proyectos de innovación, programaciones, etc.), se recabó información específica acerca de los alumnos (trabajos en clase, calificaciones, etc.) y se realizó un reportaje fotográfico a fin de cotejar los datos existentes sobre los espacios y materiales del centro.

Todo aquello que se estimaba de interés quedó reflejado en un diario. La triangulación fue una constante a lo largo de toda esta fase. Con ello se pretendía contrastar las interpretaciones de los investigadores con las de los profesores o, incluso, con las de personas ajenas a la evaluación o profundizar alguna cuestión concreta a través de diferentes estrategias de investigación.

La información obtenida en esta primera parte de la fase cualitativa resultó fundamental para poder llevar a cabo el proceso de intervención; de hecho, los centros seleccionados para este proceso fueron los mismos que se habían elegido para el análisis cualitativo.

El *proceso de intervención* tenía como finalidad introducir en el centro escolar aquellos cambios que el profesorado o los evaluadores consideraban necesarios. Se pretendía que la intervención tuviera unos objetivos de cambio realistas, que tuviera carácter preventivo y que se desarrollase de manera adaptativa y flexible, por lo que se recurrió a la metodología de la *investigación-acción*³³. La pretensión de los evaluadores era que estos *asesores de intervención*, más que adoptar el papel de expertos externos en el diagnóstico y en la búsqueda de soluciones, se implicaran en la intervención de manera activa, motivando al profesor para que reflexionase sobre su práctica educativa y proporcionándole los conocimientos necesarios para promover el cambio.

³³ La *investigación-acción* constituye un método de intervención de tipo antropológico, de gran tradición en el campo de la pedagogía, y cuyo objetivo es la introducción de mejoras en un determinado contexto o situación a partir de la información que se obtiene de su estudio. Véase, en especial, ELLIOTT (1994).

Cuadro V.7
Relación de variables más relevantes consideradas en las fases cualitativa y de intervención (por sistema)

Centro	Experiencias previas en integración Razones de la adscripción al programa Grado y tipo de acuerdo para la adscripción Naturaleza, variedad y utilización de los recursos del centro Proyecto educativo
Ambiente inmediato de aprendizaje	Procesos instruccionales Adaptaciones curriculares
Profesor	Formación Experiencia Actitudes y expectativas con respecto al programa
Alumnos integrados	Tipo y grado de déficit
Modalidad de integración	Lugar, tiempo y criterios de la modalidad
Familia	Expectativas y actitudes de los padres ante la integración y el programa

Fuente: Elaboración propia a partir de AGUILERA *et al.* (1990: 112 y ss.).

* * *

El diseño de la evaluación tuvo, en líneas generales, bastante buena acogida por parte de los distintos actores.

La opinión más favorable, como es lógico, fue la de los miembros del equipo de evaluación. Álvaro Marchesi, Rosario Martínez Arias y Ángel Rivière destacaron como rasgos más definitorios y positivos del proceso la combinación de técnicas cuantitativas y cualitativas de recogida y análisis de datos y el hecho de que en el mismo participaran prácticamente todos los actores que tenían interés en el programa³⁴.

En la misma línea se sitúan los integrantes de los EE.MM. En su opinión, el empleo de una metodología ecléctica permitió dibujar una imagen muy completa y panorámica del programa.

La visión del profesorado fue, en cambio, bastante más negativa. Aunque les pareció acertado que el enfoque metodológico fuese mixto, fueron muy críticos con los instrumentos de recogida de datos empleados en la fase cuantitativa. Tres de los cinco profesores entrevistados señalaron que a través de las pruebas aplicadas era muy difícil, si no imposible, que se llegara a conocer

³⁴ Realmente, como podremos comprobar en este capítulo, la mayoría de los interesados intervinieron, únicamente, en calidad de informantes.

cómo funcionaba realmente el programa y que en ellas no se profundizaba ni se descendía a la realidad de los centros. Por el contrario, el trabajo realizado por los EE.MM. y por el equipo de coordinación durante la fase cualitativa fue considerado, por todos los profesores entrevistados, como el aspecto más destacable del proceso evaluativo. Todos ellos insisten en que su asesoramiento les proporcionó una información de gran interés para reflexionar sobre su práctica docente e importantes claves para su mejora.

2.2.3. *Los criterios de valor*

El informe de evaluación publicado por el CIDE no hacía mención alguna de los criterios que habían sido empleados para enjuiciar el programa. La única referencia explícita a esta cuestión se encuentra en un documento publicado por el Equipo de Evaluación de la Integración en 1987 y al que hemos aludido ya en reiteradas ocasiones.

Según este documento, los criterios de valor se encontraban recogidos en los propios objetivos de la evaluación, lo que equivale a decir que de su formulación se habían encargado quienes habían definido dichos objetivos. De este modo, el equipo de diseño e, indirectamente, José María Maravall vieron reforzada, aún más si cabe, la *centralidad de su posición* en el proceso decisorio.

La finalidad del estudio, como se recordará, era «emitir un juicio valorativo sobre el P.I. y determinar en qué medida logra su doble objetivo, tanto promover el desarrollo de los niños integrados como introducir mejoras en el sistema educativo» (Aguilera *et al.* 1990: 14). El equipo de diseño entendía, aunque no lo hiciera explícito, que cuanto mayor fuera el progreso de los niños integrados y cuanto mayores fueran los avances producidos en el sistema educativo mayor sería el éxito del programa; no se planteó, en ningún momento, si los objetivos resultaban pertinentes o si habían sido correctamente formulados: su cumplimiento se equiparaba y utilizaba como criterio del éxito del programa.

2.3. **Del diseño al análisis de datos**

Durante la fase cuantitativa, la recogida de datos correspondió a aquellos EE.MM. y Servicios de Orientación Escolar y Vocacional (SOEV)³⁵ de la zona en la que se ubicaba el centro que decidieron tomar parte en la evaluación y cuyos miembros recibieron formación específica para el manejo, aplicación y corrección de las pruebas. La coordinación técnica del trabajo de campo recayó en el equipo de coordinación.

³⁵ Los SOEV se crearon en 1977 (Orden Ministerial de 30 de abril) y se ocupaban fundamentalmente de tres tareas: la orientación personal y escolar de los alumnos, el asesoramiento a padres y profesores de EGB y la colaboración en el diagnóstico de los alumnos de integración.

Como se recordará, los sujetos de los que se recabó información durante la fase cuantitativa fueron aquellos que se vincularon al programa en su primer año de implantación, esto es, en el curso académico 1985-86. Las mediciones se realizaron durante tres cursos consecutivos, entre septiembre de 1985 y junio de 1988. El cuadro VI.8 recoge el momento de aplicación de cada una de las pruebas.

Cuadro V.8
Momento de aplicación de las pruebas de la fase cuantitativa

Hoja de las características físico-ambientales del centro	Inicio del primer curso
Cuestionario sobre el funcionamiento del centro (Q.U.A.F.E.)	Línea base Inicio del primer curso Final de cada curso
Informe cualitativo sobre el centro	Una aplicación por curso
Cuestionario de aspectos curriculares e instruccionales	Una aplicación por curso
Cuestionario de opiniones pedagógicas de los profesores	Una aplicación por curso
Hoja de la modalidad de la integración	Inicio y final de cada curso
Escala de Desarrollo de la Identidad Personal (EDIP)	Línea base Final de cada curso
Hoja de seguimiento del alumno	Inicio y final de cada curso
Escala de observación de la interacción social	Inicio y final de cada curso
Escala de actitudes de los padres	Línea base Final de cada curso
Escala de actitudes de los profesores ante la integración	Línea base Final de cada curso
Escala de inteligencia de Weschler	Línea base Final del último curso
Test de matrices progresivas de Raven	Línea base Final de cada curso
Test de vocabulario de Peabody	Línea base Final del último curso
Pruebas pedagógicas de Terrassa	Final del penúltimo y último curso

Fuente: AGUILERA *et al.* (1990: 21).

Los resultados obtenidos durante la fase cuantitativa sirvieron de base para la elaboración de sendas encuestas a los profesores y los miembros de EE.MM. que se habían sumado a la experiencia en su primer año de implantación. El objetivo era conocer su opinión con respecto a su continui-

dad en el programa y comprobar hasta qué punto las valoraciones que hacían sobre el programa se confirmaban en el resto de centros y EE.MM.³⁶.

En la primera de estas encuestas tomaron parte 2.193 profesores de integración, lo que representaba cerca de un 61 por ciento del total. De los 730 que pertenecían a centros participantes en la evaluación, un 92,6 por ciento manifestó su deseo de continuar en el programa³⁷. El porcentaje, en el caso de los profesores de centros no participantes en la evaluación, descendió hasta el 87,3.

Las opiniones del profesorado de los centros sometidos a evaluación eran más favorables, en todos los ítems, que las de quienes no tomaban parte en la evaluación. Así, por ejemplo, el 88 por ciento de los docentes que pertenecían a centros evaluados hacía una valoración positiva del programa frente al 80,5 por ciento de los profesores de centros no sometidos a evaluación. La razón que explica este hecho no es otra que la especial motivación de los primeros.

El número de integrantes de EE.MM. que respondieron a la encuesta fue considerablemente menor. De las 952 encuestas repartidas solo se recibieron 207 (un 21,7 por ciento). El 96,5 por ciento manifestó su interés por seguir en el programa, a pesar de lo cual el porcentaje de quienes lo valoraron positivamente fue solo del 73,1.

En la fase cualitativa, dos integrantes del denominado equipo de coordinación asumieron la realización de las entrevistas y las tareas de observación y llevaron a cabo el análisis documental. Antes de que comenzase la recogida de datos, que se centraba en algunas de las variables de cada sistema, el equipo de coordinación se ponía en contacto con la dirección del centro seleccionado para comunicarle cuáles eran los objetivos del estudio y, más específicamente, los correspondientes al análisis cualitativo. La dirección del centro debía comprometerse a informar de todo ello al claustro y al consejo escolar.

En un segundo momento, los aplicadores contactaban con el inspector o inspectora de la zona con objeto de darle a conocer la finalidad de la evaluación y de recabar datos sobre el centro.

La reunión con el equipo multiprofesional de la zona constituía el siguiente paso. En ella se recogía también información sobre el funcionamiento del centro y el proceso de implantación del programa.

³⁶ Los resultados de las encuestas fueron publicados por la Dirección General de Renovación Pedagógica en 1988 y 1989, respectivamente. En las encuestas se pedía a estos profesionales su opinión sobre las siguientes cuestiones: el cambio inducido por el programa, los recursos administrativos, la suficiencia de los recursos materiales, el trabajo del profesorado de apoyo, la formación recibida, las potencialidades de la integración (mejora de la dinámica del aula, contribución al perfeccionamiento del profesorado, estilo de enseñanza, innovaciones pedagógicas), la estabilización de plantillas, la influencia del programa sobre las actitudes en el centro, la modificación de objetivos del currículum y la valoración de las actividades de sensibilización del Ministerio de Educación y Ciencia.

³⁷ De los 1.463 sujetos restantes, 1.445 pertenecían a centros no vinculados a la evaluación. Los otros 18 profesores se encuadraban en la categoría «no sabe/no contesta». Para mayor detalle véase DIRECCIÓN GENERAL DE RENOVACIÓN PEDAGÓGICA (1988).

Posteriormente, se celebraba un encuentro con el claustro en el que se negociaban tanto las condiciones bajo las cuales se iba a llevar a cabo el estudio (fundamentalmente en lo relativo a fechas y horario de visitas) como los compromisos que el centro y el equipo de evaluación iban a asumir.

Una vez concluía el trabajo de campo, se elaboraba un informe que era sometido a la consideración del claustro. En ocasiones, el equipo de coordinación incorporaba al informe final las objeciones realizadas por aquel.

En la *fase de intervención* se procedía de manera similar. En este caso, era el *equipo de intervención* el que asumía las tareas de recogida de datos.

El primer contacto con la dirección del centro tenía un propósito eminentemente informativo. Durante el mismo se explicaba cuáles eran los objetivos de la intervención, los papeles que corresponderían al equipo de intervención y al centro, la duración y el tipo de actividades que se iban a desarrollar y las posibles ayudas que iba a proporcionar la administración. Si el centro aceptaba tomar parte en el estudio, se convocaba una reunión con el equipo multiprofesional y con el orientador pedagógico para planificar el trabajo y evitar, así, posibles interferencias entre el trabajo de unos y otros.

En las siguientes sesiones de trabajo con el equipo directivo, el claustro, los profesores y los coordinadores de ciclo se recogían las principales demandas de todos estos colectivos. El equipo de intervención se encargaba después de contrastar estas demandas con las necesidades que había detectado durante la fase cualitativa.

* * *

Todos los entrevistados coinciden en que el trabajo de campo ocasionaba frecuentes interrupciones en la marcha de clases y restaba tiempo al que debía ser principal cometido de los EE.MM.: el diagnóstico de las necesidades de los alumnos de integración y el asesoramiento a los profesionales del centro. No obstante, como ya hemos señalado, incluso los profesores más críticos destacaron la labor de quienes se ocuparon de la recogida de datos como el elemento más positivo de la evaluación y el que mayor repercusión tuvo sobre su trabajo y sobre la dinámica del centro, algo que también ha sido señalado por los miembros de los EE.MM. De hecho, la relación que mantuvieron con los EE.MM. y con el equipo de coordinación fue mucho más estrecha y cordial que la que mantuvieron con el equipo de diseño. Con este, el trato se limitó a las sesiones formativas e informativas que se celebraban periódicamente en el marco del programa y a las visitas que algunos de los miembros del equipo –en ocasiones, el propio Marchesi– realizaron a los centros.

Tanto los centros escolares como los EE.MM. colaboraron en todo momento con el equipo evaluador, a pesar de sus críticas al proceso de implantación del programa y a la forma en que se había concebido y se estaba desarrollando la evaluación. Entendían que esta les podía servir, siquiera mínimamente, para

conocer en qué medida estaban contribuyendo a que se alcanzaran los resultados previstos.

Pero, con el paso del tiempo, las diferencias entre los centros y la administración se fueron acentuando. En ello influyeron la forma en que se había planteado la evaluación pero también, y muy especialmente, los problemas surgidos durante el proceso de generalización del programa a los que ya nos hemos referido. La evaluación no fue la causa del distanciamiento entre el Ministerio y los centros pero, sin duda, contribuyó a que este fuese aún mayor.

2.4. La etapa final de la evaluación

2.4.1. *Un documento de legitimación del programa: el informe final de evaluación*³⁸

El informe final de la evaluación del programa de integración se orientó, más que a dar cuenta de cuál era el grado de consecución de los objetivos del programa, a la verificación de las hipótesis de las que se partía en la investigación. Una vez más, fue el equipo de evaluación el que se situó en una posición central al determinar los contenidos, el estilo y el grado de difusión que se daría al informe.

En la estructura del informe final se pueden distinguir dos partes claramente diferenciadas. La primera de ellas se ocupa de los principales resultados y conclusiones de las fases 1ª y 3ª de la evaluación. En la segunda se recogen las conclusiones generales de la evaluación así como las sugerencias y recomendaciones que los equipos de diseño y coordinación dirigieron a la administración educativa. Esta última constituye, junto con los capítulos dedicados al análisis cualitativo, la única cuyo estilo resulta asequible para quien no esté familiarizado con cuestiones de carácter estadístico.

Las conclusiones³⁹ se organizaron a partir de los objetivos de la investigación:

Objetivo I. *Estudiar cómo se concreta en la realidad el programa de integración*

- Los factores que mayor incidencia tenían en la implantación del programa son los que se habían considerado en el diseño técnico. La evaluación permitió constatar que el éxito del programa dependía de la combinación y el ajuste entre variables, aun cuando parecía que eran las relativas al sistema «centro» las de mayor relevancia.

³⁸ Nos referiremos aquí, exclusivamente, al informe final de la evaluación ya que en este se recogen los informes parciales relativos a los estudios de caso realizados en los centros escolares.

³⁹ Aunque la exposición de las conclusiones generales de la evaluación resulte algo extensa, entendemos que ello es necesario para dar una idea más aproximada del estilo y el «tono» del informe.

- La existencia de un proyecto educativo de centro que considere las necesidades de todos los alumnos constituía una variable esencial para la implantación del programa pero no era determinante de su éxito.
- El tipo y grado de acuerdo en la vinculación al programa resultaban claves para el posterior proceso de toma de decisiones. Cuanto mayores eran el debate y el consenso mayores eran las oportunidades de éxito del programa.
- Si bien era importante la dotación de recursos materiales del centro al inicio del programa, aún lo era más el aprovechamiento que se hiciera de esos recursos. Se constataba que, cuanto más sólido era el proyecto educativo y mayor el grado de coordinación, mayor era también el grado de aprovechamiento de los recursos.
- La experiencia previa de los profesores como equipo docente favorecía su coordinación y la calidad del proyecto educativo. Esta variable parecía ser aún más importante que la experiencia en el trabajo con alumnos con necesidades educativas especiales.
- Las actitudes de los profesores tutores y de los profesores de apoyo revestían una singular importancia para el programa. Estas actitudes eran más positivas en aquellos centros que contaban con un mejor proyecto educativo o en aquellos en los que este había mejorado sustancialmente con el tiempo.
- La experiencia previa en integración tanto de los tutores como de los profesores de apoyo constituía uno de los principales factores de éxito del programa.
- El nivel de formación del profesorado contribuía de manera significativa al correcto desempeño de sus funciones.
- Se constató la relevancia de la variable «estilo de enseñanza» aunque no pudo afirmarse la existencia de un estilo «mejor». El más apropiado sería aquel que mejor se adecuara a las circunstancias del alumno, del aula y del centro.
- Las adaptaciones curriculares se consideraban un factor de gran relevancia para la integración. Dichas adaptaciones eran más frecuentes en aquellos centros en los que existía un proyecto educativo más innovador y una práctica educativa más coordinada.
- El nivel intelectual del alumno condicionaba en mayor medida sus resultados que el tipo de déficit que presentara.
- Ninguna variable era, en sí misma, decisiva, a pesar de lo cual las vinculadas al sistema «centro» eran las más relevantes para predecir el éxito del programa.
- El análisis cualitativo ponía de manifiesto que las divergencias entre las distintas categorías o tipos de aplicación del programa eran aún mayores de lo que inicialmente apuntaba el análisis cuantitativo.

- Ninguno de los centros con valores más altos descendió de categoría mientras que, de los que tenían peores resultados, algunos pasaron a la categoría superior.
- Las diferentes modalidades de integración no parecían tener correspondencia con las grandes formas de aplicación del programa. En líneas generales, en los centros que obtenían peores resultados, los alumnos integrados contaban con mayores apoyos por parte de profesionales distintos del tutor. En los centros mejores, la modalidad de integración estaba más adaptada, desde un comienzo, al cociente intelectual del alumno.

Objetivo II. *Analizar el impacto de la aplicación del programa sobre el medio de aprendizaje en el que se lleva a cabo*

- La participación en el programa no afectaba negativamente ni al funcionamiento ni a la organización de los centros. La integración constituía un elemento de dinamización en los centros con mejores niveles organizativos y con un alto grado de consenso en torno al proyecto educativo. Tenía, no obstante, repercusiones negativas en los centros peor organizados o en los que existía desacuerdo inicial acerca del proyecto.
- Las demandas que los centros con mejores resultados dirigían a la administración parecían estar mejor adaptadas a sus necesidades que las de otros centros, que son «menos pertinentes (...) no responden a un análisis previo de las necesidades (...) son demandas “insaciables”»⁴⁰.
- Las actitudes del profesorado con respecto a la integración mejoraban en los primeros momentos de implantación del programa para estabilizarse después. El cambio positivo era más evidente en los centros que presentaban peores condiciones de partida. Las actitudes de los padres no variaban sustancialmente con el programa.
- La motivación del profesorado a la hora de adaptar su estilo instruccional a las necesidades de los alumnos integrados era muy elevada. El estilo «liberal» parecía ser más habitual en los centros «mejores». En este tipo de centros los alumnos integrados permanecían más tiempo en las aulas ordinarias.

Objetivo III. *Analizar el impacto de la aplicación del programa sobre los niños integrados*

- Los alumnos de integración, si bien continuaban su proceso de desarrollo cognitivo y social y podían adquirir conocimientos propios del nivel que cursaban, seguían estando por debajo de sus compañeros pares en cuanto a la consecución de objetivos. El programa tenía una repercusión muy

⁴⁰ Así se afirma textualmente en el informe de evaluación. Véase AGUILERA *et al.* (1990: 309).

positiva en la adquisición, por parte de los alumnos integrados, de habilidades sociales y en el desarrollo de su autonomía.

- Las variables de mayor incidencia sobre los resultados de los alumnos eran el estilo de enseñanza, las actitudes del profesor y la modalidad de integración. Asimismo, la planificación exhaustiva, por parte del profesorado, de las actividades se reveló como una de las dimensiones más importantes del estilo de enseñanza.
- Los centros optaban por la modalidad de integración más adecuada a sus características y a las necesidades de los alumnos integrados. La modalidad de integración no pareció afectar a los resultados obtenidos por estos alumnos.
- Cuando el déficit que presentaba el alumno era de carácter específico, la modalidad de integración por la que se optaba era de especial trascendencia para que el niño alcanzara los objetivos del currículum ordinario.
- En todos los centros, el cociente intelectual se relacionaba inversamente con el nivel de apoyo que recibe el alumno.
- Tanto el desarrollo social como la adquisición de hábitos sociales y de autonomía personal de los alumnos integrados evolucionaron positivamente tras su participación en el programa de integración.

