

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS POLÍTICAS Y SOCIOLOGÍA

Departamento de Ciencia Política y de la Administración II



**LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS
DE IGUALDAD DE GÉNERO DE LOS GOBIERNOS
CENTRAL Y AUTONÓMICOS EN ESPAÑA:
1995 - 1999**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

María Bustelo Ruesta

Bajo la dirección del doctor

Rafael Bañón i Martínez

Madrid, 2001

ISBN: 84-669-2285-7

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS POLÍTICAS Y SOCIOLOGÍA
DEPARTAMENTO DE CIENCIA POLÍTICA Y DE LA
ADMINISTRACIÓN II

LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS DE
IGUALDAD DE GÉNERO DE LOS GOBIERNOS
CENTRAL Y AUTONÓMICOS EN ESPAÑA:
1995-1999

Elaborada por: MARÍA BUSTELO RUESTA
Dirigida por: RAFAEL BAÑÓN i MARTÍNEZ

Madrid, mayo de 2001

*A Carlota Bustelo García del Real,
pionera de las Políticas de Igualdad de Género en España,
con todo mi cariño y admiración.*

*A mis hijos Guillermo y Jaime,
hombres del siglo XXI,
a quienes, estoy segura, les será más fácil comprender
que también a ellos les incumbe e interesa luchar por un mundo
mejor, más justo y diverso,
en donde, de verdad, quepamos todas y todos.*

*Y a Carmelo,
por acompañarme tan acertada y brillantemente
en este gran proyecto que los dos tenemos entre manos.*

AGRADECIMIENTOS

Son muchas las personas que me han prestado ayuda y aliento y sin las que la realización de esta tesis hubiera sido mucho más difícil y costosa.

En primer lugar, quiero agradecer a Rafael Bañón, director de esta tesis, su apoyo y dirección a lo largo de este prolongado y arduo proceso. Nunca podré agradecerle lo suficiente el hecho de haberme descubierto la Ciencia Política y de la Administración y el deslumbramiento intelectual que eso supuso para mí, además de la paciencia demostrada y de no haber dejado de creer en mí en ningún momento.

Para la recopilación de información fue imprescindible la ayuda de las personas responsables política y técnicamente de los procesos de evaluación llevados a cabo en los diferentes organismos de igualdad. Gracias a sus directoras que me brindaron su tiempo para las entrevistas y, de forma muy especial, a los técnicos y técnicas (Mariano Alvaro, Adela Abarrategui, Rosa Sacacia, Chari Pizarro, Carmen Marín, Toñi Gómez, Mónica Gasol, Concha Alvarez, Paloma Rodríguez, Clemente Martín Barroso, Nines Centeno, Itziar Fernández y Ana Alberdi) quienes, de forma desinteresada, me ofrecieron la información necesaria y soportaron pacientemente mis múltiples llamadas telefónicas. Mención especial requiere la ayuda prestada por la profesora Amparo Almarcha de la Universidad de la Coruña, directora de la evaluación externa del plan gallego y Puri Causapié, entonces en la Fundación Mujeres, encargada de la evaluación del plan castellano-manchego.

En las tareas de revisión y edición de este trabajo debo agradecer la ayuda de Carlos Closa, Marta Martínez, Juan Carlos Monedero y Celia Valiente, por su paciente lectura y sus valiosas sugerencias mí.

También debo agradecer la acogida que, desde el curso 1994-95 en el que comencé dando clases como profesora asociada, me ha ofrecido tanto el Departamento de Ciencia Política y de la Administración II como, en general, la Facultad de Ciencias Políticas y Sociología de la Universidad Complutense de Madrid. Gracias a los colegas por los ánimos en los pasillos, a José Ignacio Fernández, secretario administrativo, por la siempre cariñosa y eficaz ayuda logística y, de forma muy especial, a mis alumnos y alumnas a lo largo de estos años, quienes han sido un constante estímulo y el mejor elemento de contraste y maduración de las ideas que se reflejan en esta tesis.

El Centro Superior de Estudios de Gestión, Análisis y Evaluación de la UCM, que me ha acogido como docente y responsable de la Unidad de Evaluación, ha sido también muy importante en este proceso. Debo una mención muy especial a su anterior director, Juan Luis Paniagua, por su constante y desinteresado apoyo y porque me ofreció la posibilidad de contribuir con mi experiencia a participar en un proyecto que aúna acertadamente el mundo académico y el profesional. En Juan Gómez Castañeda, su actual director, he encontrado siempre un trato humano exquisito y un apoyo incondicional a esta tarea doctoral. También han sido importantes la comprensión y los ánimos de la gente que ha trabajado conmigo de forma más cercana, en la Unidad de Evaluación, Marta Martínez y Juan Andrés Ligeró, y en el Observatorio Regional de Políticas Públicas, David Corominas, Manuel Galán y José Manuel Cruz. Agradezco también al resto de los compañeros y compañeras del Centro su aliento y, en especial, a Andrés Benito, su gerente, por cuidar y velar con mimo ese buen ambiente humano y profesional.

Fue también desde el CSEGAE que Juan Luis Paniagua, junto a Jesús Montero, me ofreció la posibilidad de asistir al "Seminario sobre Elaboración de Tesis y Diseño Curricular" en 1998. Este Seminario supuso un impulso y un estímulo muy importante para mí. La impecable, paciente y dedicada dirección del prof. Carlos Closa hizo que el aprovechamiento fuera máximo, además de resultar tremendamente divertidas y estimulantes las discusiones epistemológicas que allí mantuvimos. A partir de entonces, nunca he dejado de contar con sus sabias orientaciones. El intercambio y el contraste con el resto de los integrantes (Raúl Aguilera, Ariel Jerez, M^a Teresa Martínez del Olmo, Laura Nuño, Luis Ramiro y María Velasco), además de las risas de entonces, han servido de estímulo durante todo el proceso.