Objetivo IV. *Proponer estrategias de intervención adecuadas para crear las condiciones que hagan posible, en función de las características de cada centro escolar, la modalidad o modalidades de integración más positivas en cuanto a las repercusiones sobre todos los elementos implicados*

- La respuesta de los centros ante el proceso de intervención difería de unos casos a otros, de tal manera que la intervención tomaba rumbos también distintos.
- En los centros con mejores resultados la intervención sobre el proyecto educativo tenía efectos positivos sobre todos los sistemas. En cambio, en los otros, era necesario actuar previamente con objeto de que los profesores se comprometieran en torno a una tarea común.
- La intervención sobre las dimensiones del sistema profesor influía positivamente en los resultados tanto de los alumnos integrados como de los alumnos pares.
- La mejora de las actitudes del profesorado dependía, en gran medida, de que los recursos y la formación que se les proporcionaban respondieran a sus necesidades reales.
- La labor de apoyo de los equipos psicopedagógicos a los centros tenía una incidencia muy positiva sobre todos los sistemas.

Por lo que se refiere a las recomendaciones y sugerencias –un capítulo básico si se quiere que la evaluación resulte útil–, se afirmaba la conveniencia de que la administración siguiera proporcionando a los centros los recursos necesarios y de que recabara de estos una información «más personalizada y cualitativa», de forma que fuera más fácil responder adecuadamente a sus requerimientos.

Como puede apreciarse, el informe final presenta, en líneas generales, una imagen bastante favorable de la integración. A lo largo del mismo aparecen constantes referencias a la bondad del programa, a la actitud prudente de la administración en su puesta en marcha, a la importancia de la estrategia seguida por el Ministerio y, lo que es aún más sorprendente, a los errores en los que incurren los centros y a los que se atribuyen algunos de los fracasos. Del tono legitimatorio –que no puede sino sorprender en un informe de evaluación– son buen reflejo estas afirmaciones, un claro alegato a favor del programa y una crítica explícita a sus detractores: «El Programa de Integración, tal y como fue diseñado y aplicado por el MEC, durante estos cuatro años, *no ha sido un fracaso educativo*, como algunos esperaban»⁴¹.

Pero el equipo de evaluación no se limitó únicamente a ensalzar las virtudes del programa sino también las del propio modelo de evaluación al que había recurrido. En el informe de evaluación se subrayaba la gran capacidad predictiva del modelo y su utilidad para reflejar la realidad sistémica de los centros. Se insistía también en su valor orientativo para otros servicios educativos y se recomendaba a la administración que fomentara el debate y la difusión, «lo más ampliamente posible»⁴², de los resultados de la evaluación.

La valoración que los profesores a los que tuvimos la oportunidad de entrevistar hacen de los informes de su centro⁴³ no es precisamente positiva. A juicio de tres de ellos las conclusiones «no se deducían de los datos», reflejaban tan solo «la cara más amable del programa» y «minusvaloraban, cuando no obviaban, las dificultades y los problemas» que los profesores advertían en su trabajo diario.

2.4.2. *Los usos del proceso de evaluación*

La percepción de los distintos participantes acerca de la utilidad de la evaluación es bien distinta. Como es lógico, la opinión más favorable es la de Álvaro Marchesi, para quien «la propia evaluación dinamizó el proceso de reflexión de todos los que trabajábamos (...) y permitió constatar dos cosas. La primera, la importancia de que la integración fuera algo asumido por el centro

⁴¹ AGUILERA *et al.* (1990: 323).

⁴² AGUILERA *et al.* (1990: 337).

⁴³ Ninguno de ellos conocía el informe final completo.

y de que había que estimular, más que obligar, a los centros a elaborar sus propios proyectos de integración; la segunda, la dificultad de implantar el programa en estructuras rígidas, por lo que era preciso dar más autonomía y más capacidad de decisión a los centros».

Marchesi reconoce que la evaluación tuvo un *uso conceptual*, que fue solo una fuente de información más entre otras muchas y que la continuidad del programa y el interés del Ministerio no dependieron nunca ni del desarrollo de la evaluación ni de cuáles fueran sus resultados. La integración «era una cuestión de política educativa y de apuesta por la igualdad de oportunidades (...) la evaluación no era para decir “paramos la integración”, era para ver en qué condiciones la integración era posible».

Por su parte, Rosario Martínez Arias señala que la prueba de que la evaluación no era un mero instrumento para legitimar el programa sino que pretendía saber realmente qué estaba sucediendo era que «en ningún momento se quiso dar la visión de que todo funcionaba perfectamente e igual en todas partes»⁴⁴. Sobre esta misma cuestión, Ángel Rivière afirmaba que «el que la evaluación sea una instancia legitimadora de una política no la devalúa; lo que es más discutible es cómo se utiliza esa información. Si usted informa a los periódicos sobre lo que dice tal evaluación, usted está haciendo un uso político de la evaluación pero es que, si no informa a los periódicos, quiere decir que usted está manejando la evaluación como un instrumento secreto, que no es la vocación que debe tener una evaluación en una sociedad democrática».

Los miembros de los EE.MM. se muestran bastante críticos en relación con la utilidad de la evaluación, «que no sacaba a la luz si el programa funcionaba y en qué casos funcionaba». Entre los profesores entrevistados no existe unanimidad sobre el grado de utilidad de la evaluación. Quienes subrayan con mayor énfasis el *uso persuasivo* (legitimador) de la evaluación son los profesores de aquellos centros en los que solo se realizó el estudio cuantitativo:

«La evaluación no servía para nada, no tenía fundamento pedagógico ni sociológico, nadie se acercó a ver nuestro colegio, cómo teníamos estructuradas las clases, qué metodología utilizábamos, todo eso que podían haber preguntado y que nos hubiese servido, no lo hicieron. La evaluación sobraba porque no servía para mejorar el programa, que es para lo que debe servir una evaluación. La evaluación era un instrumento de autojustificación para el Ministerio, que había hecho un programa en el que había puesto mucho empeño y tenía que hacer algún trabajo sobre los resultados».

⁴⁴ Efectivamente, la evaluación mostraba también los problemas surgidos en algunos colegios. Sin embargo, como hemos visto en el apartado anterior, estos problemas fueron atribuidos a la situación de partida de los centros o a su deficiente capacidad de gestión pero, en ningún caso, a defectos en el diseño del programa por parte de la administración educativa.

«La evaluación se quiso hacer porque se querían presentar unos buenos resultados en un momento en que ya había muchas críticas a la integración y para justificar una considerable inversión económica. (...) Marchesi tenía que demostrar que lo que estaba haciendo tenía valor, para todo, para que hubiera más medios para la integración y para que él subiera más arriba. La evaluación no sirvió para mejorar el programa, ni para ayudar al profesorado a conocer cómo lo estaba haciendo».

«Siempre se percibía que hacían caso a las cosas positivas y obviaban un poco las negativas. Siempre querían sacar adelante la bondad de la integración»⁴⁵.

Por el contrario, aquellos otros profesores que pertenecían a centros en los que sí se llevó a cabo el estudio cualitativo (y que tuvieron, por tanto, una relación más próxima a los EE.MM., los SOEV y a los equipos de coordinación e intervención) consideran que la experiencia fue «una herramienta de gran ayuda para advertir las principales deficiencias y puntos débiles de nuestro trabajo con los alumnos de integración»⁴⁶.

Ahora bien, con respecto al valor que los entrevistados atribuían al trabajo de estos equipos es conveniente hacer dos puntualizaciones. En primer lugar, resulta prácticamente imposible discriminar qué aspectos del trabajo de los equipos encargados del trabajo de campo formaban parte de su cometido como integrantes del equipo de evaluación y cuáles se encuadraban en su labor habitual de asesoramiento; en otras palabras, determinar si la mejora de la práctica profesional de los docentes es atribuible o no a la evaluación del programa. En segundo lugar dicha mejora, de haberse producido, hubiera repercutido sobre la dinámica del centro y apenas si hubiese tenido alguna incidencia sobre los aspectos comunes y sustantivos del programa.

* * *

La evaluación del programa de integración nunca pretendió poner en tela de juicio ni el programa, tal y como estaba concebido, ni la integración misma. Ciertamente, existía un interés real por conocer cómo estaba funcionando la integración en la práctica, como también resulta evidente que el proceso de evaluación propició la reflexión y el debate tanto entre sus responsables como entre los centros docentes. La evaluación de la integración constituye, desde nuestra perspectiva, un magnífico ejemplo de lo que Argyris denomina «rutinas defensivas de la organización»⁴⁷: el equipo de evaluación no tuvo intención de

⁴⁵ Aunque este centro fue objeto de un análisis cualitativo y el profesor al que entrevistamos destacó la importante labor desarrollada por los aplicadores se mostró bastante crítico con el proceso de evaluación.

⁴⁶ El entrevistado que hizo esta afirmación reconocía que el centro tuvo «una relación muy estrecha y positiva con los equipos (...) y con el Ministerio. (...) Estábamos muy mimados por Marchesi».

⁴⁷ Este autor (1994) describe las «rutinas defensivas» como aquellas políticas, prácticas y actuaciones que hacen que los seres humanos eviten sentirse amenazados, turbados o desconcertados y, al tiempo, analizar la naturaleza y las causas de esa amenaza, turbación o desconcierto.

cuestionar, en ningún momento, ni el diseño del programa ni los principios en los que este descansaba.

Tres son las razones que nos llevan a hacer estas afirmaciones:

- En primer lugar, la opinión que sobre la finalidad legitimadora de la evaluación tienen algunos de los miembros de los EE.MM. y de los profesores de aula entrevistados.
- En segundo lugar, el hecho de que aunque se afirmara partir de una perspectiva iluminativa se apostara, en la práctica, por un enfoque de evaluación por objetivos, enfoque este, como es bien sabido, típicamente acrítico, en el que el logro de los objetivos del programa es empleado como criterio de valor de la evaluación. Coincidimos con Ewald (1986) cuando afirma que no poner en tela de juicio los objetivos de un programa no tiene solo consecuencias técnicas: al no hacerlo se dejan de cuestionar también los valores y principios en los que dicho programa se sustenta.
- Por último, la imagen tan positiva del programa e, incluso, del propio modelo evaluativo, que presenta el informe de evaluación y el que en este se atribuyan algunos de los problemas a una deficiente actuación por parte de los centros.

El equipo evaluador dirigió todos sus esfuerzos a hacer de la evaluación una herramienta para la mejora del programa (esto es, a darle un *uso instrumental*) pero también para su legitimación. Uno de los usos que se pretendió dar a la evaluación fue, desde un principio, el *persuasivo* y *simbólico* y ello condicionó la adopción de todas y cada una de las decisiones configuradoras.

CAPÍTULO VI

LA EVALUACIÓN DEL PROYECTO *ATENEA* PARA LA INTRODUCCIÓN DE LOS ORDENADORES EN LAS AULAS

1. EL PROYECTO *ATENEA* PARA LA INTRODUCCIÓN DE LOS ORDENADORES EN LAS AULAS

1.1. Los antecedentes y la puesta en marcha del proyecto

Del conjunto de programas experimentales de innovación educativa desarrollados en España en la década de los ochenta destacaron, desde su creación, los relacionados con la introducción de las nuevas tecnologías de la información y la comunicación¹. La creciente importancia social y económica que la informática iba adquiriendo hizo que, desde diversos sectores de la administración central² y desde los propios centros escolares, se comenzasen a impulsar distintos estudios y proyectos para la integración de los ordenadores en la educación.

Más concretamente, fue un grupo de trabajo de la Comisión Ministerial de Informática (que había tenido oportunidad de conocer las experiencias desarrolladas en otros países), el que sugirió a la Secretaría General Técnica del Ministerio de Educación y Ciencia que elaborase una propuesta para la incorporación de los microordenadores a la enseñanza. Los contenidos de la misma quedaron recogidos en un documento titulado *Proyecto ATENEA. Una propuesta para la integración racional de las nuevas tecnologías de la información en la enseñanza básica y media*, que fue aprobado por el Ministerio en marzo de 1985³.

¹ Véase el informe elaborado por el PNTIC sobre el proyecto *Atenea* que sirvió de base a la evaluación realizada por la OCDE y que está incluido en el correspondiente documento de este organismo (1991).

² Nos estamos refiriendo al Consejo Superior de Informática, la Comisión Ministerial de Energía del Ministerio de Industria y al Centro de Proceso de Datos del Ministerio de Educación y Ciencia.

³ En opinión de Julio Carabaña, uno de los factores que más contribuyeron a la puesta en marcha del Proyecto *Atenea* fue la presión ejercida por las grandes multinacionales del sector informático.

Este primer proyecto perseguía básicamente dos grandes objetivos: fomentar la utilización de los ordenadores como instrumento al servicio del aprendizaje de los alumnos y difundir la utilidad de la informática y sus aplicaciones entre los miembros de la comunidad escolar. Al no contar la propuesta con el VºBº del Ministerio de Economía y Hacienda, el Ministerio de Educación y Ciencia, a través de su Secretaría General Técnica –que ocupaba en aquellos momentos Joaquín Arango–, decidió impulsar un proyecto algo menos ambicioso en cuanto a su alcance pero idéntico en sus objetivos al diseñado inicialmente. Así nació, en abril de 1985 y vinculado ya al gabinete de la recién creada Secretaría General de Educación, el proyecto *Atenea*⁴.

1.2. Objetivos y ámbito de aplicación

Los principales impulsores del proyecto, el secretario general técnico del Ministerio, Joaquín Arango, y su directora, Elena Veiguela, eran conscientes de que la introducción de las nuevas tecnologías carecía de sentido si no se convertía en un aliciente para la renovación de las prácticas educativas en los centros escolares. Precisamente por ello, la integración de la informática debía hacerse sobre la base de planteamientos educativos innovadores que se orientaran a potenciar un tipo de aprendizaje más activo y autónomo. El objetivo no era, pues, que el ordenador se convirtiera en un fin en sí mismo, ni siquiera en una herramienta de apoyo para el aprendizaje de determinadas materias. La plena inclusión de la informática en el *currículum*, apenas desarrollada por los países de nuestro entorno, se convirtió en la principal finalidad del proyecto.

Como la mayoría de programas y proyectos de aquel período, el proyecto *Atenea* tuvo carácter experimental⁵. El nuevo secretario general de educación, Alfredo Pérez Rubalcaba y Elena Veiguela estimaron que, dado el carácter novedoso que aún tenía la informática, la experimentación constituía el mejor modo de conocer la «bondad» y el grado de aceptación del programa antes de su generalización y que, además, la experiencia debía ser sometida a evaluación⁶.

⁴ Orden ministerial de 19 de abril de 1985. Para entonces, Cataluña y País Vasco contaban ya con una cierta experiencia en el ámbito de las nuevas tecnologías. En la primera, los programas experimentales para la introducción de la informática en la educación habían dado comienzo en 1982, principalmente en centros de FP y bachillerato. En el curso 1985-86 se aprobó en el País Vasco un plan quinquenal para la introducción de la informática en EGB y en enseñanzas medias, aunque existían ya algunos proyectos. Remitimos, de nuevo en este punto, a la obra de ZUFIAURRE (1994).

⁵ Durante los cinco primeros años. Al término de la fase experimental dio comienzo la «fase de extensión» que supuso, entre otras cosas, la ampliación del número de centros adscritos al proyecto, la creación de los coordinadores provinciales del proyecto y la modificación de la figura del coordinador de equipo pedagógico.

⁶ De hecho, solo en el transcurso de la fase experimental se pusieron en marcha dos procesos de evaluación. Sobre las evaluaciones realizadas *cfr. infra*, apartado 1.4.

La participación de los centros del *territorio MEC*⁷ en el proyecto era voluntaria, con la única limitación de que impartiesen enseñanzas medias o ciclo superior de EGB. El Ministerio de Educación y Ciencia convocaba periódicamente una serie de ayudas⁸ para formación y adquisición de equipos informáticos y software, a la que concurrían los centros interesados en tomar parte en la experiencia⁹. Estos centros, previa autorización del claustro y del consejo escolar, debían presentar ante el Ministerio una propuesta de aplicación de la informática a la educación. En la selección de los colegios se daba prioridad a aquellos que habían desarrollado con anterioridad proyectos de innovación educativa. Los CEP a los que estaban adscritos los centros recibían el mismo equipamiento que estos.

Los cuadros adjuntos muestran el número total de centros que tomaron parte en el proyecto a lo largo de la fase experimental. En el curso 1985-86, primero de implantación del proyecto, participaron en el mismo 127 centros públicos de las once Comunidades Autónomas que en aquel momento constituían el denominado *territorio MEC*, además de Ceuta y Melilla. Durante la fase experimental el número de centros participantes se multiplicó por 7. Sin embargo, en ese mismo período y, a pesar de este notable incremento, la media de ordenadores por centro se redujo, aunque de forma poco significativa.

Cuadro VI.1
Evolución del número de centros participantes y de equipos informáticos en el proyecto *Atenea* (1985-1990)

	1985	1986	1987	1988	1989	1990	Total
Nº de ordenadores	286	1.785	1.050	1.356	1.436	2.100	8.013
Nº de centros	127	401	563	671	698	1.004	1.004

Fuente: OCDE (1991: 84).

El coste del proyecto *Atenea* fue, durante la fase experimental, muy elevado (y muy superior al que tuvo el proyecto *Mercurio*), lo cual se explica no solo por el número de centros participantes sino por el elevado coste de los equipos informáticos. Entre 1985 y 1990, el coste se incrementó en algo más de un 1.450 por ciento. Esta tendencia al alza (que, solo en los dos primeros años de vida de proyecto, fue de casi 720 millones de pesetas) se mantuvo durante toda

⁷ Véase nota a pie, núm. 26 del capítulo III.

⁸ Las órdenes de convocatoria pueden consultarse en GARCÍA SÁNCHEZ (2003).

⁹ La selección de las empresas que suministraban los equipos informáticos se realizaba, evidentemente, por concurso público. De la convocatoria y la resolución del concurso se ocupaba la Junta de Construcciones, Instalaciones y Equipamiento Escolar del Ministerio de Educación y Ciencia, que estaba vinculada a la Dirección General de Programación e Inversiones de la Subsecretaría del Ministerio.

la fase experimental, a excepción del ejercicio 1987, en el que el gasto se redujo en algo más de 22 millones de pesetas debido, probablemente, a la necesidad de inversión en el proyecto *Mercurio*.

Cuadro VI.2
Coste total del Programa de Nuevas Tecnologías. Proyectos *Atenea* y *Mercurio*, 1985-1990 (en millones de pesetas)

	1985	1986	1987	1988	1989	1990	Total
<i>Atenea</i>	129,950	848,124	825,334	1.301,224	1.507,381	2.026,800	6.663,308
<i>Mercurio</i>	–	–	134,678	355,167	831,377	946,869	2.102,596
Total	129,950	848,124	960,012	1.656,391	2.338,758	2.973,669	8.765,904

Fuente: OCDE (1991: 83).

Como se puede advertir en el siguiente cuadro, las partidas presupuestarias que mayor peso tuvieron en el coste del proyecto *Atenea* fueron, evidentemente, las relativas a adquisición de equipos informáticos y a personal. Todas las partidas se incrementaron si exceptuamos la primera, que sufrió un descenso considerable entre 1986 y 1989, y la relativa a desarrollo de software, que se mantuvo constante a lo largo de toda la fase experimental.

Cuadro VI.3
Costes del proyecto *Atenea*, 1985-1990 (en millones de pesetas)

	1985	1986	1987	1988	1989	1990	Total
Costes de equipos	111,550	645,600	500,000	460,500	387,800	579,000	2.684,450
Costes de mantenimiento	–	8,924	37,487	57,200	66,313	115,000	284,915
Costes de personal (incluyendo formación)	8,400	143,600	208,681	450,391	472,734	622,000	1.905,806
Costes de formación (materiales, viajes)	10,000	50,000	44,975	87,679	114,261	159,000	465,915
Software, materiales, concursos	–	–	–	119,358	331,307	363,000	813,665
Desarrollo de software	–	–	30,000	30,000	30,000	30,000	120,000
Publicación de materiales didácticos	–	–	4,200	23,596	25,416	43,100	96,312
Costes de funcionamiento de centros	–	–	–	72,500	79,550	115,700	292,245
Total	129,950	848,124	825,334	1.301,224	1.507,381	2.026,800	6.663,308

Fuente: OCDE (1991: 83).

1.3. La formación del profesorado

La formación del personal fue, sin lugar a dudas, la cuestión a la que desde el proyecto se prestó una mayor atención: no en vano resultaba imprescindible para la implantación del proyecto dado el carácter novedoso que, en aquellos momentos, tenía la informática en España. Se organizó un sistema de formación «en cascada». Los monitores fueron uno de los actores más importantes del proyecto *Atenea* y los primeros en recibir formación en nuevas tecnologías. Estos profesores de EGB o enseñanzas medias con dedicación a tiempo completo a las actividades formativas se encargaban de formar a los coordinadores de los equipos pedagógicos, esto es, a los profesores que se responsabilizaban de la puesta en marcha y desarrollo del proyecto en los centros escolares; los coordinadores, a su vez, se ocupaban de formar a los profesores pertenecientes a dichos equipos.