Gracias, en fin, a todas las personas a las que quiero y que me quieren, por el apoyo y las necesarias ausencias y porque han tenido que soportar la cantinela de la tesis y los agobios que ésta producían a mi persona. Debo agradecer igualmente a mi familia extensa estar ahí y el apoyo logístico prestado en los momentos en los que lo he necesitado, que han sido muchos.

Y gracias, Carmelo, modelo y referencia académica y padre exquisito en quien he podido delegar en estos años una buena parte de mis tareas de madre, con el fin de dedicar más tiempo y energías a esta empresa. Y, por supuesto, gracias a Guillermo y a Jaime, por existir y estar ahí, por ayudarme a saber cuáles son las cosas importantes de la vida y por ser siempre, y sobre todo, una fuente inagotable de satisfacción. Ellos me han dado la fuerza para llegar hasta el final.

LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO DE LOS GOBIERNOS CENTRAL Y AUTONÓMICOS EN ESPAÑA: 1995-1999

SUMARIO

INTRODUCCIÓN

PARTE I: LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS

CAPÍTULO 1: La evaluación en el marco del análisis de políticas públicas

CAPÍTULO 2: La evaluación de políticas públicas desde una perspectiva histórica

PARTE II: LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO EN ESPAÑA

CAPÍTULO 3: Las políticas públicas de igualdad de género en España en los niveles central y autonómico

PARTE III: LA PRÁCTICA DE LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO EN ESPAÑA. LOS PROCESOS DE EVALUACIÓN DE LOS PLANES CENTRALES Y AUTONÓMICOS

CAPÍTULO 4: La evaluación de las políticas públicas de igualdad de género en España: marco para el análisis y la clasificación

CAPÍTULO 5: El planteamiento general de los procesos de evaluación analizados: contexto, propósitos, estructuras y recursos

CAPÍTULO 6: Los elementos clave de la evaluación: agentes, momentos, criterios y metodología.

CAPÍTULO 7: La utilización de las evaluaciones: pertinencia, comunicación e impacto

PARTE IV: LA VALORACIÓN DE LA EVALUACIÓN. EL RETO DE LA EVALUACIÓN DE LAS POLÍTICAS DE IGUALDAD DE GÉNERO.

CAPÍTULO 8: El discurso sobre la función de evaluación en torno a las políticas de igualdad de género.

CAPÍTULO 9: Los factores que facilitan y dificultan la evaluación de las políticas de igualdad de género. Hacia una evaluación innovadora de dichas políticas.

PARTE V: CONCLUSIONES

BIBLIOGRAFÍA

ANEXOS

LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO DE LOS GOBIERNOS CENTRAL Y AUTONÓMICOS EN ESPAÑA: 1995-1999

ÍNDICE

INTRODUCCIÓN

Objeto.....	1
Marco teórico.....	3
Objetivos e hipótesis de la investigación.....	8
Metodología.....	9
Plan de exposición.....	12

PARTE I: LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS

CAPÍTULO 1. LA EVALUACIÓN EN EL MARCO DEL ANÁLISIS DE POLÍTICAS PÚBLICAS

Introducción.....	19
1. El análisis de políticas públicas: el papel de la evaluación en este marco de referencia. 21	
1.1. Concepto y características del análisis de políticas públicas.....	21
1.2. El componente descriptivo y prescriptivo del análisis de las políticas públicas.....	22
1.3. La necesidad de la evaluación en el ciclo de vida de la política.....	23
1.4. El carácter retrospectivo y prospectivo del análisis de políticas públicas.....	25
1.5. Una visión integrada del análisis de políticas públicas: el papel de la evaluación	26
2. La evaluación de políticas públicas.....	29
2.1. El concepto y el sentido de la evaluación.....	29
2.2. Las funciones de la evaluación.....	32
2.3. Propósitos y objetivos de la evaluación de las políticas públicas.....	35
2.4. Diferencias y similitudes entre la evaluación de programas y de políticas públicas .	38
3. Lo específico de la evaluación.....	41
3.1. Las diferencias con otras prácticas de análisis: evaluación vs. Investigación.....	41
3.2. El carácter político de la evaluación.....	53
3.3. El carácter práctico y aplicado a la acción: el elemento valorativo y el establecimiento de recomendaciones.....	57
3.4. El carácter práctico y aplicado a la acción: La preocupación por la utilización y el uso de la evaluación.....	65

CAPÍTULO 2. LA EVALUACIÓN DE POLÍTICAS PÚBLICAS DESDE UNA PERSPECTIVA HISTÓRICA

Introducción.....	71
1. La historia de la evaluación y los diferentes modelos o enfoques propuestos.....	73
1.1. Breve historia de la evaluación.....	73

1.2. Factores que influyen en la evolución de la evaluación.....	76
1.3. Las diferentes clasificaciones de modelos o enfoques	79
A. Enfoque de evaluación por objetivos.....	81
B. Enfoques experimentales de evaluación	82
C. Enfoques orientados a la toma de decisiones.....	83
D. Enfoque centrado en el cliente.....	89
E. Enfoque centrado en el consumidor	91
F. Enfoque orientado a la utilización.....	92
G. Enfoques participativos	96
G.a. La Evaluación de Cuarta Generación.....	99
G.b. La Evaluación para el Empoderamiento	105
G.c. La Evaluación Pluralista.....	107
2. Tipos de evaluación	109
2.1. Según el papel que cumple la evaluación.....	109
2.2. Según el contenido o aspecto de la política que se evalúa	111
2.3. Según el agente que evalúa.....	113
2.4. Según el momento en el que se evalúa.....	114
2.5. Otros tipos de evaluación	116
2.6. Otras clasificaciones.....	118
3. Los actores de la evaluación	121
4. El papel del evaluador.....	124
5. Las cuestiones éticas en evaluación.....	127
6. Hacia un enfoque integrador de evaluación de políticas públicas.....	130

PARTE II: LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO EN ESPAÑA

CAPÍTULO 3: LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO EN ESPAÑA EN LOS NIVELES CENTRAL Y AUTONÓMICO

Introducción	133
1. Concepto e historia de las políticas de igualdad de género	135
1.1. Las difíciles relaciones feminismo-Estado.....	135
1.2. El feminismo institucional.....	140
2. Las políticas de igualdad de género en España desde una perspectiva histórica.....	146
2.1. Un crecimiento rápido	146
2.2. La evolución en la definición del problema	151
2.3. La entrada en la agenda política.....	157
2.4. El marco institucional de las políticas de igualdad	162
2.5. Una revisión de los actores que contribuyen a la formulación de las políticas.....	172
3. Los planes de igualdad de género	176
3.1. El significado de los planes de igualdad	176
3.2. La evolución de los planes de igualdad en España	180
3.3. La estructura y los contenidos de los planes de igualdad.....	185

3.4. Fortalezas y debilidades de los planes.....	186
--	-----

PARTE III: LA PRÁCTICA DE LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO EN ESPAÑA. LOS PROCESOS DE EVALUACIÓN DE LOS PLANES CENTRALES Y AUTONÓMICOS

CAPÍTULO 4: LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO EN ESPAÑA: MARCO PARA EL ANÁLISIS Y LA CLASIFICACIÓN

Introducción	191
1. Fundamentación.....	193
1.1. La metaevaluación.....	193
El concepto de metaevaluación y su diferenciación con el concepto de síntesis de evaluaciones	193
El significado y las funciones de la metaevaluación.....	196
Tipos de metaevaluación	198
Criterios para la metaevaluación	202
La metaevaluación de evaluaciones de políticas públicas de igualdad de género que se presenta en este trabajo.....	211
1.2. El análisis de los procesos de evaluación de las políticas públicas de igualdad de género en España.....	213
2. Un panorama general: experiencias de evaluación en políticas de igualdad en los niveles estatal y autonómico	216
2.1. Las primeras experiencias de evaluación	217
2.2. La segunda generación de evaluaciones.....	228
2.3. La tercera generación de evaluaciones	230
3. El marco para la metaevaluación: los criterios para el análisis valorativo de las once experiencias de evaluación.....	239
3.1. El planteamiento de la evaluación y las estrategias empleadas.....	240
El contexto evaluativo	240
La claridad en los propósitos y funciones de las evaluaciones	240
Las estructuras institucionales de la evaluación.....	241
Tipos de evaluaciones realizadas	241
Recursos empleados en las evaluaciones	241
3.2. Elementos clave de los procesos evaluativos	242
Agentes y participantes en el proceso de evaluación	242
Momentos de los procesos de evaluación	243
Criterios e indicadores de evaluación: fuentes y proceso de establecimiento.....	243
Procedimientos y metodología empleada	244
3.3. La utilización de los procesos evaluativos	244
Pertinencia y utilidad de la información producida	245
Comunicación de los resultados y destinatarios de los mismos	245
El impacto de los resultados y los procesos de evaluación	246

CAPÍTULO 5: EL PLANTEAMIENTO GENERAL DE LOS PROCESOS DE EVALUACIÓN ANALIZADOS: CONTEXTO, PROPÓSITOS, ESTRUCTURAS Y RECURSOS

Introducción	249
1. Tipos de evaluación realizados	251
1.1. Un primer análisis de las once experiencias según los tipos de evaluación.....	251
1.2. Los criterios para la clasificación.....	258
2. El planteamiento de las evaluaciones y las estrategias empleadas	264
2.1. Los contextos evaluativos.....	264
Panorama general.....	264
El contexto evaluativo de las evaluaciones externas	269
El contexto evaluativo de las evaluaciones mixtas.....	270
El contexto evaluativo de las evaluaciones internas	271
2.2. La claridad en los propósitos y las funciones de las evaluaciones	272
Panorama general.....	272
Los propósitos y objetivos de las evaluaciones externas.....	273
Los propósitos y objetivos de las evaluaciones mixtas	279
Los propósitos y objetivos de las evaluaciones internas	281
2.3. Las estructuras institucionales de la evaluación.....	282
Panorama general.....	282
Las estructuras institucionales para las evaluaciones externas.....	289
Las estructuras institucionales para las evaluaciones mixtas	291
Las estructuras institucionales para las evaluaciones internas	293
2.4. Recursos empleados en las evaluaciones	294
Panorama general.....	294
Los recursos para las evaluaciones externas.....	296
Los recursos empleados en las evaluaciones mixtas	298
Los recursos empleados en las evaluaciones internas	299

CAPÍTULO 6: LOS ELEMENTOS CLAVE DE LA EVALUACIÓN: AGENTES, MOMENTOS, CRITERIOS Y METODOLOGÍA.

Introducción	301
1. Agentes y participantes en los procesos de evaluación	303
1.1. Agentes y participantes en los procesos de evaluación externa.....	304
1.2. Agentes y participantes en los procesos de evaluación mixta.....	307
1.3. Agentes y participantes en los procesos de evaluación interna.....	308
2. Momentos de los procesos de evaluación.....	310
2.1. Momentos de los procesos de evaluación externa.....	310
2.2. Momentos de los procesos de evaluación mixta	314
2.3. Momentos de los procesos de evaluación interna	315
3. Criterios e indicadores de evaluación: fuentes y proceso de establecimiento.....	317
3.1. Criterios e indicadores en las evaluaciones externas.....	318
3.2. Criterios e indicadores en las evaluaciones mixtas	325
3.3. Criterios e indicadores en las evaluaciones internas	331
4. Procedimientos y metodología empleada	333
4.1. Procedimientos y metodología de las evaluaciones externas.....	334

4.2. Procedimientos y metodología de las evaluaciones mixtas.....	340
4.3. Procedimientos y metodología en las evaluaciones internas.....	346