Con el propósito de que la formación específica en nuevas tecnologías se integrara plenamente en el proceso de formación global del profesorado, la figura del monitor quedó adscrita a los CEP, aunque mantuvo su vinculación orgánica, funcional y salarial a los Servicios Centrales del Programa. A pesar de ello, los monitores constituyeron un actor diferenciado desde el momento en que adoptaron una posición propia frente a la evaluación, si bien sus ideas e intereses eran casi siempre coincidentes con los de los Servicios Centrales del Programa. Los monitores acudían periódicamente a cada uno de los centros escolares asignados a su CEP con objeto de prestar a los coordinadores el asesoramiento necesario y de verificar la marcha del proyecto. La doble dependencia de los monitores hizo que estos recibieran en muchas ocasiones directrices y órdenes contradictorias, algo que fue tremendamente disfuncional para la marcha del proyecto.

La formación de los monitores fue encomendada por el Ministerio a un equipo de profesores vinculados al ICE de la Universidad de Murcia, cuyo director sería después responsable de la evaluación¹⁰. Para la directora del proyecto, la formación no debía quedar reducida a una mera adquisición de habilidades en el manejo de los ordenadores, sino que tenía que servir para propiciar en el docente la reflexión y la actualización didáctica. El monitor no debía ser únicamente un formador de docentes sino, ante todo, un agente de innovación en los centros. Por ello, las materias impartidas en los cursos formativos se re-

¹⁰ El convenio que sirvió de marco a la colaboración entre el Ministerio de Educación y Ciencia y la Universidad de Murcia se firmó el 15 de marzo de 1988, aunque los cursos formativos habían dado ya comienzo en el curso académico 1985-86. Los monitores que superaban el curso, que tenía 520 horas de duración, obtenían el título de «Formador de Formadores para la Integración de las Nuevas Tecnologías de la Información y la Comunicación en las Enseñanzas Primaria y Secundaria». Aunque no hemos podido tener acceso al texto del convenio, hemos podido saber que tanto el diseño del curso como la selección del profesorado que podía acceder a él correspondían al Ministerio de Educación y Ciencia a través de la dirección del proyecto y al ICE de la Universidad de Murcia. La expedición del título era responsabilidad de esta última.

lacionaban tanto con cuestiones técnicas como didácticas. Pero los intentos de los Servicios Centrales para que los monitores asumieran este doble papel –de agentes de innovación y de formadores en medios informáticos– resultaron infructuosos. Tan solo un número muy reducido de monitores veía en el proyecto una oportunidad para iniciar una experiencia de innovación educativa¹¹.

Los cursos formativos no tuvieron muy buena aceptación entre los monitores. Para estos y para algunos miembros de los servicios centrales¹², hacían un excesivo énfasis en las cuestiones puramente técnicas, en la enseñanza de software y del funcionamiento de los equipos informáticos, mientras que los aspectos pedagógicos quedaban relegados a un segundo plano. Los propios datos de la evaluación corroboran estas afirmaciones. El siguiente cuadro recoge la opinión de los monitores sobre la formación recibida.

Cuadro VI.4
Valoración por parte de los monitores de la formación recibida¹

	1+2	3	4+5
Cursos de informática educativa	13,5	20,4	65,9
Innovación y formación de profesores	40,8	37,5	20,4
Técnicas de análisis (cuestionarios, entrevistas, observación)	49,9	37,5	12,4

¹ La pregunta se formuló del siguiente modo: «Valora a continuación cada uno de los aspectos siguientes relativos a la formación, según su contribución a tu capacitación como *monitor*, en una escala de menor (1) a mayor (5)». En el cuadro aparecen, pues, agrupados los porcentajes de respuesta de las dos categorías de menor valoración y las de las dos de mayor valoración.

Fuente: ESCUDERO MUÑOZ (1989: 27).

Para casi el 66 por ciento de los monitores, la formación recibida en «informática educativa» merecía una alta o muy alta valoración. En cambio, el porcentaje descendía al 20,4 y al 12,4, respectivamente, en lo relativo a «innovación y formación» y a «técnicas de análisis». De hecho, entre un 40 y un 50 por ciento de los entrevistados valoró con las puntuaciones más bajas la formación recibida en estas materias.

Sin embargo, en opinión de Luis del Blanco, la formación de los monitores era, tanto por sus planteamientos como por el material que se proporcionaba a los asistentes, «impecable y actualizada». La valoración que los monitores hicieron estuvo muy influida, a su juicio, por el hecho de que un sector, poco numeroso pero muy activo, de los monitores estaba más interesado en aprender informática que en conocer sus aplicaciones pedagógicas.

¹¹ Esta es, al menos, la opinión que mantienen Luis del Blanco, consejero técnico del PNTIC en temas de formación, y algunos de los monitores entrevistados.

¹² Nos referimos a Carmen Candiotti, Carlos San José y Luis del Blanco.

Sea como fuere, lo cierto es que el descontento de los monitores con los cursos de formación fue la causa de su desencuentro con el equipo de la Universidad de Murcia. Las diferencias entre ambos actores se harían aún más acusadas, como tendremos ocasión de comprobar en las páginas que siguen, con ocasión de la puesta en marcha del proceso de evaluación.

1.4. La creación del Programa de Nuevas Tecnologías de la Información y la Comunicación. El proyecto *Atenea* en el marco del PNTIC

En el momento de su creación, el proyecto *Atenea* no contaba con una estructura organizativa ni con recursos financieros propios, sino que dependía de los presupuestos de las Direcciones Generales de Educación General Básica y de Enseñanzas Medias¹³. Por esta razón, se constituyó una *Comisión de Seguimiento y Coordinación* de la que formaban parte los responsables de las diferentes unidades orgánicas del Ministerio de Educación y Ciencia que habían estado implicadas en la puesta en marcha del proyecto: la Secretaría General de Educación, las Subdirecciones Generales de Educación General Básica, Bachillerato, Perfeccionamiento del Profesorado, Organización y Automación y el CIDE¹⁴.

Sin embargo, el hecho de que la coordinación del proyecto y su financiación dependiesen de distintas unidades resultaba escasamente operativo para la consecución de los objetivos planteados. Con el fin de resolver los problemas que ello generaba, a mediados de 1986 y a iniciativa de Joaquín Arango, se creó el Programa de Nuevas Tecnologías de la Información y la Comunicación (PNTIC), que unificaba los recursos humanos y financieros hasta entonces dispersos y en el cual se encuadraron los Proyectos *Atenea* y *Mercurio*¹⁵. El

¹³ Concretamente, de los presupuestos destinados a material didáctico y a personal.

¹⁴ Recordemos que según el Real Decreto 504/1985, de 8 de abril, la Secretaría General de Educación estaba integrada por las Direcciones Generales de Promoción Educativa, Educación Básica (a la cual pertenecía la Subdirección General de Educación General Básica) y Enseñanzas Medias (a la que estaba adscrita la Subdirección General de Bachillerato). Dependían también de la Secretaría General, además del Gabinete Técnico, la Subdirección General de Perfeccionamiento del Profesorado, la Subdirección General de Enseñanzas Artísticas, el CIDE y el ITE. Por su parte, la Subdirección General de Organización y Automación dependía de la Secretaría General Técnica (que, a su vez, estaba encuadrada en la Subsecretaría del Ministerio). Esta estructura se mantuvo hasta noviembre de 1986, en que el Real Decreto 2352/1986, de 7 de noviembre, suprimió las Direcciones Generales de Educación Básica y de Medias y creó las Direcciones Generales de Centros Escolares y de Renovación Pedagógica. Por este mismo Real Decreto se suprimió también la Subdirección General de Organización y Automación.

¹⁵ En 1987, la Ley de Presupuestos Generales del Estado había incluido un programa presupuestario denominado *Programa de Nuevas Tecnologías Aplicadas a la Educación*, que unificaba las partidas destinadas a los proyectos *Atenea* y *Mercurio*. Este último había dado comienzo en junio de 1985; su objetivo era fomentar el uso del vídeo como recurso didáctico en los centros escolares. Con el tiempo, a estos dos proyectos se les uniría el Proyecto *Mentor*, que se implantó también de forma experimental en los centros públicos de EGB y enseñanzas medias del territorio MEC. Se inició formalmente en 1991 (aunque los primeros cursos no se pusieron en marcha

PNTIC quedó conformado por los denominados servicios centrales, de los cuales formaban parte la directora del proyecto, sus asesores técnicos y el personal administrativo, por los monitores y por los coordinadores *de equipos pedagógicos*.

El PNTIC consolidó su situación en 1989, cuando fue formalmente creado en el seno del gabinete de la Secretaría de Estado de Educación¹⁶, lo que llevó aparejados ciertos cambios en el Programa. El primero fue la constitución de dos áreas, la de *Medios Informáticos* y la de *Medios Audiovisuales* que se hicieron cargo, respectivamente, de la coordinación de los proyectos *Atenea* y *Mercurio*. Las modificaciones más importantes, sin embargo, fueron las que afectaron a la figura del monitor y a la evaluación y el seguimiento del proyecto.

Así, se reforzó la formación de los monitores –denominados, desde 1989, *asesores de formación en medios informáticos*¹⁷– con objeto de potenciar su papel de agentes de innovación educativa y de responsables, en su ámbito funcional, de la evaluación, seguimiento y difusión de las experiencias. Para ello, se pusieron en marcha diversos cursos de actualización y se distribuyeron materiales de apoyo entre los centros (documentos de consulta, software y vídeos explicativos).

El plan de formación de los monitores, coordinadores y miembros de equipos pedagógicos se vio seriamente comprometido por la falta de software educativo. La elaboración y depuración de los prototipos de programas informáticos resultaban más lentas de lo que se había previsto inicialmente, lo que hacía que el material no estuviese disponible a tiempo para su incorporación al diseño de la asignatura. Para solventar este problema, el Ministerio de Educación y Ciencia firmó un convenio con el Ministerio de Industria y con el Centro para el Desarrollo Tecnológico e Industrial con objeto de promover la producción de software por parte de empresas especializadas. Además, el PNTIC, en colaboración con el CIDE, comenzó a convocar a partir de 1988 un concurso nacional destinado a premiar las aplicaciones y los programas informáticos elaborados por los profesores.

El seguimiento y la evaluación del proyecto ocuparon también, como ya hemos señalado, un lugar muy destacado en el proceso de implantación. De hecho, el proyecto *Atenea* fue evaluado en tres ocasiones. La primera de las evaluaciones, que dio comienzo en 1988 y se refirió también al proyecto *Mercurio*, fue realizada por el ICE de la Universidad de Murcia bajo la dirección del profesor Juan Manuel Escudero. La evaluación tuvo dos partes, una primera, denominada de progreso, que se desarrolló de manera paralela a la implantación y cuyos resultados quedaron recogidos en un informe publicado en 1989

hasta 1993) en la Comunidad Autónoma de Castilla y León con el objetivo de atender las necesidades formativas de aquellas zonas con insuficiente oferta educativa empleando para ello las nuevas tecnologías de la información y la comunicación.

¹⁶ Orden de 7 de noviembre de 1989. En aquel momento, la responsable del gabinete era Elena Angulo Aramburu.

¹⁷ En concreto, desde la aprobación del *Plan Marco de Formación Permanente del Profesorado*. Véase MINISTERIO DE EDUCACIÓN Y CIENCIA (1989b).

y una segunda, llevada a cabo durante el curso 1989-90, al término de la fase experimental.

En 1988, un equipo de expertos de la OCDE realizó una segunda evaluación por encargo de las autoridades educativas españolas. Este estudio se basó en los informes proporcionados por el PNTIC y en los datos recabados por el equipo evaluador durante su estancia en España¹⁸.

La última de las evaluaciones se desarrolló en 1992 y fue realizada por el Servicio de Inspección Técnica de Educación (SITE) del Ministerio de Educación y Ciencia. Como la primera, se refirió también al proyecto *Mercurio*.

Por su parte, las tareas de seguimiento del proceso de implantación y el asesoramiento técnico a los centros quedaron en manos de los monitores y del Gabinete de Seguimiento y Evaluación del PNTIC, a cuyo frente estuvo Carmen Candiotti durante toda la fase experimental. Precisamente, la existencia de este gabinete provocó algunos enfrentamientos entre los monitores y los Servicios Centrales del Programa. Un nutrido grupo de monitores llegó incluso a «amotinarsen» y a solicitar la supresión del Gabinete por considerarlo un instrumento de control que minusvaloraba el trabajo que venían desarrollando en el proyecto¹⁹.

No fue esta, sin embargo, la única dificultad a la que los monitores tuvieron que hacer frente. Al igual que los coordinadores de los equipos pedagógicos, desarrollaron siempre su labor fuera del horario lectivo; tan solo en algunas provincias consiguieron dos horas de «liberación» para desempeñar sus funciones y nunca lograron que la administración les proporcionase un seguro de accidente, a pesar de los continuos desplazamientos que tenían que realizar entre los centros y los CEP. La disponibilidad de tiempo era el principal problema de los monitores:

Cuadro VI.5
Principales problemas encontrados por los monitores para el desempeño de sus funciones

Disponibilidad de tiempo	40,9%
Motivación de los equipos	18,2%
Interferencias de los CEP	7,9%
Falta de apoyo material	7,9%
Escasa formación como monitores	4,6%

Fuente: ESCUDERO (1989: 126)

¹⁸ OCDE (1991).

¹⁹ Ninguno de los monitores entrevistados recuerda que surgieran estas discrepancias que, sin embargo, fueron señaladas tanto por la dirección del Programa y la responsable del Gabinete como por algunos asesores técnicos.

Además, los monitores se convirtieron en el blanco de las críticas de los coordinadores de los equipos pedagógicos, quienes se quejaban de no poder solventar los problemas técnicos que surgían ni resolver las dudas planteadas por los profesores. La causa, para los coordinadores, no era otra que la deficiente formación que habían recibido de los monitores.

2. EL PROCESO DE EVALUACIÓN DEL PROYECTO *ATENEA*

2.1. La iniciativa y el impulso de la evaluación

La evaluación del proyecto *Atenea* fue el fruto de la iniciativa de dos actores de naturaleza político-administrativa. En primer lugar, la idea estaba ya presente en el diseño inicial del proyecto que había realizado su directora, Elena Veiguela. En segundo lugar, desde la Secretaría General de Educación se había establecido, tras la llegada de Alfredo Pérez Rubalcaba, la conveniencia de evaluar todos aquellos programas que tuvieran carácter experimental, con objeto de comprobar su viabilidad y su eficacia²⁰. Tanto Veiguela como Pérez Rubalcaba entendían que era imprescindible conocer en profundidad la forma en que se estaba desarrollando el proyecto, cuáles eran sus puntos débiles y cuáles eran los resultados que se estaban alcanzando. Solo si se disponía de este tipo de información se podrían tomar las oportunas medidas correctoras durante el propio proceso de implantación y se podría proceder a la generalización del proyecto con cierta garantía de éxito.

La *posición formal* que ocupaban Veiguela y, muy especialmente, Pérez Rubalcaba y su *entorno de oportunidad*, no podían sino hacer que su iniciativa prosperase. Pero en la decisión de Elena Veiguela influyó no solo su *interés* por conocer el grado de desarrollo del proyecto sino también por legitimarlo. Cuando se inició la evaluación no existía aún el Programa de Nuevas Tecnologías y los proyectos no se habían consolidado en la estructura organizativa del Ministerio. Aunque el proyecto *Atenea* tenía para las autoridades educativas un evidente interés al margen de cuáles fueran las conclusiones de la evaluación, lo cierto es que aquel podía verse fuertemente reforzado si el estudio evaluativo arrojaba resultados positivos.

La intervención de Pérez Rubalcaba y de Elena Veiguela en el proceso de evaluación fue aún más allá. Veiguela, en una muestra de su actitud abierta ante la evaluación y de su interés porque esta fuese realmente útil al programa, encomendó su realización a un equipo de expertos –por tanto, un actor técnico– a cuyo frente se encontraba Juan Manuel Escudero²¹, quien había

²⁰ Esta consigna no afectó a las evaluaciones de reforma experimental de enseñanzas medias y del programa de integración puesto que ambas habían dado comienzo antes de que se creara la Secretaría General de Educación, en abril de 1985.

²¹ Además de Escudero integraban el equipo de evaluación Manuel Ato García, también de la Universidad de Murcia, y Amador Guarro Pallá, de la Universidad de La Laguna.

dirigido los cursos formativos para monitores en los primeros momentos de vida del proyecto. Veiguela se hizo así con el control sobre una de las decisiones configuradoras –la determinación del papel de los actores–, al tiempo que los Servicios Centrales del Programa se erigieron en clientes de la evaluación, lo que les situó en una posición de evidente superioridad frente al equipo evaluador.

Además de la antigua vinculación al proyecto del equipo de Escudero, en la decisión de Elena Veiguela influyeron otros factores, como ella misma apunta. En primer lugar, este equipo era de los pocos que, en aquellos momentos, tenía alguna experiencia en nuevas tecnologías aplicadas a la educación y en evaluación de programas. En segundo lugar, su visión del proyecto era muy similar a la que, inicialmente, tenían los Servicios Centrales. Tanto unos como otros entendían que la informática no era el fin sino el medio para promover la innovación educativa en los centros escolares. Las restantes propuestas de evaluación que llegaron al PNTIC fueron desestimadas precisamente por hacer un excesivo énfasis en la dimensión más técnica, en detrimento del plano pedagógico del proyecto. Sin embargo, según Elena Veiguela, con el paso del tiempo se puso de manifiesto que el equipo evaluador prestaba también, en realidad, más atención a la primera que al segundo.

A juicio de Escudero, las divergencias en torno a esta cuestión estaban claras desde un principio. En su opinión, eran los miembros de los Servicios Centrales quienes consideraban que el proyecto representaba básicamente una oportunidad para la introducción de las nuevas tecnologías en los centros. El equipo evaluador entendía, por el contrario, que el proyecto debía servir, ante todo, para promover la innovación educativa.

Sea como fuere, en lo que sí coinciden Elena Veiguela y Juan Manuel Escudero es que la participación del equipo de la Universidad de Murcia en la evaluación repercutió de manera muy negativa en las relaciones entre actores. Veiguela afirma que, en aquellos primeros momentos, nadie cuestionó esa decisión ni advirtió los problemas que podía acarrear²².

2.2. El diseño: «pluralismo» y negociación

2.2.1. *Los objetivos y el modelo de evaluación*

Aunque la evaluación dio comienzo en 1988, la relación entre los evaluadores y la dirección del proyecto no se formalizó hasta 1990, cuando la investiga-

²² Luis del Blanco afirma, sin embargo, que desde el PNTIC se era plenamente consciente de las consecuencias que podía tener esta medida y, por ello, se solicitó de la OCDE la realización de una evaluación externa.

ción entraba ya en su fase final²³. El presupuesto asignado para la evaluación fue de 9 millones de pesetas y procedió ya de fondos propios del programa²⁴.

El objetivo principal de la evaluación, en su primera fase, fue realizar un estudio exploratorio y valorativo que permitiera determinar (i) el grado de desarrollo alcanzado por el proyecto y (ii) en qué medida dicho proyecto estaba propiciando la renovación pedagógica en las escuelas. Todo ello, como ya hemos señalado, con vistas a la adopción de las oportunas medidas correctoras durante el proceso de implantación. En el caso de la evaluación final el propósito era conocer el grado de consecución de los objetivos que el Ministerio pretendía alcanzar a través del proyecto. En la definición de los objetivos de la evaluación intervinieron el equipo evaluador y los Servicios Centrales del Programa pasando, así, a convertirse en *actores centrales*.

La propuesta sobre el modelo evaluativo que presentaron los evaluadores fue negociada con los Servicios Centrales y con los monitores. En la decisión final influyeron, básicamente, tres factores: la concepción (esto es, el *marco cognitivo*) que los Servicios Centrales y el equipo evaluador tenían del proyecto, los propios objetivos de la evaluación y los *intereses* y preferencias metodológicas de su director.

Para Escudero, el mejor modo de determinar cuál era el grado de desarrollo del proyecto y de reflejar su complejidad era recabar información acerca de un amplio número de variables y dimensiones, mediante diferentes tipos de técnicas cuantitativas y cualitativas de recogida y análisis de datos. Ello permitía, además, contrastar la información obtenida por unos y otros medios. Los dos cuadros que figuran a continuación dan cuenta de las variables que fueron consideradas, respectivamente, en la evaluación de progreso y en la evaluación final.

Cuadro VI.6
Variables consideradas en la evaluación de progreso del proyecto *Atenea* e informantes consultados sobre ellas

Monitores
Valoración global del proyecto y de su estado de desarrollo
Valoración de la formación recibida
Dificultades para el cumplimiento de funciones
Grado de ilusión en el proyecto
Grado de capacitación para el cumplimiento de determinadas funciones
Sugerencias y propuestas acerca del proyecto

²³ El convenio de colaboración entre la Secretaría de Estado de Educación y la Universidad de Murcia, al que no pudimos tener acceso, fue firmado el 10 de marzo de 1990.

²⁴ En opinión de Juan Manuel Escudero, el presupuesto fue ajustado pero suficiente.