CAPÍTULO 7: LA UTILIZACIÓN DE LAS EVALUACIONES: PERTINENCIA, COMUNICACIÓN E IMPACTO

Introducción	347
1. Pertinencia y utilidad de la información producida en las evaluaciones externas	349
1.1. Pertinencia y utilidad de la información producida en las evaluaciones externas ...	349
1.2. Pertinencia y utilidad de la información producida en las evaluaciones mixtas	354
1.3. Pertinencia y utilidad de la información producida en las evaluaciones internas	356
2. Comunicación de los resultados y destinatarios de los mismos	358
2.1. Comunicación de los resultados en las evaluaciones externas.....	358
2.2. Comunicación de los resultados en las evaluaciones mixtas	367
2.3. Comunicación de los resultados en las evaluaciones internas	374
3. El impacto de los resultados y los procesos de evaluación	376
3.1. El impacto de los resultados y los procesos de evaluación externa	376
3.2. El impacto de los resultados y los procesos de evaluación mixta	381
3.3. El impacto de los resultados y los procesos de evaluación interna	385

PARTE IV: LA VALORACIÓN DE LA EVALUACIÓN. EL RETO DE LA EVALUACIÓN DE LAS POLÍTICAS DE IGUALDAD DE GÉNERO

CAPÍTULO 8: EL DISCURSO SOBRE LA FUNCIÓN DE EVALUACIÓN EN TORNO A LAS POLÍTICAS DE IGUALDAD DE GÉNERO

Introducción	387
1. El discurso sobre la evaluación y sobre las políticas de igualdad de género de las responsables políticas	389
2. La evaluación: ideas asociadas y actitudes ante la misma	394
2.1. Entre las expectativas mágicas sobre la evaluación y una visión más realista.....	394
2.2. La inevitable necesidad de la evaluación vs. la conveniencia de la misma	396
2.3. La evaluación como algo propio de los organismos de igualdad.....	396
2.4. Una evaluación más integral vs. una evaluación asociada exclusivamente a los resultados.....	397
2.5. La evaluación como algo técnico	398
2.6. La evaluación como medición cuantitativa y relacionada con indicadores	398
2.7. La evaluación o bien como seguimiento o bien como evaluación externa	399
3. ¿Para qué sirve evaluar? La utilidad de la evaluación	400
4. ¿Qué se debe evaluar y cómo?	404
5. Dificultades percibidas para la evaluación de los planes de igualdad	406
5.1. Dificultades relacionadas con la falta de cultura de evaluación.....	406
5.2. Dificultades relacionadas con la complejidad de los planes	407

5.3. Dificultades relacionadas con la estrategia de la transversalidad de las políticas de igualdad de género	408
---	-----

CAPÍTULO 9: LOS FACTORES QUE FACILITAN Y DIFICULTAN LA EVALUACIÓN DE LAS POLÍTICAS DE IGUALDAD DE GÉNERO. HACIA UNA EVALUACIÓN INNOVADORA DE DICHAS POLÍTICAS.

Introducción	417
1. Factores previos que facilitan o dificultan la evaluación	420
1.1. Los factores que facilitan la evaluación	420
La concepción y estructuración en planes de igualdad.....	420
La institucionalización de compromisos respecto a los planes y a su evaluación.....	421
Aval y grado de acuerdo internacional	422
La fortaleza y estabilidad institucional de los organismos de igualdad	423
El reconocimiento simbólico de la evaluación en los últimos años	425
1.2. Los factores que dificultan la evaluación	426
La abstracción o generalidad de los planes.....	426
Su carácter de horizontalidad e integralidad.....	428
Su carácter de sensibilización y persuasión.....	429
Falta de cultura y experiencia previa en evaluación.....	430
Falta de equipos preparados.....	431
No pensar la evaluación a tiempo debido a la falta de normalización de su función	432
2. Factores inherentes al propio proceso de evaluación	434
2.1. Falta de claridad en el propósito de la evaluación:	
Evaluación vs. investigación de políticas de igualdad	434
2.2. Falta de una visión global de la actuación de los poderes públicos:	
La evaluación de las políticas vs. la evaluación de los planes	437
2.3. Falta de reconocimiento del carácter político de la evaluación.....	438
2.4. Percepción de la evaluación como una función secundaria:	
El papel crucial de los organismos de igualdad	439
2.5. Visión parcial de la evaluación:	
Una mala gestión de los diferentes tipos de evaluación.....	442
2.6. La optimización de agentes externos en la evaluación:	
Las evaluaciones mixtas vs. las evaluaciones externas.....	445
2.7. La falta de claridad en el QUÉ evaluar:	
La dictadura de la metodología y las técnicas en los procesos de evaluación	447
2.8. Problemas en la recopilación y los flujos de información:	
Sistemas, cultura de recopilación, estructuras de coordinación	451
2.9. La importancia de los tiempos y el calendario de los procesos:	
La temporalidad como elemento clave de la evaluación.....	455
2.10. El claro déficit de “elaboración práctica”.....	456
2.11. Los deficientes procesos de comunicación y difusión de la evaluación:	
La accesibilidad de los informes	458
2.12. La escasez de recursos dedicados.....	460
3. Factores relacionados con el impacto de la evaluación	463
3.1. La relativa utilidad y pertinencia de los procesos	463
3.2. La imposibilidad de una síntesis de evaluaciones	467

PARTE V: CONCLUSIONES

Sobre la evaluación en el marco del análisis de políticas públicas	471
Sobre el concepto de evaluación de políticas públicas.....	472
Sobre la historia de la evaluación y los diferentes enfoques o modelos propuestos	475
Sobre los tipos de evaluación, los actores y agentes y sus implicaciones éticas.....	476
Sobre las políticas de igualdad de género en España	478
Sobre la metaevaluación como marco para el análisis y la valoración de las evaluaciones de los planes de igualdad de género	480
Sobre las experiencias de evaluación de planes de igualdad de género en España y los criterios de selección para la metaevaluación	481
Sobre el planteamiento de las evaluaciones de la segunda generación.....	482
Sobre los elementos clave de las evaluaciones	485
Sobre la utilización de los procesos de evaluación	487
Sobre el discurso de las responsables sobre la evaluación y las políticas de igualdad de género	489
Sobre los factores que facilitan y dificultan la evaluación y cómo conseguir evaluaciones de las políticas de igualdad de género más eficaces	491
Sobre el grado de verificación de las hipótesis iniciales	495
Sobre líneas abiertas para futuras investigaciones	496