LA EVALUACIÓN DEL PROYECTO *ATENEA* PARA LA INTRODUCCIÓN DE LOS ORDENADORES
EN LAS AULAS

Coordinadores
Valoración global del proyecto y de su estado de desarrollo
Características de los equipos pedagógicos y dinámica de trabajo
Motivos de la incorporación de los miembros de los equipos al proyecto
Expectativas de los miembros de los equipos acerca del proyecto y evolución
Estadio de desarrollo del proyecto en el aula
Funciones atribuidas a los medios informáticos
Frecuencia de uso de medios informáticos en las aulas
Posible cambio operado en los alumnos por el uso de medios informáticos (rendimiento, actitudes, grado de interés, formación)
Repercusiones del proyecto sobre el trabajo del profesor en el aula
Relaciones entre coordinadores y monitores
Apoyo esperado y recibido de los monitores
Repercusiones del proyecto sobre la organización y las infraestructuras de los centros
Actitudes de profesores no implicados con respecto al proyecto
Frecuencia de asistencia a actividades de los CEP
Autovaloración
Necesidades de formación
Necesidad de reducción de horarios y número de horas de reducción que precisan
Directores de centros
Grado de información sobre el proyecto
Aportaciones del proyecto
Directores de CEP
Problemas de organización generados por el proyecto
Grado de aceptación del proyecto
Problemas ocasionados por el proyecto
Profesores
Características de los equipos pedagógicos y dinámica de trabajo
Motivos de la incorporación de los miembros de los equipos al proyecto
Expectativas de los miembros de los equipos acerca del proyecto y evolución
Estadio de desarrollo del proyecto en el aula
Funciones atribuidas a los medios informáticos
Frecuencia de uso de medios informáticos en las aulas
Posible cambio operado en los alumnos por el uso de medios informáticos (rendimiento, actitudes, grado de interés, formación)
Repercusiones del proyecto sobre el trabajo del profesor en el aula
Relaciones entre profesores y monitores
Apoyo esperado y recibido de los monitores
Repercusiones del proyecto sobre la organización y las infraestructuras de los centros
Actitudes de profesores no implicados con respecto al proyecto
Frecuencia de asistencia a actividades de los CEP

Fuente: Elaboración propia a partir de ESCUDERO (1989).

Cuadro VI.7
Variables consideradas en la evaluación final del proyecto *Atenea*

Creación de condiciones para el surgimiento, adopción e implicación en los proyectos
Creación de una filosofía compartida sobre el proyecto y existencia de una visión estratégica para su desarrollo
Tipología y calidad de los materiales ofertados
Formación y capacitación del profesorado, los coordinadores y los monitores
Coordinación interna de los equipos pedagógicos
Tipo de asesoramiento y apoyo externo ofertado a los profesores y a los equipos
Planificación y desarrollo evolutivo de los proyectos
Reestructuración organizativa y funcionamiento de los equipos pedagógicos
Relaciones de los proyectos con los centros
Práctica educativa de aula

Fuente: ESCUDERO (1992).

En opinión de Escudero, la evaluación debía propiciar la participación de todos aquellos que tuvieran algún interés en el programa pero no solo en calidad de informantes sino también de decisores «para identificar desde los objetivos hasta los ámbitos de la evaluación, para decidir los procesos y metodologías que se iban a utilizar y el análisis y la interpretación de los resultados». El equipo evaluador, en consonancia con esta perspectiva, apostó por un modelo de evaluación que diera cabida no solo a los responsables del programa sino también a los monitores.

Los actores que podían tener *a priori* algún interés en su evaluación –los CEP, los centros escolares, las familias o los alumnos– quedaron, también en esta evaluación, excluidos de este proceso decisorio. Su papel en la evaluación fue, por tanto, eminentemente *pasivo*, si bien es cierto que los monitores se ocuparon de hacer llegar a los evaluadores las opiniones de los directores y del profesorado²⁵.

A pesar de la ausencia de estos actores, el enfoque de la evaluación del proyecto *Atenea* podría, pues, calificarse de mixto y parcialmente pluralista. Pero, de algún modo, estos rasgos de la evaluación se terminarían volviendo en su contra. La presencia de un importante número de actores con intereses contrapuestos y la indefinición de los papeles que habían de corresponder a algunos de ellos comprometieron seriamente el proceso evaluativo hasta llegar casi a impedir su conclusión.

2.2.2. *Las técnicas de recogida y análisis de datos*

En consonancia con el pluralismo metodológico que presidía la evaluación, se recurrió a diversos tipos de técnicas para la recogida y el análisis de datos. Se

²⁵ En el caso de la evaluación final se recogieron también las opiniones de los alumnos participantes en el proyecto.

trataba, de este modo, de salvar las deficiencias que presentaban tanto las técnicas cuantitativas como las cualitativas y de obtener información más contrastada, completa y fiable. El diseño de los diferentes instrumentos correspondió al equipo evaluador, aunque su contenido fue después negociado con los monitores y con los Servicios Centrales.

Concretamente, los métodos empleados para la recogida de información fueron la encuesta, la entrevista semiestructurada y la observación participante²⁶. En la fase final se realizaron también estudios de caso²⁷ y análisis documental.

Las pruebas fueron aplicadas en la totalidad de centros participantes en el proyecto pero no siempre a todos los sujetos. Así, fueron encuestados todos los monitores y directores de CEP (un total de 317) pero tan solo tres profesores participantes (1.210) y otros tantos no participantes (799) por centro. Se realizaron, además, un total de 88 entrevistas a profesores implicados en el proyecto.

La información obtenida en las entrevistas fue codificada y analizada mediante la utilización de técnicas de reducción. El análisis de los datos procedentes de los cuestionarios se orientó, por un lado, a conocer la distribución de los porcentajes de respuesta para cada uno de los ítems y, por otro, a determinar la presencia de asociación entre variables²⁸. Puesto que de la aplicación de pruebas multivariantes no se obtuvieron conclusiones significativas, se decidió incluir únicamente en el informe los resultados expresados en porcentajes y los derivados de la aplicación de la prueba de chi-cuadrado.

Por lo que se refiere a las observaciones de aula, se acudió a la técnica del relato en la situación de observación. A pesar de que se había previsto que todos los monitores se responsabilizaran de los procesos de observación participante de forma que se cubrieran todas las áreas curriculares y todos los niveles educativos, diversas circunstancias hicieron que este objetivo no pudiera lograrse y que la calidad de las observaciones realizadas no cubriese las expectativas de los evaluadores.

Si algo hemos de destacar con respecto a la recogida de datos en la evaluación del proyecto *Atenea* es la importancia que los evaluadores concedieron a la calidad y al interés de la información, aun cuando los procedimientos utilizados para su obtención no cumplieran totalmente con los requisitos de las técnicas estadísticas. Los evaluadores eran plenamente conscientes de los problemas que esto podía ocasionar y, por ello, advirtieron en el informe de evaluación de la cautela con la que había que interpretar los resultados.

²⁶ Los cuestionarios y las guías de las entrevistas y la observación pueden consultarse en ESCUDERO (1989: 203 y ss.).

²⁷ Los estudios de caso fueron realizados por nueve investigadores contratados *ex profeso* para esta tarea.

²⁸ Concretamente, se recurrió a la prueba de chi-cuadrado (χ^2) y a procedimientos multivariantes, como el análisis factorial y el análisis de clusters.

2.2.3. *Las opiniones de los actores implicados como criterios de valor de la evaluación*

Al igual que sucedió en las evaluaciones de los programas de integración y de reforma experimental de enseñanzas medias, tampoco en la del proyecto *Ate-nea* encontramos ninguna referencia explícita a la cuestión de los criterios de valor. Estos se deducen de los propios objetivos de la evaluación. Así, se consideró que el éxito del proyecto dependía de la medida en que hubiera propiciado la innovación pedagógica en los centros y del grado de desarrollo que hubiera alcanzado. Puede decirse pues que, indirecta e implícitamente, los criterios de valor fueron determinados, al igual que los objetivos en los que se basaban, a través de un proceso de negociación entre actores. Pero lo realmente significativo de este proceso de evaluación y una de sus principales diferencias con las otras dos evaluaciones analizadas es que *el cumplimiento de estos objetivos no se determinaba a partir de pruebas objetivas, del criterio de la administración educativa o de la opinión de los evaluadores, sino de las apreciaciones y valoraciones de los principales protagonistas del proyecto*. Ello constituye buena prueba del talante que presidió el proceso de evaluación y de la importancia que en él se concedió a los actores que estaban implicados en el mismo.

2.3. El desarrollo del proceso evaluativo²⁹

2.3.1. *El trabajo de campo*

Una vez que todos los actores que tomaron parte en el diseño de las pruebas de recogida de datos habían llegado a un cierto consenso acerca de su contenido, estas fueron editadas y enviadas a quienes iban a ocuparse de su aplicación, los monitores del proyecto. Estos se encargaron de hacer llegar los cuestionarios a cada uno de los centros así como de realizar las entrevistas y la observación participante. Para ello recibieron del equipo evaluador la formación necesaria.

Aunque la aplicación de las pruebas correspondía a los monitores, dentro del equipo evaluador existía un coordinador por cada una de las técnicas utilizadas que se responsabilizaba de todo lo relativo a la recogida de datos. El equipo evaluador realizaba un primer análisis de los resultados obtenidos en el trabajo de campo para, posteriormente, someterlos a discusión con los monitores y los Servicios Centrales.

²⁹ El trabajo de campo de la evaluación de progreso se desarrolló entre los meses de febrero y mayo de 1988. Recordemos que, en el transcurso de la evaluación externa, dio comienzo la evaluación de la OCDE. A pesar de que las conclusiones de una y otra son parcialmente coincidentes, la evaluación de la OCDE presenta una imagen más positiva del proyecto que la que ofrece la de la Universidad de Murcia. Sin duda, esta es la razón por la que su acogida por parte del PNTIC fue mucho más favorable. Es de destacar que el informe de la OCDE subrayara el rigor en el tratamiento de datos de la evaluación de la Universidad de Murcia.

La participación de los monitores en la recogida de datos marcó un auténtico punto de inflexión en el proceso de evaluación. En opinión de Elena Veiguela, Carmen Candiotti y de los monitores a los que entrevistamos, este hecho contribuyó a enturbiar las ya de por sí difíciles relaciones que estos mantenían con los miembros del equipo evaluador.

Como se recordará, los monitores venían desempeñando sus funciones en condiciones nada favorables. Parece lógico, pues, que no fuera de su agrado tener que asumir tareas adicionales. Ciertamente, con ello se les obligaba a desempeñar distintos papeles (diseñadores y aplicadores) y a ocupar tan pronto una posición central como a situarse en la periferia del proceso decisorio.

Para el equipo evaluador, el que los monitores se hicieran cargo del trabajo de campo se justificaba por razón del propio enfoque participativo del estudio. Sin embargo, para Elena Veiguela y para Carmen Candiotti, la responsabilidad que los monitores tuvieron que asumir fue excesiva.

A pesar de que estas circunstancias crearon un clima de no poca tensión y de que, ya para entonces, era evidente que el «reparto» de papeles entre los diferentes actores no había sido el adecuado³⁰, la evaluación fue desarrollándose tal y como estaba previsto. Se concluyeron los informes de progreso, de cuya redacción se ocupó el propio Escudero, y se presentaron ante los monitores y los Servicios Centrales del programa para su discusión.

2.3.2. *El informe de progreso*³¹

La estructura del informe de progreso del proyecto *Atenea* refleja perfectamente cuáles eran los objetivos de la evaluación y cuáles eran las cuestiones que tenían mayor relevancia para el equipo evaluador. Los diferentes apartados se dedican a presentar y analizar los datos obtenidos de los miembros de los equipos pedagógicos y de los monitores, a estudiar las relaciones entre los distintos actores y a conocer su valoración sobre el proyecto y sobre su grado de desarrollo. El informe de progreso avanzaba ya algunas conclusiones:

- El proyecto gozaba de un elevado grado de aceptación por parte de todos los sujetos implicados. Para todos ellos, el proyecto posibilitaba no solo la adquisición de equipos informáticos sino también la formación del profe-

³⁰ Esta es la opinión tanto de Juan Manuel Escudero como de los miembros de los Servicios Centrales entrevistados.

³¹ Hemos analizado únicamente los datos contenidos en este informe ya que el final nunca llegó a publicarse y no fue posible obtenerlo de los protagonistas de la evaluación. Disponemos, no obstante, de un documento presentado por el director de la evaluación en unos encuentros nacionales sobre las nuevas tecnologías en la educación y en el que expuso los resultados más significativos obtenidos en la fase final. Creemos, pues, que la información disponible permite hacerse una idea bastante aproximada del proceso de evaluación y de los factores que marcaron su desarrollo. Véase ESCUDERO MUÑOZ (1992).

sorado, la motivación de los alumnos y la introducción de prácticas de innovación pedagógica. El proyecto, además, se encontraba «muy enraizado en la estructura de funciones y expectativas mantenidas por los profesores, coordinadores y monitores» (Escudero 1989: 164).

- La práctica en el aula era todavía bastante escasa, debido principalmente a la falta de software y de guías orientativas para la integración curricular, a la insuficiente capacitación del profesorado, así como al escaso y asistemático apoyo prestado por los monitores.
- El nivel educativo constituía la variable de mayor influencia sobre el desarrollo del proyecto. EGB parecía ser un ámbito más propicio que BUP o FP.
- En definitiva, los datos de la evaluación señalaban que el proyecto se encontraba aún en los momentos iniciales de la fase de movilización y adopción³².

El informe de progreso incluyó también algunas recomendaciones, con vistas a posibilitar el uso de la información producida. En primer lugar, se consideraba necesario proceder a una clarificación de las percepciones, expectativas y funciones de los distintos agentes participantes en el proyecto, así como incrementar el número y la calidad de los ordenadores y los programas informáticos. Asimismo, se entendía que era preciso fomentar un nuevo tipo de formación para los monitores, coordinadores y miembros de equipos pedagógicos que permitiera el desarrollo de experiencias de innovación educativa en las aulas. Se señalaba también la conveniencia de determinar una secuencia lógica y flexible para la integración de las nuevas tecnologías de la información en el currículum y de potenciar el sistema de recogida de información sobre las distintas experiencias y el análisis y la reflexión sobre las mismas.

El estilo del informe de progreso del proyecto *Atenea* es, en línea con el enfoque que presidió la evaluación, radicalmente distinto al que preside los informes de las evaluaciones de reforma de enseñanzas medias y del programa de integración. Está redactado en términos fácilmente comprensibles para quien no domine las cuestiones de carácter técnico y metodológico. En él se realiza un análisis objetivo del proyecto, desde una postura de cierto distanciamiento del tema. Las conclusiones se derivan fielmente de los datos obtenidos en la aplicación de las pruebas y no se encuentran pronunciamientos, ni siquiera velados, a favor o en contra del programa.

³² Esta fase suele caracterizarse por la puesta en marcha de actividades de difusión del proyecto, de formación y sensibilización de los profesionales, de preparación y/o elaboración de materiales y de clarificación de las funciones que a cada figura le corresponde desempeñar. La *fase de desarrollo* supone la puesta en marcha del proyecto en las aulas, la progresiva integración de las propuestas de innovación en la práctica educativa y la asimilación de dichas propuestas por parte de los sujetos participantes. La tercera y última, denominada *de institucionalización*, implicaría la incorporación plena de los principios y estrategias del proyecto a la realidad de los centros. Sobre las fases de los proyectos y la valoración de los monitores y los coordinadores de los equipos pedagógicos acerca del grado de desarrollo del proyecto, véase ESCUDERO (1989: 32 y 33).

El malestar de los monitores se hizo aún más patente cuando se presentaron los resultados que arrojaba el informe de progreso. Según este, el apoyo que recibieron profesores y coordinadores de equipos pedagógicos por parte de los monitores no respondió a las expectativas que aquellos tenían (Escudero 1989: 104 y 105). Se señalaba, además, que el avance cualitativo del proyecto dependía, en gran medida, de la redefinición de las funciones que los monitores habían de desempeñar: era necesario que estos dejaran de insistir en la formación inicial del profesorado para concentrar esfuerzos y recursos en el desarrollo de proyectos de innovación.

Los monitores no ocultaron su descontento con estas conclusiones. Consideraban, por un lado, que era una demostración palpable del escaso valor que se atribuía a su trabajo y a la dedicación que venían mostrando. Por otro, entendían injusto que se les criticara por no cumplir ciertas funciones cuando, a su juicio, no habían recibido la formación necesaria para ello (Escudero 1989: 127).

En opinión de Escudero, la reacción de los monitores era algo perfectamente comprensible porque, como él mismo señala, «siempre que hay una actuación educativa que toca cuestiones personales, que toca roles, que toca funciones, responsabilidades, que toca en último caso la profesionalidad de los que están implicados, las resistencias son inevitables».

El enfrentamiento entre monitores y evaluadores terminó por afectar negativamente al proceso de evaluación, como no podía ser de otra forma. La relación del equipo de evaluación con los Servicios Centrales –que cerraron filas en torno a los monitores, con quienes compartían el interés por la continuidad y consolidación del proyecto– fue derivando paulatinamente hacia el «cansancio y la desconfianza mutua»³³.

2.4. El final de la evaluación

2.4.1. *La presentación del informe final y la ruptura de las relaciones entre clientes y evaluadores*

El informe final, presentado por los evaluadores en 1990, no hizo sino confirmar lo que había apuntado ya el informe de progreso. En él se ponía de manifiesto que, si bien el proyecto había alcanzado algunos de sus objetivos (como dotar a los centros del equipamiento informático necesario o familiarizar al profesorado con las nuevas tecnologías) subsistían todavía algunas deficiencias:

- La falta de adecuación y de especificidad de algunos de los materiales y programas informáticos seguía siendo una realidad denunciada por la mayoría del profesorado, a pesar del notable incremento registrado, a nivel general, tanto en el volumen como en la calidad de la oferta.

³³ Así es, al menos, como Elena Veiguela describe las relaciones en aquellos momentos. La visión de Escudero es muy similar, aunque no se manifieste de manera tan explícita.

- El papel desempeñado en la práctica por los coordinadores de los equipos pedagógicos distaba de ser el esperado. Su trabajo se orientaba a la resolución de los problemas técnicos que iban surgiendo en los centros o en la mera gestión del proyecto, más que a la constitución y dinamización de los equipos pedagógicos, la elaboración de proyectos o la evaluación de las experiencias.
- El trabajo de los monitores se centraba todavía en la «formación inicial del profesorado, en la difusión de programas y materiales y en la atención a problemas variados no estrictamente pedagógicos muchas veces y menos en el apoyo al desarrollo, focalizado sobre la planificación, seguimiento, análisis, valoración de experiencias y dinamización pedagógica de los equipos» (Escudero 1992: 263).
- Los proyectos no parecían haber contribuido a mejorar los procesos educativos de los centros, probablemente porque su desarrollo había estado vinculado exclusivamente a los equipos pedagógicos.

En definitiva, la conclusión principal de la evaluación era que el proyecto se encontraba aún en un estadio muy inicial y que su incidencia pedagógica en los centros era todavía muy escasa.

Las relaciones entre el equipo evaluador y los Servicios Centrales, que habían ido deteriorándose progresivamente, se volvieron insostenibles en el momento de la presentación del informe final. Todos los miembros de los Servicios Centrales a los que entrevistamos insisten en que el informe ofrecía una imagen poco reconocible, distorsionada y en exceso negativa del proyecto. En opinión de Carmen Candiotti, en el informe se hacía evidente que «el equipo de evaluación tenía una idea bastante ingenua de lo que era el PNTIC, porque pretendía que en unos años de experimentación este programa hubiera alcanzado unos resultados claramente discernibles. Los evaluadores olvidaron las dificultades que entraña todo cambio educativo, especialmente si se trata de una herramienta tan novedosa como la informática». Elena Veiguela, por su parte, tenía la impresión de que los evaluadores «no se daban cuenta de que habían recibido un encargo» y se empeñaron en hacer una evaluación «que no era la que se había acordado o que, por lo menos, si era eso lo que se había acordado, no se había hecho lo suficientemente explícito como para que todo el mundo supiese de qué estábamos hablando. El talante «pretendidamente» progresista de los evaluadores hacía que considerasen cualquier actuación de la administración educativa como «intrínsecamente mala». Para Luis del Blanco «los evaluadores no tuvieron en cuenta las dificultades de alcanzar el objetivo del proyecto, ni el contexto en el que se tenía que implantar. No destacaban suficientemente los logros del proyecto».

La dirección del proyecto solicitó a los evaluadores que se realizasen en el informe algunos cambios que suavizasen la imagen tan desfavorable que se ofrecía del proyecto. Sin embargo, Escudero decidió mantener sus valoraciones tal y como las había formulado inicialmente porque «no sabía cómo matizar algunas expresiones sin maquillar los datos».

El informe final nunca fue publicado³⁴. Los Servicios Centrales, que tenían en sus manos el control de esta decisión, adujeron problemas presupuestarios para hacer frente a los gastos de edición. Escudero cree sin embargo –y esta es también nuestra opinión– que la verdadera razón fue el descontento de los responsables del proyecto con las conclusiones que arrojaba el informe final.

La directora del proyecto considera hoy que la evaluación no debería haberse llevado a cabo en un momento tan inicial de la vida del proyecto y que fue un error que quienes habían tomado parte en el mismo se hiciesen después cargo de su evaluación.

2.4.2. *La influencia de los resultados en las percepciones sobre la utilidad de la evaluación*

A la vista de estos datos, resulta lógico que quienes formaban parte de los Servicios Centrales consideren que la evaluación dirigida por el profesor Escudero no tuvo ninguna repercusión positiva. Todos ellos insisten en que la verdadera evaluación del proyecto *Atenea* fue la realizada por la OCDE, «un estudio que entraba más al detalle y permitía conocer el proyecto con mayor profundidad»³⁵.

La opinión de los monitores entrevistados es aún más crítica que la de los Servicios Centrales. El proceso de evaluación, afirman, lejos de tener algún efecto positivo no hizo sino perjudicar y entorpecer el desarrollo del proyecto³⁶.

Juan Manuel Escudero tiene, como se puede intuir, una opinión mucho más favorable sobre el impacto de la evaluación. Desde su perspectiva, la evaluación hizo que los monitores desarrollaran una mentalidad más crítica acerca de lo que estaban haciendo y de cuál debía ser su papel en el proyecto y sirvió también para promover el debate y la discusión en torno al proyecto entre los miembros de los servicios centrales. Cree también que la experiencia fue interesante para el propio equipo evaluador, ya que le permitió conocer de primera mano las dificultades propias de los procesos de evaluación e, incluso, que puede ser de utilidad para quienes realicen evaluaciones sobre proyectos relacionados con las nuevas tecnologías.

³⁴ Según Luis del Blanco, ello hizo que se magnificaran los problemas surgidos durante la evaluación cuando, en sí mismos, no tenían tanta importancia.

³⁵ Al igual que Escudero, no coincidimos con esta opinión. La evaluación de la OCDE no pasó de ser una empresa de aval político del proyecto que ofrecía, además, un conocimiento mucho más superficial que el que proporcionaba la evaluación realizada por la Universidad de Murcia.

³⁶ Esta apreciación está lógicamente condicionada por su enfrentamiento con el equipo evaluador. Durante las entrevistas, el discurso de los monitores giraba insistentemente en torno a la deficiente formación que recibieron.

CAPÍTULO VII

CONCLUSIONES

Sobre la relevancia de la metaevaluación como instrumento al servicio de las evaluaciones

- *La metaevaluación constituye la culminación analítica de los procesos de evaluación.*

De la misma forma que las evaluaciones pueden ser consideradas como la conclusión analítica de los programas, los estudios de metaevaluación han de ser entendidos como el punto final, la culminación de esos procesos de análisis que son las evaluaciones. Ello no quiere decir que toda evaluación deba finalizar necesariamente con un proceso metaevaluativo, aunque parezca aconsejable tanto desde un punto de vista técnico como político.

- *La metaevaluación está llamada a desempeñar un papel central en el desarrollo y el progreso de la evaluación de políticas públicas.*

La contribución de las evaluaciones a la mejora de los procesos decisivos que están en la base de políticas y programas y al fortalecimiento de la *accountability* pasa, necesariamente, por la revisión y el perfeccionamiento de los instrumentos de los que se sirve, por contrastar la validez de sus juicios y por comprobar que, tal y como ha sido diseñada, resulta útil para lograr «mejores» políticas. Es en este punto en el que la metaevaluación cobra sentido y adquiere relevancia. Gracias a ella podemos determinar la «calidad» de las evaluaciones, distinguir las que han sido o son útiles de las que no lo son contribuyendo, de ese modo, al desarrollo de la práctica evaluativa, a fortalecer los pilares teóricos sobre los que dicha práctica se sustenta y al avance y consolidación del campo profesional de la evaluación. El trabajo metaevaluativo requiere de un cierto grado de desarrollo –cuantitativo y cualitativo– de la evaluación y, por tanto, de un acervo suficiente de experiencias que permitan realizar análisis concluyentes sobre las evaluaciones y/o sobre la función de evaluación. No quiere ello decir, sin embargo, que la metaevaluación se justifi-

que únicamente en contextos en los que la evaluación esté sólidamente asentada. Antes al contrario, su importancia se revela mayor en las etapas iniciales o intermedias del desarrollo institucional de la evaluación, cuando en éste no se han dado aún pasos firmes.

- *La metaevaluación se revela como una excelente herramienta para comprender el «estilo de evaluación» de los poderes públicos.*

Como se ha apuntado en diversas ocasiones (Nilsson y Hogben 1983; Bustelo 2001 y 2002; García Sánchez 2003), la metaevaluación no ha de ocuparse solo de evaluaciones concretas sino que puede y debe atender a las políticas de evaluación, lo que implica dirigir la mirada hacia cómo se conciba esta desde los poderes públicos, de qué modo se articula el sistema de evaluación, a qué actores se abre la participación, qué medidas se adoptan para que las evaluaciones contribuyan al diseño y planificación de las políticas y qué control se pretende ejercer sobre la información producida.

Efectivamente, el análisis pormenorizado de tres de las evaluaciones de programas educativos más importantes y representativas de las llevadas a cabo en España nos ha proporcionado información de extraordinaria utilidad sobre las actitudes y el comportamiento de las autoridades educativas en dichos procesos de evaluación. Hemos podido, de este modo, conocer los enfoques metodológicos por los que se han decantado y la importancia que han concedido a ciertas cuestiones centrales en la evaluación, como la explicitación de los criterios de valor, la participación de los distintos *stakeholders*, la propiedad de la información producida o la utilidad y la difusión de los resultados.

Sobre las fortalezas del esquema de metaevaluación

- *La metaevaluación realizada ha permitido ratificar la pertinencia de las premisas de partida y la utilidad del esquema propuesto.*

Nuestra investigación tomó como punto de partida la tesis de los enfoques pluralistas que subraya el carácter político de las evaluaciones (Weiss 1987; Palumbo 1987; MacDonald 1977; House 1980, 1993; Guba y Lincoln 1981, 1987, 1989; Schwandt 2005; Kushner 2005; Greene 2003; MacDonald y Kushner 2005; Stake 2006). Estas son, es evidente, procesos de investigación, pero también *empresas eminentemente políticas* que se desarrollan *en un contexto también político*. Parece lógico, por tanto, que la metaevaluación no pueda quedar reducida al mero examen de la calidad técnica de las evaluaciones. En esta idea se sustentan las dos premisas de partida sobre las que se construye el esquema metaevaluativo propuesto y cuya conveniencia ha quedado confirmada en esta investigación:

- (i) El reconocimiento (analítico) de la doble dimensión de las evaluaciones. La metaevaluación ha de analizar y enjuiciar tanto la dimensión técnico-metodológica –aquella a la que cabe asociar los objeti-

vos formales, el enfoque evaluativo y el metodológico o los métodos de recogida y análisis de datos— como la dimensión política de las evaluaciones —vinculada a la presencia de intereses contrapuestos no solo con respecto a la política evaluada sino también en relación con el propio proceso de evaluación.

- (ii) Las decisiones que los distintos actores adoptan y las relaciones que se establecen entre ellos afectan e inciden en las dos dimensiones contribuyendo, de ese modo, a configurar los procesos de evaluación.

Paralelamente, el esquema se conforma a partir de un conjunto de cuestiones asociadas a la dimensión técnico-metodológica y a la dimensión política de las evaluaciones —por tanto, subdimensiones— que han sido identificadas a partir de los distintos enfoques de metaevaluación revisados y cuyo análisis permite reconstruir, comprender y enjuiciar el proceso evaluativo. La aplicación del esquema consistirá en determinar, analizar y valorar las decisiones (o no decisiones) que los actores implicados adoptaron con respecto a cada una de las subdimensiones consideradas en el esquema.

Asimismo, la utilización del esquema nos ha permitido obtener no solo un retrato fiel de los procesos evaluativos sino una comprensión de los mismos más profunda que la que nos hubiera proporcionado un enfoque meramente descriptivo basado de estándares. Así pues, la principal fortaleza del esquema propuesto radica en su notable capacidad explicativa, en la medida en que no se limita a dar cuenta de qué sucede en las evaluaciones sino a desvelar por qué sucede. La orientación de la propuesta presentada es eminentemente práctica, si bien creemos que también puede contribuir a enriquecer la reflexión teórica. El esquema permite orientar tanto a los metaevaluadores como a los evaluadores a la hora de identificar qué decisiones han contribuido o pueden contribuir al éxito o al fracaso de un proceso evaluativo, al afectar positiva o negativamente la utilidad, la pertinencia, el rigor o la viabilidad. El objetivo último es lograr aprendizajes que permitan mejorar la gestión —política, técnica, económica, administrativa— de las evaluaciones.

- *Ello revela, a su vez, el interés de conjugar la Ciencia Política y la evaluación de políticas públicas no solo para construir un esquema capaz de comprender y enjuiciar los procesos de evaluación sino para descubrir nuevos objetos de estudio a la Ciencia Política¹.*

Los estudios clásicos de Ciencia Política nos han permitido ahondar en las causas por las que cabe afirmar el carácter político de las evaluaciones —la presencia de actores con distintos intereses que luchan por influir en la adopción de las asignaciones de valor (Easton 1969/2006) —fortaleciendo

¹ Nos referimos a las evaluaciones, ámbitos más políticos, si cabe, que las propias políticas públicas.

así las tesis pluralistas— redescubrir a los actores y acuñar una definición precisa y operativa del propio término de *actor*. Los trabajos sobre los procesos decisorios, por su parte, nos han ofrecido las claves para identificar los factores que influyen en la adopción de decisiones por parte de los actores, determinar su perfil y el poder que consiguen ejercer en las evaluaciones. Por último, en los distintos modelos y enfoques de metaevaluación hemos hallado las referencias necesarias —esto es, las cuestiones y los ítems a los que la bibliografía especializada ha venido prestando atención— para articular un esquema que posibilita la evaluación de las evaluaciones.

Sobre la centralidad del análisis de actores en el esquema de metaevaluación

- *Se confirma la importancia de que la metaevaluación —muy en especial la que se emprenda desde una visión comprensiva, desde el análisis de políticas públicas o desde la Ciencia Política— aborde, como elemento central, el análisis de actores.*

La investigación realizada ha puesto de relieve que una evaluación estará tanto más politizada cuanto mayor sea la divergencia de intereses en su seno. Piénsese, por ejemplo, en las notables diferencias que, en este sentido, se dan entre las evaluaciones de la reforma experimental de enseñanzas medias y la del proyecto *Atenea*. El análisis de esos intereses —en definitiva, de los actores que los defienden— se vuelve, pues, imprescindible para comprender y valorar adecuadamente los procesos de evaluación.

Ello no supone desconocer el importante papel que pueden desempeñar los elementos estructurales, entendidos aquí como los escenarios o contextos en los que los actores se desenvuelven.

- *La metaevaluación, en tanto incorpore un análisis de actores, ha de atender a (i) la identificación de los actores implicados en los procesos de evaluación, (ii) su caracterización o determinación de su perfil y (iii) el análisis de las decisiones en las que cada uno de ellos participa.*

La metaevaluación debe proporcionar información sobre cuáles son los actores implicados en el proceso evaluativo, cuáles son sus características distintivas, cuáles son las decisiones que adoptan y/o en las que participan y en qué medida esas decisiones han configurado una evaluación capaz de describir y analizar apropiadamente el programa.

- *La utilización de los criterios de naturaleza del actor y modalidad de participación en la evaluación ha demostrado ser operativa, en el análisis empírico, a la hora de trazar el perfil de cada uno de los actores presentes en las evaluaciones analizadas.*

La diversidad de formas en que los actores intervienen en un proceso de evaluación hacía necesaria la utilización de diversos criterios de clasificación. Recurrimos, por ello, a dos categorías:

(i) La *naturaleza* del actor nos permite distinguir entre actores *político-administrativos* (las instituciones, los partidos políticos o las autoridades político-administrativas), *actores técnicos* (por ejemplo, las unidades organizativas y los expertos o profesionales que se encuentran vinculados al programa y/o a su evaluación) y, finalmente, *actores corporativos-asociativos* (los centros docentes, los centros de profesores, las asociaciones o sindicatos de profesores, los colectivos de padres o alumnos o los movimientos educativos).

(ii) La *modalidad de participación* tiene que ver con la forma en la que los actores intervienen en el proceso evaluativo: *activa* (en el caso de que participen en calidad de decisores) o *pasiva* (si actúan, fundamentalmente, como meros objetos o destinatarios de las decisiones adoptadas por otros). Dentro de la categoría de *actores activos* es posible distinguir, a su vez, entre quienes ocupan una *posición central* en el proceso decisorio al tomar parte en la adopción de las decisiones configuradoras de la evaluación y quienes se sitúan en una *posición periférica* al permanecer al margen de las mismas, sea por voluntad propia, sea porque son excluidos por otros actores.

Si bien se confirma la utilidad del segundo de los criterios a la hora de determinar cómo se distribuye el poder en las evaluaciones y cómo ello incide sobre la calidad de estas, como veremos más adelante, no sucede lo mismo en lo que respecta al primero. Las categorías identificadas a partir de este resultan insuficientes a la hora de encuadrar a algunos de los actores. Es lo que sucede con los Servicios Centrales del PNTIC en la evaluación del proyecto *Atenea*, la DGEM en la evaluación de la reforma experimental de enseñanzas medias y el equipo de evaluación de la integración. Si, *a priori*, los dos primeros podrían haber quedado encuadrados en la categoría de *actores político-administrativos* y el tercero en la correspondiente a *actores técnicos*, el análisis empírico apunta más bien hacia un equilibrio entre factores político-ideológicos y científico-técnicos en el seno de cada uno de ellos. Dicho en otros términos, y por paradójico que pudiera parecer, el comportamiento del equipo de evaluación de la integración estuvo más condicionado por elementos políticos que el de la DGEM o el de los Servicios Centrales del PNTIC. Nos encontramos, así pues, con actores de *naturaleza híbrida*, aunque ello no impidió que sus integrantes actuaran de manera unitaria frente al resto de actores intervinientes en los escenarios de la evaluación.

Sobre el análisis de las decisiones y los roles de los actores como elemento clave de la metaevaluación

- *La capacidad de un actor para intervenir en la adopción de una decisión y el contenido de esta dependen de la acción combinada de cinco factores: (i) la posición del actor, (ii) sus intereses, (iii) su sistema de creen-*

cias, (iv) sus habilidades y (v) el entorno de oportunidad en el que despliega su actuación.

Entendemos por *posición* aquella que se deriva del cargo o de su situación en la estructura informal de la organización. Al hablar de *intereses* nos referimos a las ambiciones, deseos, predisposiciones y expectativas que inducen a una persona a actuar y a hacerlo en un sentido determinado. Bajo la denominación *sistema de creencias* se incluyen tanto las asunciones causales sobre ciertos fenómenos que limitan las alternativas que los actores consideran eficaces o útiles –los llamados *paradigmas cognitivos*– como las creencias y los valores que hacen que los actores opten por ciertas decisiones por entender que son aceptables y legítimas desde el punto de vista moral. Las *habilidades*, por su parte, son aquellas capacidades del individuo que se derivan de su personalidad, su formación académica y su experiencia vital y profesional y que le llevan a poner en práctica sus ideas de manera eficaz y exitosa. Influye, asimismo, el denominado *entorno de oportunidad*, es decir, el contexto en el que se desarrolla la actuación del actor y que le brinda recursos y posibilidades pero también restricciones.

La importancia que cada uno de estos factores adquiere en la práctica varía extraordinariamente de un caso a otro; por tanto, la determinación de su influencia requiere de un análisis empírico. Así por ejemplo, en la evaluación de la reforma experimental de enseñanzas medias, Mariano Álvaro, a pesar de no partir de una posición formal privilegiada, logró participar en la adopción de las decisiones configuradoras valiéndose de sus *habilidades* –su excelente formación académica y su amplia experiencia como investigador– e *intereses* –dirigir la primera gran empresa de evaluación (no de investigación) que se acometería desde las instituciones educativas y hacerlo, además, sobre un programa claramente novedoso–. Por supuesto, el *entorno* le fue también favorable –su proximidad a Julio Carabaña y el interés, para las autoridades ministeriales, de la evaluación de uno de los programas más emblemáticos de la política educativa.

Sin embargo, en la evaluación del programa de integración, la *posición* de Álvaro Marchesi, en tanto que Subdirector General de Educación Especial, fue clave, aunque influyeran también sus *habilidades* y su *interés* para rentabilizar políticamente el programa y el *entorno de oportunidad* le fuera más que favorable con el apoyo directo del ministro Maravall.

- *De los cuatro factores mencionados, son las creencias y los intereses de los actores los que mayor influencia tienen en el contenido de sus decisiones –y, por tanto, en la evaluación– llegando, en algunos casos, a gozar de mayor importancia que las consideraciones de carácter científico o técnico.*

Así, Mariano Álvaro y Julio Carabaña optaron por un enfoque metodológico de corte cuantitativo y por determinar la calidad de la reforma experimental a partir del rendimiento de los alumnos no porque enten-

dieran que era lo más adecuado desde el punto de vista metodológico sino porque era el enfoque *en el que ellos creían* y el que había predominado en su formación. No parece probable que hubieran barajado otra posibilidad que no fuera un enfoque cuantitativo.

Del mismo modo, Marchesi y el denominado equipo de diseño de la evaluación de la integración se decantaron, sobre el papel, por el modelo iluminativo² por su *interés* en acudir a un enfoque que gozara de cierto reconocimiento en la comunidad educativa y no despertara las reticencias que provocaban, ya entonces, los más clásicos y cuantitativistas.

- *Consiguientemente, la metaevaluación ha de ocuparse no sólo de las características explícitas de las evaluaciones sino también de aquellas otras, si cabe más reales, pero que permanecen implícitas, con objeto de determinar en qué medida unas y otras reflejan los intereses y las creencias de los actores.*
- *Se confirma la importancia de las denominadas decisiones configuradoras por cuanto son ellas las que confieren a la evaluación sus características distintivas.*

El análisis empírico llevado a cabo en la presente investigación ha permitido confirmar la especial relevancia de cierto tipo de decisiones, concretamente, aquellas que afectan (i) al objeto y los objetivos de la evaluación, (ii) al enfoque metodológico, (iii) a los criterios de valor, (iv) a la selección de los actores participantes y al rol que se les atribuye (v) al contenido, estilo de los informes y al grado de difusión de la información.

- *En consecuencia, puede concluirse que el poder de un actor en un proceso de evaluación se vincula a la intensidad de su participación en la adopción de las decisiones configuradoras.*

Efectivamente, en las evaluaciones analizadas, los actores que poseen una mayor cuota de poder son aquellos que logran adoptar –unilateralmente o en colaboración con otros actores– las decisiones que confieren a la evaluación sus características distintivas. El grado de poder de un actor será tanto mayor cuanto mayor sea el número de decisiones configuradoras en cuya adopción pueda participar. La presencia en ella de un mayor número de actores hace de la relativa al proyecto *Atenea* la evaluación más pluralista de las analizadas.

- *Son las decisiones relativas al enfoque metodológico las que han tenido una mayor relevancia en las evaluaciones estudiadas, al condicionar el resto de decisiones configuradoras.*

² Véase la nota a pie núm. 25 del capítulo V.

El enfoque metodológico elegido tuvo una incidencia más que evidente en el papel desempeñado por los equipos de evaluación y en el grado de «pluralismo» de las evaluaciones. Tanto en el caso de la evaluación de la reforma experimental de enseñanzas medias como en la del programa de integración, los evaluadores asumieron el rol tradicional de expertos y se hicieron con el control de todas y cada una de las decisiones configuradoras. En ninguno de los dos casos se promovió la participación de otros actores en la toma de decisiones —y en la evaluación de la integración a pesar de afirmarse que se partía de un enfoque iluminativo—. Por el contrario, el papel que en la evaluación del proyecto *Atenea* correspondió a los evaluadores se asemejó más al de un consejero cuya valoración acerca del programa no se consideraba ni la única ni la más importante; un consejero que, además, promovía la participación de otros actores en el afán de llegar a una decisión lo más consensuada posible. Aunque no se trata de una evaluación «plenamente» pluralista, lo cierto es que, de los tres casos analizados, el de la evaluación del proyecto *Atenea* fue el que contó con un mayor número de participantes y en el que, además, la implicación de estos fue mayor. Todo ello evidencia que las decisiones sobre el enfoque metodológico tienen implicaciones que exceden el plano técnico.

La perspectiva metodológica afectó también a los *criterios de valor* de la evaluación. En las dos primeras evaluaciones analizadas, los evaluadores entendieron que el criterio que permitía juzgar la bondad de los programas era el cumplimiento, por parte de los alumnos y de los centros, de las directrices fijadas por el Ministerio. Cuanto más se aproximaran los resultados de unos y otros a lo establecido por la administración mayor sería el éxito que cabría atribuir al programa. El que los objetivos no fueran no solo no cuestionados sino, al contrario, tomados como referencia para los criterios de valor, redujo de manera considerable la capacidad crítica de las evaluaciones. La reflexión de Ewald (1986) sobre las políticas públicas resulta especialmente apropiada en este punto: solo si se ponen en tela de juicio sus objetivos se pueden cuestionar realmente los valores y principios en los que dichas políticas se basan.