BIBLIOGRAFÍA	499
---------------------------	-----

ANEXOS

- A. Documentación consultada
- B. Entrevistas realizadas

LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO DE LOS GOBIERNOS CENTRAL Y AUTONÓMICOS EN ESPAÑA: 1995-1999

ÍNDICE TABLAS

Capítulo 1

<i>Tabla 1.1.</i>	
Tres perspectivas para el análisis de políticas públicas	26
<i>Tabla 1.2.</i>	
Procedimientos analíticos generales según el momento y el tipo de cuestión.....	27
<i>Tabla 1.3.</i>	
El análisis valorativo en evaluación y sus cuatro niveles	31
<i>Tabla 1.4.</i>	
Diferencias entre la evaluación de programas y la evaluación de políticas.....	40
<i>Tabla 1.5.</i>	
Diferencias entre investigación y evaluación.....	51-52

Capítulo 2

<i>Tabla 2.1.</i>	
Esquema de Stake para la recopilación de datos de un programa.....	90
<i>Tabla 2.2.</i>	
Variables situacionales de Patton.....	95
<i>Tabla 2.3.</i>	
Tabla comparativa de Paradigmas.....	101
<i>Tabla 2.4.</i>	
Tipos de evaluación. Clasificación según las cuatro tipologías básicas	115
<i>Tabla 2.5.</i>	
Criterios de evaluabilidad según Chelimsky.....	117
<i>Tabla 2.6.</i>	
Listado de agentes críticos (<i>Stakeholders</i>) según Guba y Lincoln	122

Capítulo 3

<i>Tabla 3.1.</i>	
Estructuras gubernamentales para la igualdad de género en las comunidades autónomas....	149
<i>Tabla 3.2.</i>	
Evolución del problema de la igualdad	155
<i>Tabla 3.3. a.</i>	
Marco institucional de las políticas de igualdad autonómicas. Año 1998.....	165
<i>Tabla 3.3. b.</i>	
Marco institucional de las políticas de igualdad autonómicas. Año 2001	170
<i>Tabla 3.4.</i>	
Evolución de los planes estatales y autonómicos de igualdad de género en España.....	181
<i>Tabla 3.5.</i>	
Duración de los planes de igualdad.....	182
<i>Tabla 3.6.</i>	
Número de áreas, objetivos y acciones. Planes de igualdad estatales y autonómicos	183

Capítulo 4

<i>Tabla 4.1.</i>	
Tipos de metaevaluación.....	201
<i>Tabla 4.2.</i>	
Criterios para la metaevaluación.....	205
<i>Tabla 4.3.</i>	
<i>Key Evaluation Checklist (KEC)</i> de Michael Scriven.....	206
<i>Tabla 4.4.</i>	
Normas para la evaluación de programas.	
<i>The Joint Committee on Standards for Educational Evaluation.</i> 2ª Edición. 1994.....	208-210

Capítulo 5

<i>Tabla 5.0.</i>	
Tipos de evaluación. Segunda generación de evaluaciones.....	252
<i>Tabla 5.0.a.</i>	
Tipos de evaluación. Evaluaciones externas.....	260
<i>Tabla 5.0. b.</i>	
Tipos de evaluación. Evaluaciones mixtas.....	262
<i>Tabla 5.0. c.</i>	
Tipos de evaluación. Evaluaciones internas.....	263
<i>Tabla 5.1.</i>	
Contextos evaluativos	266
<i>Tabla 5.2.</i>	
Estructura institucional.....	288
<i>Tabla 5.3.</i>	
Costes directos de los estudios externos de evaluación	296

Capítulo 6

<i>Tabla 6.1. a.</i>	
Equipos externos de evaluación. Evaluaciones externas	305
<i>Tabla 6.1. b.</i>	
Equipos externos de evaluación. Evaluaciones internas mixtas	307
<i>Tabla 6.2. a.</i>	
Tiempos y momentos de las evaluaciones externas.....	311
<i>Tabla 6.2. b.</i>	
Tiempos y momentos de las evaluaciones internas mixtas	314
<i>Tabla 6.2. c.</i>	
Tiempos y momentos de las evaluaciones internas.....	316
<i>Tabla 6.3. a.</i>	
Criterios y preguntas de evaluación. Evaluaciones externas	318-321
<i>Tabla 6.3. b.</i>	
Criterios y preguntas de evaluación. Evaluaciones internas mixtas	326-328
<i>Tabla 6.3. c.</i>	
Criterios y preguntas de evaluación. Evaluaciones internas	331
<i>Tabla 6.4. a.</i>	
Procedimientos y metodología. Evaluaciones externas	335-337
<i>Tabla 6.5. a.</i>	

Instrumentos de recopilación de información utilizados. Evaluaciones externas.....	339
<i>Tabla 6.4. b.</i>	
Procedimientos y metodología. Evaluaciones mixtas	341-343
<i>Tabla 6.5. b.</i>	
Instrumentos de recopilación de información utilizados. Evaluaciones mixtas	345
<i>Tabla 6.4. c.</i>	
Procedimientos y metodología. Evaluaciones internas	346

Capítulo 7

<i>Tabla 7.1.</i>	
Pertinencia y utilidad de la información producida. Cuadro-resumen	357
<i>Tabla 7. 2. a.</i>	
Contenidos de los informes de evaluación. Evaluaciones externas	359-362
<i>Tabla 7. 2. b.</i>	
Contenidos de los informes de evaluación. Evaluaciones internas mixtas	367-369
<i>Tabla 7. 2. c.</i>	
Contenidos de los informes de evaluación. Evaluaciones internas.....	374

INTRODUCCIÓN

En el ámbito del análisis de políticas públicas, la fase de la evaluación ha sido la menos tratada y estudiada. Sin embargo, la evaluación de programas y políticas públicas es una realidad que, aunque tímidamente, ha ido creciendo de forma estable y positiva en las administraciones públicas y, en este momento, reclama la atención tanto de profesionales como de académicos. En España, se empiezan a realizar evaluaciones y esto está generando referencias empíricas en algunos sectores de política, aunque éstas distan mucho de estar sistematizadas. En general, son prácticamente inexistentes las investigaciones que tienen como objeto de estudio las propias evaluaciones que se están realizando de las diferentes políticas públicas.

Existe, por lo tanto, una necesidad de estudios sobre los incipientes procesos de evaluación que se están llevando a cabo en las administraciones públicas. Un enfoque novedoso para hacer dichos estudios es la realización de metaevaluaciones (evaluación de evaluaciones), que consisten en la recopilación, análisis y valoración sistemática de un conjunto determinado de procesos de evaluación. La metaevaluación es, por lo tanto, una herramienta especialmente útil para la mejor comprensión de la función de evaluación en determinado ámbito o sector y para la realización de propuestas de mejora y desarrollo de dicha función. De esta manera, se conseguirá aumentar, por un lado, la demanda de evaluaciones y, por otro, la oportunidad, la relevancia, el rigor y la validez de las evaluaciones que ya se realizan.

OBJETO

El objeto de estudio de la tesis doctoral que aquí se presenta es la evaluación de políticas públicas, aplicado al caso concreto de las políticas de igualdad de género¹. Para esto, hemos realizado una metaevaluación de las evaluaciones de los planes de igualdad de oportunidades entre hombres y mujeres (a partir de ahora planes de igualdad de género), realizadas en España en los niveles de gobierno central y autonómico, en la segunda mitad de la década de los años noventa.

La evaluación de políticas públicas se define aquí como la valoración de la acción emprendida por los poderes públicos para abordar un problema, en este caso concreto, de la desigualdad de género. Las evaluaciones analizadas, por lo tanto, implican valorar

¹ Por políticas de igualdad de género se entienden las políticas encaminadas a paliar la desigualdad existente entre los géneros o, si se quiere, en su denominación más común, las políticas de igualdad de oportunidades para las mujeres o entre mujeres y hombres. Para entender mejor el uso del término “igualdad de género”, véase el capítulo tres, especialmente el punto referido a la evolución en la definición del problema (punto 2.3.).

la acción de los poderes públicos durante un período determinado. Por lo tanto, distinguimos la evaluación del simple seguimiento (comprobación de que las acciones se están ejecutando tal y cómo se tenía previsto), que sólo se tiene en cuenta si forma parte de una estrategia más global de valoración final de dichas acciones.

Por otro lado, el caso escogido resulta especialmente adecuado e ilustrativo. Las políticas de igualdad de género en España son políticas innovadoras, no sólo porque son relativamente recientes sino también porque se conciben de forma integral y transversal, implicando varias áreas gubernamentales en su formulación y ejecución. Dichas políticas se conciben, en este trabajo, no sólo como el conjunto de medidas suficientemente coherentes entre sí que los gobiernos adoptan para evitar o paliar la desigualdad entre los géneros, sino también como los procesos que se generan con dicha intervención, es decir, el análisis y la definición del problema, la introducción en la agenda de las cuestiones y los problemas pertinentes, la toma de decisiones, la participación de agentes críticos, la coordinación de actores, la implantación de medidas, etc.. Lo que en esta investigación se analiza es cómo los gobiernos evalúan su acción en dicho terreno, así como el papel que cumple dicha evaluación en el proceso de la política pública.

Para evaluar una política pública, es imprescindible evaluar las acciones públicas que se realizan en el marco de dicha política, aunque la evaluación de la política vaya más allá de la simple evaluación de las acciones concretas. Así, el objeto general de la tesis doctoral que aquí se presenta se centra en una unidad de análisis más concreta, la evaluación de los planes centrales y autonómicos de igualdad de género. Efectivamente, en España, se elaboran los llamados “planes de igualdad” que consisten en un conjunto de medidas, para un período de tiempo determinado, aprobados por el Ejecutivo y que pretenden abordar el problema de la desigualdad de género de una forma integral y transversal, es decir, implicando a diferentes áreas o sectores gubernamentales. La elaboración de planes se convierte, a partir de la década de los años noventa, en el principal instrumento a través del que se articulan las políticas de igualdad de género central y autonómicas. Por lo tanto, siendo el objeto de la investigación la evaluación de las *políticas* públicas de igualdad de género, el cuerpo empírico de la misma coincide con la evaluación de los *planes* de igualdad. De hecho, no podría ser de otro modo, ya que las evaluaciones existentes son de los planes y no de las políticas. Más allá, los planes son el principal y prácticamente único instrumento de articulación de las políticas de igualdad de género y, además, pretenden abarcar todas las acciones que en dicha materia se realicen. Por lo tanto, en este caso concreto, el análisis y valoración de las evaluaciones realizadas de los planes de igualdad permite, de forma clara, una reflexión sobre la evaluación de las políticas de igualdad de género en general.

Los principales impulsores de las políticas de igualdad de género en España son los organismos de igualdad, como el Instituto de la Mujer -del gobierno central-, y sus homólogos en el nivel autonómico -Institutos autonómicos de la Mujer, Direcciones Generales de la Mujer o similares-. Por otro lado, las situaciones de evaluación cuentan siempre con un comanditario o “cliente”, es decir, alguien que requiere, pide o encarga la evaluación, incluso aunque ésta sea de carácter interno. En todos los casos analizados, dicho papel ha sido ejercido por el organismo de igualdad correspondiente, encargado de impulsar y coordinar el plan evaluado. Las evaluaciones, por lo tanto, han de responder ante dichos comanditarios y otros agentes críticos de la política evaluada. Así, la acción de conocimiento está muy mediatizada por los destinatarios que vayan a “consumirla” y sean susceptibles de utilizarla. De esta manera, el objeto de la presente

investigación no se limita a estudiar cómo se ha resuelto técnicamente la recopilación y el análisis de la información que permite la valoración de los planes, sino que se centra y le da especial importancia al contexto y a los procesos políticos y organizativos subyacentes a las situaciones de evaluación. Efectivamente, la evaluación supone procesos complejos fuertemente ligados a los procesos políticos (Palumbo, 1987).