La importancia que en la evaluación del proyecto *Atenea* se concedía a la participación se refleja incluso en el tema de los criterios de valor, aunque estos no se formularan de manera explícita. De hecho, el elemento que servía de base para juzgar la calidad del proyecto no era tanto el cumplimiento de sus objetivos cuanto la opinión que los distintos implicados tenían acerca del mismo. La evaluación del proyecto *Atenea* fue, por esta razón, mucho «menos amable» con el proyecto de lo que lo habían sido las evaluaciones de la integración y de reforma de enseñanzas medias.

El *estilo de los informes* de evaluación está igualmente relacionado con el enfoque metodológico elegido. De nuevo, los informes de las evaluaciones de los programas de reforma experimental de enseñanzas medias y de integración presentan unas diferencias claras con el de la evaluación del proyecto *Atenea*. El estilo de ambos, muy en especial el de enseñanzas medias, era eminentemente técnico y difícilmente accesible para quien no tuviera conocimientos estadísticos; por otro lado, la imagen que ofrecían de sus respectivos programas, sobre todo el informe de evaluación del programa de integración, no puede ser más favorable. Sin embargo, el estilo narrativo del informe de progreso de la evaluación del proyecto *Atenea* hace que sea comprensible para un público no especializado lo que, a su vez, favorece su utilización.

- *En todos los casos, los actores político-administrativos y técnicos intervinieron de manera activa mientras que a los actores de naturaleza corporativa-asociativa correspondió siempre un papel pasivo en los procesos de evaluación. Asimismo, ambas categorías compartieron el control de las decisiones configuradoras.*

Si los *actores corporativo-asociativos* se limitaron, en todos los casos, a actuar como meros receptores de las decisiones adoptadas por los *actores político-administrativos* o por los *técnicos* fue porque, desde estos últimos, no hubo interés en promover su participación activa. No disponemos de datos definitivos acerca de las razones por las que ello fue así, al menos en las evaluaciones del programa de integración y del proyecto *Atenea*. Probablemente, los actores políticos y los actores técnicos temieran con ello incorporar perspectivas excesivamente críticas o consideraron, por sus «creencias» metodológicas, que las evaluaciones eran tarea exclusiva de los expertos o de la administración. Desde nuestra perspectiva, su participación activa hubiera enriquecido de modo considerable las evaluaciones, al ofrecer una imagen más panorámica y equilibrada de los programas.

Sobre la utilidad, el pluralismo y la viabilidad de las evaluaciones

- *El análisis empírico permite confirmar las dificultades existentes, en la práctica, para determinar el grado de utilización de las evaluaciones.*

De las relacionadas con la evaluación, pocas cuestiones hay tan espinosas y controvertidas como las relacionadas con la utilización de los procesos evaluativos, como demuestra el elevado número de estudios dedicados a este tema (Patton 1986, 1997; Chelmsky 1987, 1995; Guba y Lincoln 1981; Alkin 1985 y Alkin, Daillak y White 1979). La principal dificultad no estriba solo en determinar si la evaluación ha contribuido a la mejora del programa (*uso instrumental*, según la definición de Leviton y Hughes 1981), ha servido para promover la reflexión sobre el mismo (*uso conceptual*, para Rossi y Freeman 1989) o si únicamente se ha empleado como instrumento de legitimación o de apoyo (*uso persuasivo*, en

palabras de Leviton y Hughes 1981). En el caso de las evaluaciones formativas, aquellas que discurren paralelamente al programa, el problema fundamental consiste en discriminar cuáles de las decisiones adoptadas son atribuibles a la evaluación y cuáles al propio proceso de implantación del programa. Esta dificultad es aún mayor cuando se trata de un programa educativo de naturaleza abierta y experimental en el que las formas en que dicho programa se concreta en la práctica difieren sustancialmente, ya desde el inicio, entre sí; en definitiva, cuando, en la realidad, existen tantos programas como centros participantes. Para conocer la incidencia de la evaluación sobre un programa de estas características sería necesario o bien limitar el análisis a aquellos cambios que afectan a los aspectos más generales del programa (tarea que resulta, desde un punto de vista estrictamente operativo, bastante compleja por el número y la dispersión de los centros participantes) o bien estudiar las modificaciones introducidas tan sólo en un número reducido de aplicaciones del programa. Partiendo de estos datos y considerando, como sugiere Chelimsky (1995), que cada actor tiene una idea distinta de lo que es una evaluación «útil», entendemos que una forma de salvar los problemas planteados es acudir a las valoraciones globales que sobre el uso de la evaluación hacen los distintos implicados en el programa.

- *El análisis empírico pone de manifiesto, asimismo, el escaso impacto instrumental de las evaluaciones estudiadas, tal y como destacan los responsables de los programas y la mayoría de los miembros de los centros educativos.*

Efectivamente, aunque las razones esgrimidas por cada uno de ellos fueran distintas (desacuerdo con el enfoque metodológico en el caso de José Segovia o disconformidad con los resultados de la evaluación en el de Elena Veiguela), lo cierto es que todos los directores insisten en señalar que las evaluaciones tuvieron una escasa, cuando no nula, incidencia en el perfeccionamiento y mejora de los programas o proyectos de los que eran responsables. La valoración de Álvaro Marchesi acerca de las repercusiones de la evaluación sobre el programa fue, como es obvio, muy positiva.

Por su parte, la negativa opinión de los profesores con respecto al impacto de las evaluaciones sobre los programas tiene que ver, como se ha puesto de manifiesto en el análisis empírico, no tanto con el hecho de que su implicación en los procesos evaluativos fuera eminentemente pasivo cuanto con el tipo de información producida en ellos: estadísticas extensas y complejas, en exceso generales, en las que no se reconocían y a las que no veían ninguna utilidad. Por el contrario, tanto los profesores como los miembros de los equipos multiprofesionales consideraron mucho más provechosos los datos obtenidos mediante técnicas cualitativas.

- *Los datos que arroja la presente investigación no permiten confirmar la relación entre el grado de «pluralismo» de las evaluaciones y la utilidad*

de sus resultados, al contrario de lo que algunos autores como Patton (1986; 1997) han venido señalando.

En efecto, la opinión que sobre la utilidad de las evaluaciones tienen algunos de los actores que tomaron parte en el proyecto *Atenea*, la experiencia que cabría considerar como más «pluralista», no es mucho más favorable que la de quienes estuvieron implicados en una evaluación como la de la reforma experimental de enseñanzas medias. En otras palabras, la percepción de los actores acerca de la utilidad de las evaluaciones no dependió de que estas fueran más o menos participativas ni de que dichos actores hubieran estado directamente implicados en la adopción de las decisiones configuradoras. Así, como ya hemos señalado, la intervención de los Servicios Centrales y de los monitores en la toma de decisiones de la evaluación del proyecto *Atenea* no sirvió para mejorar su negativa opinión sobre la utilidad del proceso. No obstante es cierto que, en términos generales, la valoración que hacen los actores centrales sobre la utilidad es más positiva que la de quienes actuaron como meros informantes o quedaron situados en la periferia del proceso decisorio.

- *Se puede advertir, igualmente, que son las evaluaciones más «pluralistas» las que ofrecen una imagen más completa del programa.*

No es casual que la evaluación que presenta un mayor grado de pluralismo, la relativa al proyecto *Atenea*, sea también la que proporciona una descripción más panorámica del proyecto, mientras que la evaluación menos pluralista, la de reforma experimental de enseñanzas medias, sea la que ofrece una descripción más parcial y reduccionista del programa.

En principio, es más probable que una evaluación refleje el punto de vista de un actor si este ha podido intervenir e influir en las decisiones adoptadas durante el proceso. Una evaluación ofrecerá un «retrato» tanto más detallado y completo del programa cuanto mayor sea el número de actores que han podido participar en ella en calidad de decisores. No obstante lo anterior, no parece posible afirmar que exista una clara relación de «causalidad» entre estos dos elementos.

- *Resulta conveniente definir, en los momentos iniciales de la evaluación, el papel que habrá de desempeñar cada uno de los actores implicados.*

En la evaluación del proyecto *Atenea*, la viabilidad del proceso se vio comprometida no solo porque monitores y Servicios Centrales estuvieran en desacuerdo con los resultados sino por la indefinición del papel que había de corresponder a los primeros. Puesto que, con frecuencia, en las evaluaciones se reproducen (atenuándose o magnificándose) los problemas surgidos previamente en los programas, una clara (pero flexible) delimitación de roles puede evitar desencuentros en el transcurso del proceso de evaluación. Entendemos, incluso, recomendable que de ello se deje constancia, en su caso, en el propio contrato de evaluación.

ANEXO I

REFERENCIAS BIBLIOGRÁFICAS Y FUENTES DOCUMENTALES

- AGUADO ODINA, M.J. (1991), *Evaluación de programas pertenecientes al plan experimental para la reforma de la educación infantil*. Madrid: Universidad Nacional de Educación a Distancia.
- AGUILAR VILLANUEVA, L.F. (2000), *El estudio de las políticas públicas*. México, D.F.: Miguel Ángel Porrúa.
- AGUILERA, M.J. *et al.* (1990), *Evaluación del programa de integración escolar de alumnos con deficiencias*. Madrid: Centro de Investigación y Documentación Educativa / Centro Nacional de Recursos para la Educación Especial.
- AINSCOW, M. y MUNCEY, J. (1989), *Meeting Individual Needs in the Primary School*. Londres: Fulton.
- ALKIN, M.C. (1985), *A Guide for Evaluation Decision Makers*. Beverly Hills, CA: Sage.
- (1990), *Debates on Evaluation*. Newbury Park, CA: Sage.
- ALKIN, M.C., ed. (2004), *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks, CA: Sage.
- ALKIN, M.C., DAILLAK, R. y WHITE, P. (1979), *Using Evaluations: Does Evaluation Make a Difference?*, Beverly Hills, CA: Sage.
- ALONSO, L.E. (1994), «Sujeto y discurso: el lugar de la entrevista abierta en las prácticas de la sociología cualitativa», en J.M. Delgado y J. Gutiérrez, coords., *Métodos y técnicas cualitativas de investigación en ciencias sociales*. Madrid: Síntesis.
- ALONSO HINOJAL, I. (1982), «Nuevas investigaciones educativas: continuidad y riesgo», en Subdirección General de Investigación Educativa, *Investigaciones educativas de la red INCIE-ICE, 1978-1982*. Madrid: Ministerio de Educación y Ciencia.
- ÁLVARO PAGE, M. (2010), «La evaluación externa de la experiencia» en M. Menor y J. Moreiro, coords., *La reforma experimental de las enseñanzas medias (1983-1987). Crónica de una ilusión*. Las Rozas: Wolters Kluwer.
- ÁLVARO PAGE, M. *et al.* (1988), *Evaluación externa de la reforma experimental de enseñanzas medias (I)*. Madrid: Centro de Investigación y Documentación Educativa.

- (1990), *Evaluación externa de la reforma experimental de enseñanzas medias (II)*. Madrid: Centro de Investigación, Documentación y Evaluación.
- (1992), *Evaluación externa de la reforma experimental de enseñanzas medias (III)*. Madrid: Centro de Investigación, Documentación y Evaluación.
- ALVIRA, F. (1991), *Metodología de la evaluación de programas*. Madrid: Centro de Investigaciones Sociológicas.
- ALZAGA, O., dir., *Comentarios a las leyes políticas: Constitución Española de 1978*. Madrid: Ed. Revista de Derecho Privado / Editores de Derecho Reunidos.
- AMERICAN EVALUATION ASSOCIATION (2004), *Guiding Principles for Evaluators*. En línea: <<http://www.eval.org/Publications/aea06.GPBrochure.pdf>> (consulta: 11 febrero 2011).
- ANGULO RASCO, J.F. (1993), «Evaluación del sistema educativo. Algunas respuestas críticas», en VV.AA., *Evaluar las evaluaciones*, número monográfico de *Cuadernos de Pedagogía*, 219: 8-15.
- ARGYRIS, C. (1994), «Good Communication that Blocks Learning», *Harvard Business Review*, 72, 4: 77-85.
- ASHFORD, D. (1976), *Democracy, Decentralization and Decision in Sub-National Politics*. Beverly Hills, CA: Sage.
- BAJO SANTOS, N. (2004), «Los cinco ejes de la Ley Orgánica de Calidad de la Educación (LOCE)», *Anuario Jurídico y Económico Escurialense*, XXXVII: 741-768.
- BALLART, X. (1992), *¿Cómo evaluar programas y servicios públicos? Aproximación sistemática y estudios de caso*. Madrid: Instituto Nacional de Administración Pública.
- BANCO MUNDIAL (2007), *Sourcebook for Evaluating Global and Regional Partnership Programs. Indicative Principles and Standards*, Washington D.C. En línea: <<http://siteresources.worldbank.org/EXTGLOREGPARPROG/Resources/sourcebook.pdf>> (consulta: 5 febrero 2011).
- BERG, B.L. (1998), *Qualitative Research. Methods for the Social Sciences*. Boston, MA: Allyn & Bacon.
- BERICAT, E. (1998), *La integración de los métodos cuantitativo y cualitativo en la investigación social*. Barcelona: Ariel.
- BERK, R., ed. (1981), *Educational Evaluation: The State of the Art*. Baltimore: The Johns Hopkins University Press.
- BERNSTEIN, I. y FREEMAN, H.E. (1975), *Academic and Entrepreneurial Research: Consequences of Diversity in Federal Evaluation Studies*. Nueva York: Russell Sage.
- BLAU, P.M. (1963), *The Dynamics of Bureaucracy*. Chicago: The University of Chicago Press.
- BONAL, X. (1998), «La política educativa: dimensiones de un proceso de transformación (1976-1996)», en R. Gomá y J. Subirats, coords., *Políticas públicas en España. Contenidos, redes de actores y niveles de gobierno*. Barcelona: Ariel.

- BUCHANAN, J.M. y TULLOCK, G. (1962), *The Calculus of Consent*. Ann Arbor: University of Michigan Press.
- BUCHANAN, J.M. y TOLLISON, R.D., eds. (1972), *Theory of Public Choice. Political Applications of Economics*. Ann Arbor: University of Michigan Press.
- BUSTELO, M. (1999), «Diferencias entre evaluación e investigación: una distinción necesaria para la identidad de la evaluación de programas», *Revista Española de Desarrollo y Cooperación*, 4: 9-29.
- (2001), *La evaluación de las políticas públicas de igualdad de género de los gobiernos central y autonómicos en España: 1995-1999*. Director: Rafael Bañón i Martínez. Tesis doctoral, Universidad Complutense de Madrid, Departamento de Ciencia Política y de la Administración II.
- (2002), «Metaevaluation as a tool for the improvement and development of the evaluation function in public administrations», ponencia presentada en la *V Bienial Conference of the European Evaluation Society*, Sevilla, 10-12 de octubre.
- CACIAGLI, M. (1986), *Elecciones y partidos en la transición española*. Madrid: Centro de Investigaciones Sociológicas.
- CAMPBELL, D.T. (1956), *Leadership & Its Effects Upon the Group*. Columbus: Ohio State University Press.
- CAMPBELL, D.T. y STANLEY, J.C. (1973), *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- CARACELLI, V.J. y GREENE, J.C. (1997), «Crafting mixed-method evaluation designs», en J.C. Greene y V.J. Caracelli, eds., *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms*, *New Directions for Evaluation*, 74: 19-32.
- CARDOZO BRUM, M. (2006), *La evaluación de políticas y programas públicos. El caso de los programas de desarrollo social en México*. México D.F.: Miguel Ángel Porrúa editores.
- CENTRO DE INVESTIGACIÓN, DOCUMENTACIÓN Y EVALUACIÓN (1992), *El Sistema Educativo Español*. Madrid.
- (1995), *El Sistema Educativo Español*. Madrid.
- CHELIMSKY, E. (1987), «Linking Program Evaluation to User Needs», en D.J. Palumbo, *The Politics of Program Evaluation*. Newbury Park, CA: Sage.
- (1995), «New Dimensions in Evaluation», en R.C. Rist y R. Picciotto, eds., *Proceedings of the 1994 World Bank Conference*.
- (1997), «The Coming Transformations in Evaluation», en E. Chelimsky y W. Shadish, eds., *Evaluation for the 21st Century*. Thousand Oaks, CA: Sage.
- CHELIMSKY, E. y W. SHADISH (1997), eds., *Evaluation for the 21st Century*. Thousand Oaks, CA: Sage.
- CHRISTIE, C.A. y ALKIN, M. (2004), «An Evaluation Theory Tree», en M. Alkin, ed., *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks, CA: Sage.

- COBO SUERO, J.M. (1995), «La realidad socioeducativa en España (1982-1993)» en M. Juárez, dir., *V Informe sociológico sobre la situación social en España: sociedad para todos en el año 2000*. Madrid: Fundación FOESSA.
- COHEN, E. y FRANCO, R. (1993), *Evaluación de proyectos sociales*. Madrid: Siglo XXI.
- COLL, C. (1988), «La reforma del bachillerato. Algunas consideraciones desde el punto de vista del curriculum», *Cuadernos de Pedagogía*, 163: 71-74.
- COLL, C. et al. (1990), «Un marco de referencia psicológico para la educación escolar: la concepción constructivista del aprendizaje y de la enseñanza», en C. Coll, J. Palacios y A. Marchesi, comps., *Desarrollo psicológico y educación, II. Psicología de la educación*. Madrid: Alianza.
- COLL, C., PALACIOS, J. y MARCHESI, A., comps. (1990), *Desarrollo psicológico y educación, II. Psicología de la educación*. Madrid: Alianza.
- CONSEJO DE UNIVERSIDADES (2001), *Plan nacional de evaluación de la calidad de las universidades. Informe final de la segunda convocatoria*. Madrid: Ministerio de Educación, Cultura y Deporte.
- (2002), *Plan nacional de evaluación de la calidad de las universidades. Informe final de la tercera convocatoria*. Madrid: Ministerio de Educación, Cultura y Deporte.
- CONSEJO ESCOLAR DEL ESTADO (1991), *Informe sobre el estado y situación del sistema educativo. Curso 1989-90*. Madrid.
- COOK, T.D. y CAMPBELL, D.J. (1979), *Quasi-experimentation. Design, Analysis. Issues for Field Experimentation*. Cambridge, MA: Houghton Mifflin Co.
- CROZIER, M. (1974), *El fenómeno burocrático*. Buenos Aires: Amorrortu.
- CROZIER, M. y FRIEDBERG, E. (1977), *L'acteur et le système*. París: Seuil.
- DAHL, R. (1957), «The Concept of Power», *Behavioral Science*, 2, 3: 201-215.
- DE ESTEBAN, J. y LÓPEZ GUERRA, L. (1982), *Los partidos políticos en la España actual*. Barcelona: Planeta / Instituto de Estudios Económicos.
- DE BLAS, P. (2010), «Entre la política y el aula: una reforma posible», en M. Menor y J. Moreiro, coords., *La reforma experimental de las enseñanzas medias (1983-1987). Crónica de una ilusión*. Las Rozas: Wolters Kluwer.
- DEGEVAL / BEYWL, W., eds. (2003), *Selected Comments to the Standards for Evaluation of the German Evaluation Society*. En línea: <<http://www.degeval.de/calimero/tools/proxy.php?id=71>> (consulta: 27 diciembre 2010).
- DE LA ORDEN, A. (1985), «Investigación evaluativa», en A. de la Orden, dir., *Investigación educativa. Diccionario de Ciencias de la Educación*, Madrid: Anaya.
- DE LA ORDEN, A., dir. (1985), *Investigación educativa. Diccionario de Ciencias de la Educación*. Madrid: Anaya.
- DELGADO, J.M. y GUTIÉRREZ, J. coords. (1994), *Métodos y técnicas cualitativas de investigación en ciencias sociales*. Madrid: Síntesis.
- DENZIN, N.K. (1978), *The Research Act*. Nueva York: McGraw-Hill.

- DÍAZ ROSALES, L.F. (2001), «La metaevaluación y su método», *Revista de Ciencias Sociales*, 92-93: 171-192.
- DIRECCIÓN GENERAL DE RENOVACIÓN PEDAGÓGICA (1988), *Evaluación de la integración escolar. 1er. Informe*. Madrid.
- (1989), *Evaluación de la integración escolar. 2º Informe*, Madrid.
- DOMÈNECH, J. y VIÑAS, J. (1992), «Movimientos de Renovación Pedagógica», *Cuadernos de Pedagogía*, 199: 72-81.
- DYE, T.R. (1987), *Understanding Public Policy*. Englewood Cliffs: Prentice Hall.
- EASTON, D. (1969/2006), *Esquema para el análisis político*. Buenos Aires: Amorrortu.
- ECHETA, G. (1991), «El programa de integración del MEC», *Cuadernos de Pedagogía*, 191: 72-75.
- ELLIOTT, J.H. (1994), *La investigación-acción en educación*. Madrid: Morata.
- EQUIPO DE EVALUACIÓN DE LA INTEGRACIÓN (1987), «Informe sobre la evaluación del proceso de integración: Marco general, hipótesis y fases del proyecto», *Revista de Educación*, 1: 7-44.
- ESCUADERO, T. (2003), «Desde los tests hasta la investigación evaluativa actual. Un siglo, el XX, de intenso desarrollo de la evaluación en educación», *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 9, 1. En línea: <http://www.uv.es/RELIEVE/v9n1/RELIEVEv9n1_1.htm> (consulta: 12 enero 2011).
- ESCUADERO MUÑOZ, J.M. (1989), *Proyecto Atenea. Informe de progreso. Fase exploratoria*. Madrid: Ministerio de Educación y Ciencia.
- (1992), «La evaluación de los proyectos *Atenea* y *Mercurio*», en A. García Manzano, *Las nuevas tecnologías en la educación*. Santander: ICE de la Universidad de Cantabria.
- (1995), «Sobre la utilidad de los informes de evaluación», en M.J. Sáez Brezmes, ed., *Conceptualizando la evaluación en España*. Madrid: Universidad de Alcalá.
- ESTEVE, J.M. (1987), *El malestar docente*. Barcelona: Laia.
- EWALD, F. (1986), *L'État-Providence*. París: Grasset.
- FEITO, R. (1991), «CONCAPA y CEAPA: dos modelos de intervención de los padres en la gestión de la enseñanza», *Educación y Sociedad*, 9: 35-58.
- (2003), «La LOCE y el fracaso escolar», *Investigación en la escuela*, 49: 61-77.
- (2005), «LOE: límites y logros del talante socialdemócrata», *Cuadernos de Pedagogía*, 37: 82-86.
- FERNÁNDEZ ENGUITA, M., et al. (1985), *La aplicación experimental de la reforma de enseñanzas medias: una evaluación sociológica*, mimeo.
- FERNÁNDEZ ENGUITA, M. (1987), «Reforma y contrarreforma en la enseñanza media», *Andecha Pedagógica*, 17: 6-10.

- FERNÁNDEZ MELLIZO-SOTO, M. (2001), *Socialismo, igualdad en la educación y democracia. La experiencia de González y Mitterrand*. Madrid: Instituto Juan March de Estudios e Investigaciones.
- FERNÁNDEZ-MIRANDA, A. (1983), «Artículo 27. Enseñanza», en O. Alzaga, dir., *Comentarios a las leyes políticas: Constitución Española de 1978*. Madrid: Ed. Revista de Derecho Privado / Editores de Derecho Reunidos.
- FERNÁNDEZ-RAMÍREZ, B. *et al.* (2010), «Metaevaluación del sistema de evaluación de las titulaciones universitarias en Andalucía», en VV.AA., *Gestión y Análisis de Políticas Públicas*, número monográfico coordinado por S. Kushner y E. García Sánchez, 3: 139-167.
- FETTERMAN, D.M., ed. (1984), *Ethnography in Educational Evaluation*. Beverly Hills, CA: Sage.
- FETTERMAN, D.M. y PITMAN, M.A., eds. (1985), *Beyond the Status Quo: Theory, Politics and Practice in Ethnographic Evaluation*. Washington, D.C.: Cato Institute
- FIELDING, N. (1994), «Qualitative Interviewing», en N. Gilbert, *Researching Social Life*. Londres: Sage.
- FIERRO, A. (1994), «Diseño y desafíos de la reforma educativa española», *Revista de Educación*, 305: 13-36.
- FIRESTONE, W.A. y HERRIOT, R.E. (1984), «Multisite Qualitative Policy Research: Some Design and Implementation Issues», en D.M. Fetterman, ed., *Ethnography in Educational Evaluation*. Beverly Hills, CA: Sage.
- GARCÍA FERRANDO, M., ALVIRA, F. e IBÁÑEZ, J. comps., (1996), *El análisis de la realidad social*. Madrid: Alianza.
- GARCÍA MANZANO, A. (1992), *Las nuevas tecnologías en la educación*, Santander: ICE de la Universidad de Cantabria.
- GARCÍA SÁNCHEZ, E. (2003), *La evaluación de programas de reforma educativa en España. Tres estudios de caso desde un enfoque de metaevaluación*. Director: Francisco J. Vanaclocha Bellver. Tesis doctoral, Universidad Complutense de Madrid: Departamento de Ciencia Política y de la Administración I.
- (2004), «La institucionalización de la evaluación educativa no universitaria en España», *Gestión y Análisis de Políticas Públicas*, 30-31: 29-40.
- (2007), «El concepto de actor. Reflexiones y propuestas para la Ciencia Política», *Andamios. Revista de Investigación Social*, 3, 6: 199-216.
- (2009), «»Metaevaluación de políticas públicas. Una visión desde la Ciencia Política», *Revista del CLAD. Reforma y Democracia*, 49: 127-154.
- GARCÍA FERRANDO, M., ALVIRA, F. e IBÁÑEZ, J., comps., (1996), *El análisis de la realidad social*, Madrid: Alianza.
- GILBERT, N. (1994), *Researching Social Life*. Londres: Sage.
- GIMENO SACRISTÁN, J. (1987), «Luces y sombras de un proyecto», *Cuadernos de Pedagogía*, 152: 84-88.

- (1993), «Reformas educativas: utopía, retórica y práctica», *Cuadernos de Pedagogía*, 209: 62-68.
- GLASS, G.V. y WORTHEN, B.R. (1971-1972), «Evaluation and Research: similarities and differences», *Curriculum Theory Network*, 8/9: 149-165.
- GOMÁ, R. y SUBIRATS, J. coords. (1998), *Políticas públicas en España. Contenidos, redes de actores y niveles de gobierno*. Barcelona: Ariel.
- GONZÁLEZ MUÑOZ, M.C. (1987), «La reforma de las enseñanzas medias: dificultades de un proceso experimental», *Andecha Pedagógica*, 17: 16-18.
- GREEN, J., CAMILLI, G. y ELMORE, P.B. (2006), *Handbook of Complementary Methods in Education Research*. Londres: Routledge.
- GREENE, J. (2001), «Dialogue in Evaluation: A Relational Perspective», *Evaluation*, 7, 2: 181-187.
- (2002), «With a Splash of Soda, Please: Towards Active Engagement with Difference», *Evaluation*, 8, 2: 259-266.
- (2003), «War and Peace... and Evaluation», *Studies in Educational Policy and Educational Philosophy*, 2, Universidad de Uppsala. En línea: <<http://www.upi.artisan.se/docs/Doc186.pdf>> (consulta: 12 octubre 2008).
- GREENE, J. y CARACELLI, V.J., eds. (1997), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms, New Directions for Evaluation*, 74: 19-32
- GREENE, J., BENJAMIN, L. y GOODYEAR, L. (2001), «The Merits of Mixing Methods in Evaluation», *Evaluation*, 7, 1: 25-44.
- GUBA, E.G. y LINCOLN, Y.S. (1981), *Effective Evaluations. Improving the Usefulness of Evaluation Results Through Responsive and Naturalistic Approaches*. San Francisco: Jossey-Bass.
- (1987), «The Countenances of Fourth-Generation Evaluation: Description, Judgement and Negotiation», en D.J. Palumbo, *The Politics of Program Evaluation*. Beverly Hills, CA: Sage.
- (1989), *Fourth Generation Evaluation*. Newbury Park, CA: Sage.
- GUBA, E.G. y STUFFLEBEAM, D.L. (1970), «Evaluation: The Process of Stimulating, Aiding and Abetting Insightful Action», *Monograph Series in Reading Education*, Indiana University, 1.
- HAMILTON, D., ed. (1977), *Beyond the Numbers Game*. Londres: McMillan.
- HAMMERSLEY, M. (1995), *The Politics of Social Research*. Londres: Sage.
- HARRÉ, R. (1981), «Philosophical Aspects of the Micro-macro Problem», en K.C. Knorr-Cetina y A.V. Cicourel, eds., *Advances in Social Theory and Methodology: Towards an Integration of Micro- and Macro-Sociologies*. Londres: Routledge.
- HARVARD FAMILY RESEARCH PROJECT (2005), «A conversation with Jennifer Greene», *The Evaluation Exchange*, XI, 3: 10-11.

- HIRSCHMAN, A.O. (1967), *Development Projects Observed*. Washington: Brookings Institute.
- HOUSE, E.R. (1980), *Evaluating with Validity*. Beverly Hills, CA: Sage (traducción al castellano en 2000, *Evaluación: ética y poder*. Madrid: Morata).
- (1993), *Professional Evaluation. Social Impact and Political Consequences*. Newbury Park, CA: Sage.
- (2006), «Democracy and Evaluation», *Evaluation*, 12,1: 119-127.
- HOUSE, E.R. y HOWE, K.R. (1999), *Values in Evaluation and Social Research*. Thousand Oaks: Sage.
- (2000), «Deliberative Democratic Evaluation Checklist». En línea: <http://www.wmich.edu/evalctr/checklists/dd_checklist.PDF> (consulta: 14 septiembre 2008)
- HOUSE, E.R. et al. (1977), *No Simple Answer. Critique of the Follow Through Evaluation*. Urbana: CIRCE / Universidad de Illinois.
- HUNEEUS, C. (1985), *La Unión de Centro Democrático y la transición a la democracia en España*. Madrid: Centro de Investigaciones Sociológicas.
- INSTITUTO NACIONAL DE ESTADÍSTICA (2011), *Enseñanzas anteriores a la universidad*. En línea: <<http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft13%2Fp001&file=inebase&L=0>> (consulta: 8 julio 2011).
- JEREZ MIR, M. (1999), *Ciencia política, un balance de fin de siglo*. Madrid: Centro de Estudios Políticos y Constitucionales.
- JIMÉNEZ JIMÉNEZ, B.J. ed., *Evaluación de Programas, Centros y Profesores*. Madrid: Síntesis.
- JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (1988), *The Personnel Evaluation Standards*. Thousand Oaks, CA: Corwin Press.
- (1994), *The Program Evaluation Standards*. Thousand Oaks, CA: Sage.
- JUÁREZ, M., dir. (1995), *V Informe sociológico sobre la situación social en España: sociedad para todos en el año 2000*. Madrid: Fundación FOESSA.
- KING, W.R. y CLELAND, D.L., eds. (1987), *Strategic Planning and Management Handbook*. Nueva York: Van Nostrand Reinhold.
- KNORR-CETINA, K.C. y CICOUREL, A.V., eds. (1981), *Advances in Social Theory and Methodology: Towards an Integration of Micro- and Macro- Sociologies*. Londres: Routledge.
- KUSHNER, S. (2005), «Democratic Theorizing: From Noun to Participle», *American Journal of Evaluation*, 26: 579-581.
- KUSHNER, S. y ADELMAN, C. (2006), «Program Evaluation: a Democratic Practice», en J.L. Greene, G. Camilli y P.B. Elmore, *Handbook of Complementary Methods in Education Research*. Londres: Routledge.

- KUSHNER, S. y GARCÍA SÁNCHEZ, E. (2010), «Educación, política y políticas de evaluación educativa», VV.AA., *Gestión y Análisis de Políticas Públicas*, número monográfico coordinado por S. Kushner y E. García Sánchez, 3: 7-15.
- LARRÚ, J.M. (2003), «Análisis de los resultados y metaevaluación del sistema de evaluaciones de Europeaid», documento para la discusión 03-01. Madrid: CECOD.
- LASSWELL, H.D. (1951-2000), «La orientación hacia las políticas», en L.F. Aguilar Villanueva, estudio introductorio y edición, *El estudio de las políticas públicas*. México, D. F.: Miguel Ángel Porrúa.
- LEAVITT, H. *et al.* (1974), *Organizations of the Future*. Nueva York: Praeger Publishers.
- LEVITON, L.C. y HUGHES, E.F.X. (1981), «Research on the Utilization of Evaluations», *Evaluation Review*, 5: 525-548.
- LINDBLOM, CH.E. (1959), «The Science of Muddling Through», *Public Administration Review*, 19: 79-88.
- (1991), *El proceso de elaboración de políticas públicas*. Madrid: Instituto Nacional de Administración Pública.
- MACDONALD, B. (1977), «A political classification of evaluation studies», en D. Hamilton, ed., *Beyond the Numbers Game*. Londres: MacMillan.
- MACDONALD, B. y KUSHNER, S. (2005), «Democratic Evaluation», en S. Mathison, ed., *Encyclopedia of Evaluation*. Thousand Oaks, CA: Sage.
- MADAUS, G.F., SCRIVEN, M. y STUFFLEBEAM, D.L., eds. (1983), *Evaluation Models. Viewpoints on Educational and Human Services Evaluation*. Boston: Kluwer Nijhoff.
- MAJONE, G. (1997), *Evidencia, argumentación y persuasión en la formulación de políticas*. México: Fondo de Cultura Económica.
- MARAVALL, J.M. (1984), *La reforma de la enseñanza*. Barcelona: Laia.
- MARCH, J. (1974), «For a Technology of Foolishness», en H. Leavitt *et al.*, *Organizations of the Future*. Nueva York: Praeger Publishers.
- MARCHESI, A. (1986), «La integración escolar: un proyecto de escuela integradora», *Quinesia. Revista de Educación Especial*, 6: 5-17.
- MARCHESI, A. y FIERRO, A. (1990), «Los objetivos de la reforma del sistema educativo» en *Razón y Fe*, 220 (1091-1092): 153-160.
- MARCHESI, A. y MARTÍN, E. (1998), *Calidad de la enseñanza en tiempos de cambio*. Madrid: Alianza.
- MARCHESI, A., COLL, C. y PALACIOS, J. (1990), «Del lenguaje del trastorno a las necesidades educativas especiales», en A. Marchesi, C. Coll y J. Palacios, comps., *Desarrollo psicológico y educación III. Necesidades educativas especiales y aprendizaje escolar*. Madrid: Alianza.
- MARCHESI, A. COLL, C. y PALACIOS, J. comps (1990), *Desarrollo psicológico y educación III. Necesidades educativas especiales y aprendizaje escolar*. Madrid: Alianza.
- MARSH, D. y SMITH, M. (2000), «Understanding Policy Networks: towards a Dialectical Approach», *Political Studies*, 48, 1: 4-21.

- MARTINÓN, R. (2006), *Las políticas de drogas en España (1982-1996) a través del marco de las coaliciones promotoras. Un estudio del cambio en las políticas públicas*. Director: Francisco J. Vanaclocha Bellver. Tesis doctoral, Universidad Carlos III de Madrid: Departamento de Ciencia Política y Sociología.
- MATHISON, S., ed. (2005), *Encyclopedia of Evaluation*. Thousand Oaks, CA: Sage.
- MAXWELL, J.A. *et al.* (1985), «Combining Ethnographic and Experimental Methods in Educational Research: A Case Study», en M.D. Fetterman y M.A. Pitman, eds., *Beyond the Status Quo: Theory, Politics and Practice in Ethnographic Evaluation*. Washington, DC: Cato Institute.
- MENDELOW, A.L. (1987), «Stakeholder Analysis for Strategic Planning and Implementation», en W.R. King y D.L. Cleland, eds., *Strategic Planning and Management Handbook*. Nueva York: Van Nostrand Reinhold.
- MENOR, M. y MOREIRO, J., coords. (2010), *La reforma experimental de las enseñanzas medias (1983-1987). Crónica de una ilusión*. Las Rozas: Wolters Kluwer.
- MINISTERIO DE EDUCACIÓN Y CIENCIA (1983), *Hacia la reforma*. Madrid.
- (1987), *Proyecto para la reforma de la enseñanza. Propuesta para debate*. Madrid.
- (1987a), *Papeles para el debate*. Madrid.
- (1988), *Estadística de la enseñanza en España: niveles no universitarios*. Madrid.
- (1989), *Libro Blanco para la Reforma del Sistema Educativo*. Madrid.
- (1989a), *Estadística de la enseñanza en España: niveles de preescolar, general básica y enseñanzas medias, 1986/87*. Madrid.
- (1989b), *Plan Marco de Formación Permanente del Profesorado*. Madrid.
- (1990), *Memoria económica del Proyecto de Ley de Ordenación General del Sistema Educativo*. Madrid.
- (1991), *Estadística de la enseñanza en España: niveles de preescolar, general básica y enseñanzas medias, 1987/88*. Madrid.
- (1992), *Estadística de la enseñanza en España: niveles de preescolar, general básica y enseñanzas medias, 1988/1989*. Madrid.
- (1993), *Estadística de la enseñanza en España: niveles de preescolar, general básica y enseñanzas medias, 1989/90*. Madrid.
- (1994), *Estadística de la enseñanza en España: niveles de preescolar, general básica y enseñanzas medias, 1990/1991*. Madrid.
- (1994a), *Estadística de la enseñanza en España: niveles de preescolar, infantil, general básica y enseñanzas medias, 1991/92*. Madrid.
- MINISTERIO DEL INTERIOR (2010), *Resultados electorales de las elecciones generales de 1 de marzo de 1979*. En línea: <<http://www.infoelectoral.mir.es/min/busqueda-AvanzadaAction.html?vuelta=1&codTipoEleccion=2&codPeriodo=197903&codEstado=99&codComunidad=0&codProvincia=0&codMunicipio=0&codDistrito=0&codSeccion=0&codMesa=0>> (consulta: 15 enero 2010).

- (2010b), *Resultados electorales de las elecciones generales de 28 de octubre de 1982*. En línea: <<http://www.infoelectoral.mir.es/min/busquedaAvanzadaAction.html?vuelta=1&codTipoEleccion=2&codPeriodo=198210&codEstado=99&codComunidad=0&codProvincia=0&codMunicipio=0&codDistrito=0&codSeccion=0&codMesa=0>> (consulta: 15 enero 2010).
- (2010c), *Resultados electorales de las elecciones generales de 22 de junio de 1986*. En línea: <<http://www.infoelectoral.mir.es/min/busquedaAvanzadaAction.html?vuelta=1&codTipoEleccion=2&codPeriodo=198606&codEstado=99&codComunidad=0&codProvincia=0&codMunicipio=0&codDistrito=0&codSeccion=0&codMesa=0>> (consulta: 15 enero 2010).
- (2010d), *Resultados electorales de las elecciones generales de 29 de octubre de 1989*. En línea: <<http://www.infoelectoral.mir.es/min/busquedaAvanzadaAction.html?vuelta=1&codTipoEleccion=2&codPeriodo=198910&codEstado=99&codComunidad=0&codProvincia=0&codMunicipio=0&codDistrito=0&codSeccion=0&codMesa=0>> (consulta: 15 enero 2010).
- MINISTERIO PARA LAS ADMINISTRACIONES PÚBLICAS (1992), *Plan de Modernización de la Administración del Estado*. Madrid.
- MORGAN, G. (1993), «Intereses, conflictos y poder: las organizaciones como sistemas políticos», en *Lecturas de teoría de la organización*, vol. II (selección de textos de C. Ramíó y X. Ballart). Madrid: Ministerio para las Administraciones Públicas / Boletín Oficial del Estado.
- MURCIANO, J. (2004), *Metaevaluación de proyectos de empleo. Iniciativa Comunitaria Equal en Andalucía 2001-2004*. Directora: M. Bustelo Ruesta. Memoria de Grado. Máster de Evaluación de programas y políticas públicas. Universidad Complutense de Madrid, mimeo.
- (2007), «Evaluar la evaluación: una herramienta para la mejora. Aplicación al caso de la Iniciativa Equal en Andalucía», en M.J. Terol Becerra, dir., *Nuevas Políticas Públicas, Anuario Multidisciplinar para la modernización de las Administraciones Públicas*, 3: 173-199.
- NEVO, D. (2006), «Evaluation in education», en I.F. Shaw, J.C. Greene y M.M. Mark, eds., *Handbook of Evaluation*. Londres: Sage.
- NILSSON, N. y HOGBEN, D. (1983), «Metaevaluation», *New Directions for Program Evaluation*, 19: 83-97.
- OCDE (1986), *Examen de la política educativa española por la OCDE*. Madrid: Centro de Publicaciones del Ministerio de Educación y Ciencia.
- (1991), *Proyecto Atenea. Informe de evaluación*. Madrid: Ministerio de Educación y Ciencia.
- (2010), *Estándares de calidad para la evaluación del desarrollo*. En línea: <<http://www.oecd.org/dataoecd/7/40/46297655.pdf>> (consulta: 15 febrero 2011).
- ORTÍ, A. (1996), «La apertura y el enfoque cualitativo o estructural: la entrevista abierta semidirectiva y la discusión de grupo», en M. García Ferrando, F. Alvira y J. Ibáñez, comps., *El análisis de la realidad social*. Madrid: Alianza.