MARCO TEÓRICO

La evaluación de políticas públicas se enmarca en esta investigación en el enfoque del **análisis de políticas públicas**. La evaluación se ha considerado tradicionalmente como una “fase” del ciclo de vida de una política pública -la fase final-. Y esto a pesar de que la evaluación ha ido ampliando su objeto de estudio y puede resultar un instrumento para el análisis de las otras fases (es decir, se puede hacer evaluación del diseño o formulación de una política, de su implementación o ejecución, etc.). Sin embargo, al haber sido erróneamente considerada como una fase exclusivamente “final”, es la que menos atención ha captado por parte de los analistas de políticas públicas.

Efectivamente, en el marco del análisis de políticas públicas, el aspecto menos desarrollado en la literatura especializada es el de la evaluación. De hecho, en muchos casos, los capítulos o secciones correspondientes a la evaluación suelen ser marginales o prácticamente inexistentes en obras o antologías generales importantes sobre políticas públicas, como ocurre en Meny & Thoening (1992) o en la antología sobre políticas públicas en cuatro volúmenes de Aguilar Villanueva (1996).

Sin embargo, el ámbito del análisis de las políticas públicas resulta especialmente adecuado para entender y situar los procesos de evaluación en el contexto político en el que se producen. Además, puede ayudar a resolver preguntas como: ¿Cuál es el papel que cumple la evaluación de políticas públicas? ¿Puede incidir sustancialmente en su análisis y desarrollo? ¿Se están llevando a cabo realmente evaluaciones? ¿De qué tipo? ¿Están sirviendo para algo? ¿Contribuyen a la institucionalización y normalización de la función de evaluación en las instituciones públicas? ¿Qué añaden?

Debido al relativo desarrollo de la evaluación en el marco del análisis de políticas públicas, otra de las bases teóricas sobre las que sustenta este trabajo de investigación es el campo concreto conocido como la **evaluación de programas**², en el que participan activamente algunos teóricos y científicos de la Ciencia Política y de la Administración que conciben la evaluación desde el análisis de políticas públicas (como por ejemplo, Palumbo, Chelimsky, Vedung o Monnier, y en España, Ballart o Zapico), pero en el que también participan profesionales y académicos de otras disciplinas, como son la sociología, las ciencias de la educación y la pedagogía, la

² Aquí se entiende programa en un sentido amplio, incluyendo en este concepto proyectos, planes, programas, políticas, medidas, servicios, materiales, etc.

psicología y las ciencias de la salud, la economía o el trabajo social, por no mencionar campos más concretos como la salud pública, la ecología humana, los recursos humanos y la teoría de la organización o el desarrollo regional. Este campo de la evaluación de programas se desarrolla a lo largo de los últimos años a la luz de la necesidad práctica de evaluar los programas públicos sociales y educativos que afloran en la segunda mitad del siglo XX y que conforman los intentos de los poderes públicos para solucionar o paliar los complejos problemas sociales de la época. Sin embargo, y a pesar de que el motor que ha movido el desarrollo de este campo conocido como la evaluación de programas haya sido la propia práctica evaluativa, se ha generado paralelamente todo un desarrollo teórico con sus características específicas concretas (Shadish, Cook y Levinton, 1991; Ballart, 1996).

El tercer elemento que sustenta este trabajo es el de los **estudios de género** (conocidos en el mundo anglosajón como *Women's Studies*), en los que también la ciencia política ha ido haciendo importantes aportaciones, especialmente en el terreno concreto del papel del Estado, los poderes públicos y los movimientos sociales y de mujeres en la lucha contra la discriminación de género. De forma especial, en los últimos años se ha desarrollado, aunque todavía sea tímidamente, la perspectiva del análisis de las acciones públicas que los gobiernos realizan para evitar la discriminación de género (estudios sobre el llamado Feminismo institucional o Feminismo de Estado) y que se han ido generalizando a lo largo del último cuarto del siglo XX (Sasoon, 1987; Meehan y Sevenhuijsen, 1991; Stetson y Mazur, 1995; Gardiner, 1997; en España, Valiente, 1995 y ss.; Gil Ruíz, 1996; Martínez, 1997 y Granados, 1999).

Se deben señalar también aquí las **razones de relevancia y oportunidad científica y profesional** del tema de investigación escogido. En concreto, al escaso desarrollo teórico de la evaluación en el campo concreto de las políticas públicas, se corresponde un también tímido desarrollo práctico de la evaluación de políticas públicas en el ámbito de las administraciones públicas, especialmente en algunos países que, como España, no cuentan con una cultura de evaluación enraizada en su sistema político-administrativo (precisamente la consolidación de la función de evaluación en las instituciones públicas de algunos países, como por ejemplo EE.UU., Canadá, Australia y en Europa el Reino Unido o Suecia, es lo que provoca el desarrollo teórico en el campo de la evaluación de programas). Sin embargo, sí parece haber un consenso cada vez más generalizado, tanto en el ámbito académico como en el de la gestión pública y la ciudadanía en general, sobre la conveniencia y la necesidad de evaluar la acción de los poderes públicos. Esta necesidad práctica de la evaluación se asienta sobre la información que aquélla puede potencialmente arrojar y que resulta indispensable para la formulación, ejecución y mejora de las políticas públicas, es decir, como una herramienta clave en la gestión pública (Monnier, 1992). Por lo tanto, la percepción cada vez más clara de la conveniencia y la necesidad de incorporar la evaluación como una herramienta útil para el mejor gobierno y gestión pública, hace que sean imprescindibles las investigaciones que, con una sólida base empírica y científica, analicen la realidad, ayuden a la reflexión y propongan algunas líneas para introducir, mejorar y fortalecer la evaluación en las organizaciones públicas.