- OSUNA LLANEZA, J.L. *et al.* (2011), *Metaevaluación de las políticas evaluadas por EVALÚA DF en 2009. Informe final*. México D.F.: Evalúa DF. En línea: <http://www.evalua.df.gob.mx/transparencia/2011/inf_finmet_09.pdf> (consulta: 8 julio 2011).
- PALUMBO, D.J. (1987), *The Politics of Program Evaluation*. Beverly Hills, CA: Sage.
- PARLETT, M. y HAMILTON, D. (1977), «Evaluation as Illumination: A new approach to the study of innovatory programmes», en D. Hamilton, ed, *Beyond the Numbers Game*. Londres: McMillan.
- PATTON, M.Q. (1982), *Practical Evaluation*. Beverly Hills, CA: Sage.
- (1986), *Utilization-Focused Evaluation*. Thousand Oaks, CA: Sage (2ª edición).
- (1997), *Utilization-Focused Evaluation*. Thousand Oaks, CA: Sage (3ª edición).
- PERROW, C.N. (1990), *Sociología de las organizaciones*. Madrid: McGraw-Hill.
- PFEFFER, J. (1981), *Power in Organizations*. Cambridge: Ballinger (traducción al castellano en 1993, *El poder en las organizaciones: política e influencia en una empresa*. Madrid: McGraw-Hill).
- PITKIN, H. (1985), *El concepto de representación*. Madrid: Centro de Estudios Constitucionales.
- PUELLES, M. (1987), «El bachillerato como problema: antecedentes y situación actual», *Cuenta y Razón*, 27: 70-78.
- (1991), *Educación e ideología en la España contemporánea*. Barcelona: Labor.
- RAMIÓ, C. y BALLART, X. (1993), *Lecturas de teoría de la organización*, vol. II, selección de textos. Madrid: Ministerio para las Administraciones Públicas / Boletín Oficial del Estado.
- REAL ACADEMIA ESPAÑOLA DE LA LENGUA (2001), *Diccionario de la Lengua Española*, 22ª edición, Madrid. En línea: <http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=m%E9rito> (consulta: 30 enero de 2011).
- REICHARDT, C.S. y COOK, T.D., eds. (1979), *Qualitative and Quantitative Methods in Evaluation Research*. Beverly Hills, CA: Sage.
- RIST, R.C. y PICCIOTTO, R., eds. (1995), *Proceedings of the 1994 World Bank Conference*.
- ROSSI, P.H. y FREEMAN, H.E. (1989), *Evaluación. Un enfoque sistemático para programas sociales*. México: Trillas.
- RYAN, K. y DE STEFANO, L. (2001), «Dialogue as a Democratizing Evaluation Method», *Evaluation*, 7, 2: 188-203.
- SABATIER, P. A. y HUNTER, S. (1989), «The Incorporation of Casual Perceptions into Models of Elite Belief Systems», *The Western Political Quarterly*, 42, 3: 229-261.
- SÁEZ BREZMES, M.J. (1995), «Caracterización y problemática de los informes de evaluación», en M.J. Sáez Brezmes, ed., *Conceptualizando la evaluación en España*. Madrid: Servicio de Publicaciones de la Universidad de Alcalá.

- SÁEZ BREZMES, M.J., ed. (1995), *Conceptualizando la evaluación en España*. Madrid: Servicio de Publicaciones de la Universidad de Alcalá.
- SANTOS GUERRA, M.A. (1999), «Meta-evaluación: rigor, mejora, ética y aprendizaje», en B.J. Jiménez Jiménez, ed., *Evaluación de Programas, Centros y Profesores*. Madrid: Síntesis.
- SANTOS GUERRA, M.A. y MORENO OLIVOS, T. (2004), «¿El momento de la metaevaluación educativa?», *Revista Mexicana de Investigación Educativa*, 9, 23: 913-931.
- SANTOS GUERRA, M.A. et al. (2008), *Nieve y barro. Metaevaluación del Plan de Evaluación de Centros Escolares de Andalucía*. Alcalá de Guadaíra: MAD.
- SCHARPF, F.W. (1997), *Games Real Actors Play. Actor-Centered Institutionalism in Policy Research*. Boulder: Westview Press.
- SHAW, I.F., GREENE, J.C. y MARK, M.M., eds. (2006), *Handbook of Evaluation*. Londres: Sage.
- SCHUMAN, E.A. (1967), *Evaluative Research*. Nueva York: Russell Sage Foundation.
- SCHWANDT, T.A. (2003), «In Search of the Political Morality of Evaluation Practice», *Studies in Educational Policy and Educational Philosophy*, 2, Universidad de Uppsala. En línea: <<http://www.upi.artisan.se/docs/Doc183.pdf>> (consulta: 15 octubre 2008).
- (2005), «Politics of Evaluation», en S. Mathison, ed., *Encyclopedia of Evaluation*. Thousand Oaks, CA: Sage.
- SCHWANDT, T.A. y HALPERN, E.S. (1988), *Linking Auditing and Metaevaluation. Enhancing Quality in Applied Research*. Newbury Park, CA: Sage.
- SCRIVEN, M. (1969), «An Introduction to Metaevaluation», *Educational Product Report*, 2: 36-38.
- (1972), «Pros and Cons About Goal-Free Evaluation», *Evaluation Comment: The Journal of Educational Evaluation*, 3, 4: 1-7.
- (1991), *Evaluation Thesaurus*. Newbury Park, CA: Sage.
- (1993), *Hard-Won Lessons in Program Evaluation*. San Francisco: Jossey-Bass.
- (2007), *Key Evaluation Checklist*, The Evaluation Center, Western Michigan University. En línea: <http://www.wmich.edu/evalctr/archive_checklists/kec_feb07.pdf> (consulta: 20 diciembre 2010).
- SEGOVIA PÉREZ, J. (2010), «La reforma de las enseñanzas medias: el marco político», en M. Menor y J. Moreiro, coords., *La reforma experimental de las enseñanzas medias (1983-1987). Crónica de una ilusión*. Las Rozas: Wolters Kluwer.
- SIBEON, R. (1999), «Agency, Structure and Social Change as Cross-disciplinary Concepts», *Politics*, 19, 3: 139-144.
- SIMON, H.A. (1957), *Administrative Behavior: A Study of Decision-Making Process in Administrative Organization*. New York: Free Press.
- SOLANA, J. (1992), «La educación en España en el decenio 1982-1992», en J.F. Tezanos y A. Guerra, eds., *La década del cambio. Diez años de gobierno socialista, 1982-1992*. Madrid: Sistema.

- STAKE, R.E. (1986), *Quieting Reform. Social Science and Social Action in an Urban Youth Program*. Urbana / Chicago: University of Illinois Press.
- (2006), *Evaluación comprensiva y evaluación basada en estándares*. Barcelona: Editorial Graó (original en inglés en 2003, *Standard-Based and Responsive Evaluation*. Thousand Oaks, CA: Sage).
- STENMARK, D. (2000-2001), «Leveraging Tacit Organizational Knowledge», *Journal of Management Information Systems*, 17, 3: 9-24.
- STUFFLEBEAM, D.L. (1971), «The relevance of the CIPP evaluation model for educational accountability», ponencia presentada en el *Annual Meeting of the American Association of School Administrators*, Atlantic City, N.J., 24 de febrero.
- (1974), «Meta-Evaluation», *The Evaluation Center's Occasional Paper Series*, 3, reeditado en 2011 en *Journal of Multidisciplinary Evaluation*, 7, 15: 99-158.
- (1981), «Metaevaluation: Concepts, Standards and Uses», en R. Berk, ed., *Educational Evaluation: The State of the Art*. Baltimore: The Johns Hopkins University Press.
- (1999), *Program Evaluations Metaevaluation Checklist*. The Evaluation Center, Western Michigan University. En línea: <http://www.wmich.edu/evalctr/archive_checklists/program_metaeval_10point.pdf> (consulta: 19 febrero 2011).
- (2001), «The Metaevaluation Imperative», *American Journal of Evaluation*, 22, 2: 183-209.
- STUFFLEBEAM, D.L., et al. (1971), *Educational Evaluation and Decision Making*. Itasca, Illinois: F. E. Peacock Publishers Inc.
- STUFFLEBEAM, D.L. y GOODYEAR, L., MARQUART, J. y JOHNSON, E. (2005), *Guiding principles checklist for Evaluating Evaluations* (based on the 2004 *Guiding Principles for Evaluators*), The Evaluation Center, Western Michigan University. En línea: <http://www.wmich.edu/evalctr/archive_checklists/guidingprinciples2005.pdf> (consulta: 19 febrero 2011).
- STUFFLEBEAM, D.L. y SHINKFIELD, A.J. (1993), *Evaluación sistemática. Guía teórica y práctica*. Madrid: Paidós / Ministerio de Educación y Ciencia.
- (2007), *Evaluation Theory, Models & Applications*. San Francisco, CA: Jossey-Bass.
- SUBDIRECCIÓN GENERAL DE INVESTIGACIÓN EDUCATIVA, *Investigaciones educativas de la red INCIE-ICE, 1978-1982*. Madrid: Ministerio de Educación y Ciencia.
- SUBIRATS, J. (1990-1991), «La administración pública como problema. El análisis de políticas públicas como propuesta», en VV.AA., *Políticas públicas y organización administrativa*, número monográfico de *Documentación Administrativa*, 224-225: 15-57.
- SWISS EVALUATION SOCIETY (2000), *Evaluation Standards*. En línea: <http://www.seval.ch/en/documents/SEVAL_Standards_2000_en.pdf> (consulta: 27 diciembre 2010).
- TEROL BECERRA, M.J. dir., *Nuevas Políticas Públicas, Anuario Multidisciplinar para la modernización de las Administraciones Públicas*, 3: 173-199.

- TEZANOS, J.F. y GUERRA, A., eds. (1992), *La década del cambio. Diez años de gobierno socialista, 1982-1992*. Madrid: Sistema.
- TIANA FERRER, A. (2009), *Por qué hicimos la LOE*. Las Rozas: Wolters Kluwer.
- TYLER, R.W. (1950), *Basic Principles of Curriculum and Instruction*. Chicago: University of Chicago Press.
- UUSIKYLÄ, P. y VIRTANEN, P. (2000), «Meta-Evaluation as a Tool for Learning. A Case Study of the European Structural Fund Evaluations in Finland», *Evaluation*, 6, 1: 50-65.
- VEDUNG, E. (1997), *Evaluación de Políticas Públicas y Programas*. Madrid: Ministerio de Trabajo y Asuntos Sociales.
- VÉLAZ DE MEDRANO, C. et al. (1995), *Evaluación de programas y de centros educativos. Diez años de investigación*. Madrid: Centro de Investigación y Documentación Educativa.
- VON BEYME, K. (1994), *Teoría política del siglo XX. De la modernidad a la posmodernidad*. Madrid: Alianza Editorial.
- VV.AA. (1991), *Políticas públicas y organización administrativa*, número monográfico de la revista *Documentación Administrativa*, 224-225.
- VV.AA. (1993), *Evaluar las evaluaciones*, número monográfico de *Cuadernos de Pedagogía*, 219.
- VV.AA. (2010), *Gestión y Análisis de Políticas Públicas*, número monográfico coordinado por S. Kushner y E. García Sánchez, 3.
- WARNOCK, M. (1981), *Meeting Special Educational Needs*. Londres: Her Britannic Majesty's Stationary Office.
- WEISS, C.H. (1987), «Where Politics and Evaluation Research Meet», en D.J. Palumbo, *The Politics of Program Evaluation*. Beverly Hills, CA: Sage.
- (1990), «Evaluations for Decisions: Is Anybody There? Does Anybody Care?», en M.C. Alkin, *Debates on Evaluation*. Newbury Park, CA: Sage.
- (1998), *Evaluation*. New Jersey: Prentice Hall.
- WHEELER, P., HAERTEL, G.D. y SCRIVEN, M. (1992), *Teacher Evaluation Glossary. Kalamazoo*, MI: CREATE Project, The Evaluation Center, Western Michigan University.
- WHOLEY, J.S. y WHITE, B.F. (1973), «Evaluation's impact on Title I elementary and secondary education program management», *Evaluation*, 1: 73-76.
- YARBROUGH, D.B. et al. (2011), *The program evaluation standards: A guide for evaluators and evaluation users*. Thousand Oaks, CA: Sage
- YUKL, G. (1994), *Leadership in Organizations*. Englewood Cliffs: Prentice-Hall.
- ZUFIAURRE, B. (1994), *Proceso y contradicciones de la reforma educativa, 1982-1994*. Barcelona: Icaria.

ANEXO II

DATOS SOBRE LAS ENTREVISTAS REALIZADAS

Álvarez Fernández, Manuel

Doctor en Filosofía y Ciencias de la Educación por la Universidad Complutense de Madrid. Fue Vicepresidente y Presidente del Fórum Europeo de Administradores de la Educación, fundador y director de la revista *Gestión Educativa*, Jefe del Departamento de Organización y Coordinador de Investigación en el CRIF *Las Acacias* de la Comunidad de Madrid y Director del Máster en Dirección y Gestión de Centros Educativos de la Universidad de Deusto y la Universidad Autónoma de Barcelona.

20 de marzo de 1996

Álvaro Page, Mariano

Doctor en Psicología por la Universidad Nacional de Educación a Distancia. Experto en psicometría, evaluación del rendimiento y metodología de la investigación, trabajó en el Instituto Nacional de Ciencias de la Educación (INCIE) y en el *Educational Testing Service* de la Universidad de Princeton. Fue Jefe del Servicio de Evaluación del CIDE (1986-1992).

14 de marzo de 1996

Angulo Aramburu, Elena

Directora del Gabinete Técnico de la Secretaría General de Educación (1987-1988), Directora del Gabinete del Secretario de Estado de Educación (1988-1992). Ha sido también Directora de Fomento y Promoción Cultural (1996-2003) de la Junta de Andalucía.

27 de noviembre de 1995

Arango Vila-Belda, Joaquín

Doctor en Ciencias Políticas y Sociología por la Universidad Complutense de Madrid. En el Ministerio de Educación y Ciencia estuvo al frente de la Se-

cretaría General Técnica (1982-1985), la Secretaría General de Educación (1985-1986) y la Subsecretaría (1986-1988). Fue, además, Presidente del Centro de Investigaciones Sociológicas (1991-1996) y Vicerrector de la Universidad Internacional Menéndez Pelayo (1989-1991). Es Catedrático de Sociología de la Universidad Complutense de Madrid.

23 de enero de 2001

Candioti López-Pujato, Carmen

Licenciada en Filosofía y Letras por la Universidad Complutense de Madrid. Responsable del Gabinete de Seguimiento y Evaluación del Programa Nacional de Nuevas Tecnologías de la Información y las Comunicaciones (PNTIC), asesora técnico-docente para la formación del profesor en nuevas tecnologías (1986-1987) y Jefa del Gabinete del Proyecto Mentor (1990-1993). En la actualidad es Jefa del Área de Contenidos Web y Medios Audiovisuales del Centro Nacional de Información y Comunicación Educativas del Ministerio de Educación.

12 de junio de 1997

Carabaña Morales, Julio

Doctor en Filosofía y Letras por la Universidad de Valencia. Asesor Ejecutivo en el Gabinete del Ministro de Educación (1982-1983) y director del Centro de Investigación y Documentación Educativa (1983-1986). Actualmente es Catedrático de Sociología de la Educación de la Universidad Complutense de Madrid.

11 de diciembre de 1995

Casanova González, M^a Antonia

Maestra y Licenciada en Pedagogía. Inspectora Central de Educación. Jefa del Gabinete de Planes y Programas de Educación General Básica de la Dirección General de Educación Básica del Ministerio de Educación y Ciencia (1984-1985), Jefa del Servicio de Evaluación de la Dirección General de Educación Básica del Ministerio de Educación y Ciencia (1986), Subdirectora General de Educación Especial y de Atención a la Diversidad en el Ministerio de Educación y Cultura (1996-1999) y Directora General de Promoción Educativa de la Consejería de Educación de la Comunidad de Madrid (1999 a 2007).

13 de febrero de 1995

Del Blanco Díez, Luis

Jefe del Departamento de Formación del Instituto de Ciencias Educativas de Alcalá de Henares. Consejero Técnico de los Servicios Centrales del Programa Nacional de Nuevas Tecnologías de la Información y la Comunicación (PNTIC).

23 de abril de 1998

Escobar Mercado, Modesto

Doctor en Sociología por la Universidad Complutense de Madrid y experto en técnicas de investigación social. Miembro del equipo de evaluación externa de la reforma experimental de enseñanzas medias. En la actualidad es Catedrático de Sociología de la Universidad de Salamanca.

7 de diciembre de 1995

Escudero Muñoz, Juan Manuel

Diplomado en Psicología y Doctor en Ciencias de la Educación. Director de las evaluaciones de los proyectos *Atenea* y *Mercurio* y de la evaluación de la implantación de la LOGSE en la Comunidad de Murcia. Actualmente es Catedrático de Didáctica y Organización Escolar de la Universidad de Murcia.

14 de mayo de 1997

Maravall Herrero, José María

Doctor en Sociología. Ministro de Educación y Ciencia entre 1982 y 1988 y diputado del Grupo Parlamentario Socialista del Congreso desde 1986 a 1989. En la actualidad es Catedrático de Sociología de la Universidad Complutense de Madrid y Miembro del Consejo Asesor del Centro de Estudios Avanzados en Ciencias Sociales del Instituto Juan March de Estudios e Investigaciones.

25 de enero de 2001

Marchesi Ullastres, Álvaro

Doctor en Psicología y Catedrático de Psicología Evolutiva y de la Educación de la Universidad Complutense de Madrid. Fue Subdirector General de Educación Especial (1983-1986), Director General de Renovación Pedagógica (1986-1992) y Secretario de Estado de Educación (1992-1996). En la actualidad es Secretario General de la Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura.

6 de marzo de 1996

Marcos Pérez, Susana

Técnico del CIDE. Miembro del equipo de evaluación externa de la reforma experimental de enseñanzas medias.

8 de febrero de 2000

Martínez Arias, Rosario

Doctora en Psicología. Catedrática del Dpto. de Metodología de las Ciencias del Comportamiento de la UCM. Miembro de los equipos de evaluación de la

reforma experimental de enseñanzas medias y del programa de integración. Fue Presidenta del Comité Científico del Instituto Nacional de Calidad y Evaluación.

10 de diciembre de 1997

Pérez Mas, Pilar

Subdirectora de Formación y Perfeccionamiento del Profesorado (1983-1986) y Secretaria General del Boletín Oficial del Estado (1992-1993).

7 de febrero de 1995

Rivière Gómez, Ángel

Catedrático de Psicología Evolutiva de la Universidad Autónoma de Madrid. Experto en educación especial y autismo, fue director del CIDE entre 1987 y 1989.

3 de enero de 1996

San José Villacorta, Carlos

Catedrático de Instituto. Miembro y Consejero Técnico de los Servicios Centrales del Programa Nacional de Nuevas Tecnologías de la Información y la Comunicación (PNTIC).

29 de noviembre de 1995

Segovia Pérez, José

Catedrático de Instituto. Fue Director General de Enseñanzas Medias (1982-1986) y Director General de Promoción Educativa (1986-1988).

11 de abril de 1997

Tiana Ferrer, Alejandro

Doctor en Ciencias de la Educación. Fue Director del CIDE (1989-1993), Director del INCE (1993-1996), Secretario General de Educación (2004-2008) y Vicerrector de Evaluación e Innovación de la UNED (1999-2003). Es Catedrático de Historia de la Educación de la Universidad Nacional de Educación a Distancia y experto en historia de la educación, educación comparada y evaluación educativa. Actualmente es Director del Centro de Altos Estudios Universitarios de la Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI).

9 de febrero de 1995

Veiguela Martínez, Elena

Licenciada en Ciencias Exactas y Catedrática de Bachillerato. Jefa del Departamento de Formación del ITE de Alcalá de Henares y Consejera Técnica de la

Dirección General de Educación Básica para la introducción de los microordenadores en la educación básica. Directora de los proyectos *Atenea* y *Mercurio* (1985-1989) y del Programa de Nuevas Tecnologías de la Información y la Comunicación (PNTIC). Fue asimismo Coordinadora General en la Subdirección General de Tratamiento de la Información del Ministerio de Educación, Cultura y Deportes.

12 de junio de 1997

El siguiente cuadro refleja el número de entrevistas realizadas para los estudios de caso e incluye tanto las ya mencionadas en este mismo anexo como las realizadas a profesores, monitores, miembros de equipos multiprofesionales y miembros de equipos pedagógicos. Las fechas de estas últimas figuran al final.

Número de entrevistas realizadas para los estudios de caso

	Reforma experimental de enseñanzas medias	Programa de integración	Proyecto <i>Atenea</i>
(i) Profesores	5	5	5
(ii) Profesores de apoyo	–	–	–
(iii) Monitores	–	–	5
(iv) Directores de programa	1	1	1
(v) Personal del programa –no incluye la categorías (iii) ni (iv)	–	–	3
(vi) Directores de la evaluación	1	1	1
(vii) Personal investigador de la evaluación o de la organización encargada de la evaluación –no incluye la categoría (vi)	3	2	–
(viii) Miembros de EE.MM.	–	2	–
Total	10	11	15

Fechas de entrevistas

(i) *Profesores*

Programa de reforma experimental de enseñanzas medias: 14 de abril de 1997, 9 de marzo, 7 de mayo, 12 de mayo y 16 de mayo de 1998.

Profesores de aulas de integración: 27 de enero, 2 de febrero, 6 de febrero, 9 de febrero y 5 de marzo de 1998.

Coordinadores de equipos pedagógicos del proyecto Atenea: noviembre de 1995 - junio de 1996.

(iii) *Monitores:* noviembre de 1995 - junio de 1996.

(v) *Miembros de los EE.MM.:* 17 y 18 de diciembre de 1997.



GOBIERNO
DE ESPAÑA

MINISTERIO
DE POLÍTICA TERRITORIAL
Y ADMINISTRACIÓN PÚBLICA

SECRETARÍA DE ESTADO
PARA LA FUNCIÓN PÚBLICA

INSTITUTO NACIONAL DE
ADMINISTRACIÓN PÚBLICA

INAP

ISBN 978-84-7351-406-4



9 788473 514064

P.V.P.
14,00 €
(IVA incluido)