La perspectiva descriptiva de la fase de evaluación de las políticas públicas ha sido sistemáticamente olvidada en el ámbito de las políticas públicas y de la evaluación de programas. Así como se han realizado estudios sobre las fases de la *definición del*

problema (Moore, 1976; Bardach, 1981) y de la *entrada en agenda* (Cobb y Elder, 1972; Elder y Cobb, 1984; Nelson, 1978), la *toma de decisiones* (Lowi, 1964; Etzioni, 1967; Allison, 1969; Majone, 1978; Lindblom, 1979) y de la *ejecución o implementación* (Pressman y Wildavsky, 1973; Rein y Rabinovitz, 1978; Berman, 1978; Sabatier y Mazmanian, 1981; O'Toole, 1989), no se encuentran en la literatura descripciones monográficas sobre la fase de *evaluación*. Incluso recurriendo al campo de la evaluación de programas, el también escaso desarrollo de la metaevaluación (evaluación de la evaluación) está más relacionado con el control de calidad de las evaluaciones y la posible acreditación de los evaluadores (Scriven, 1991; Schwandt y Halpern, 1988; Love, 1994) que con el estudio y análisis descriptivo de los procesos de evaluación como objeto de estudio.

De esta manera, aparece como interesante y novedoso la realización de análisis descriptivos de la fase de evaluación: ¿Cómo se está desarrollando la función de evaluación en torno a unas determinadas políticas e instituciones públicas? ¿Cuáles son las dificultades que se encuentran? ¿Cómo se inserta esta fase en el ciclo general de vida de una política? Insistimos aquí en que, además, la metaevaluación de los procesos de evaluación que en torno a unas determinadas políticas se realizan (y no tanto de los resultados de las evaluaciones), permite un análisis y una valoración que puede arrojar conclusiones especialmente ilustrativas y útiles sobre la función de evaluación en un determinado sector.

La **elección del estudio de caso**, las políticas de igualdad de género autonómicas y centrales en España, responde a una serie de razones que se exponen a continuación:

- Son políticas innovadoras y relativamente recientes en el tiempo. Esto las hace especialmente adecuadas, dado que las políticas innovadoras son más complicadas de llevar a cabo y, por lo tanto, la evaluación resulta indispensable para ofrecer retroalimentación sobre la acción, sugerir mejoras y estimar si es oportuna su generalización. Pero, además, los programas innovadores se realizan en un contexto de no-rutina, por lo que los actores carecen de un sistema de referencias previo para evaluar su repercusión. Más allá, la propia evaluación puede inducir a que se redefina el problema que ha motivado la actuación (Monnier, 1995: 14).

- Estas políticas responden a una nueva definición de la desigualdad de género como un problema susceptible de ser abordado por los poderes públicos. El perfil del problema a tratar, complejo y multiforme, y que requiere de una redefinición continua es un ejemplo interesante del tipo de problemas a los que han de enfrentarse las complejas y fragmentadas sociedades democráticas del S. XXI; especialmente, teniendo en cuenta que el origen de dicho problema está firmemente enraizado en siglos de historia y en la discriminación sistemática de la mitad de la población. La valoración de la respuesta, en muchos casos contradictoria además de novedosa, que los poderes públicos ofrecen ante dicho problema, aparece como un objeto especialmente atractivo de estudio.

- Son políticas que cuentan, en la actualidad, con un aval y un apoyo internacional. Crucial ha sido el impulso de las Organización de las Naciones Unidas y las Conferencias Mundiales sobre la Mujer (México, 1975; Copenhague, 1980; Nairobi, 1985; y, de forma especial, Pekín, 1995), así como la formulación de políticas

y medidas de acción positiva en algunas organizaciones internacionales y supranacionales (destacando el desarrollo de los “Programas de Acción” de la Unión Europea). En este sentido, son políticas que acaparan la atención de organizaciones internacionales y gobiernos nacionales, al menos de manera formal, y, por lo tanto, y de forma creciente, de profesionales, académicos e investigadores.

- En España, las políticas de igualdad de género han tenido un rápido desarrollo y, en cierto modo, paralelo al ritmo de construcción autonómica en España. En la actualidad están presentes tanto en el nivel central como en todas las Comunidades Autónomas. A pesar de que las originales competencias en materia de “promoción de la mujer” se transfirieron desde un principio incluso a las Comunidades Autónomas que accedieron a la autonomía por la “vía lenta” (acogiéndose al artículo 143 de la Constitución Española), el desarrollo inicial de las políticas de igualdad de género en el nivel central fue el punto de partida y un referente crucial para el desarrollo de dichas políticas en el ámbito autonómico. Por otro lado, a lo largo de los últimos quince años, las políticas de igualdad de género autonómicas han ido adquiriendo tal potencia y protagonismo propio que se tendría una visión muy sesgada si sólo se analizara el nivel central. Por lo tanto, parece lógico que el objeto de estudio se centre en los dos niveles gubernamentales mencionados.

- En España, las políticas públicas de igualdad de género son, en cierto modo, homogéneas entre sí, en el sentido de que cuentan con una estrategia común general en los dos niveles gubernamentales estudiados y que consiste, como ya se ha comentado, en la elaboración, aprobación y ejecución de planes de igualdad. Esto las hace especialmente idóneas para formar parte de un estudio de caso que permite una comparación transversal entre los diferentes subcasos. Esto es, el campo de estudio es diverso pero con elementos lo suficientemente homogéneos como para establecer pautas de comparación de una experiencia a otra.

- Finalmente, son políticas horizontales, que pretenden comprometer transversalmente a todas las instancias gubernamentales implicadas. En este sentido, cada vez más se plantean soluciones integrales a los problemas complejos que los poderes públicos abordan. Este tratamiento integral es una forma innovadora y cada vez más generalizada de abordar los problemas sociales. La evaluación de estas políticas complejas es lógicamente también compleja y representa un nuevo reto público y social.

Además de los fundamentos teóricos, su relevancia científica y las razones de selección del estudio de caso, queremos señalar, como último punto de este apartado, las posibles **limitaciones** de esta investigación.

1. La primera limitación posible está relacionada con un dilema inevitable y que ha sido ampliamente reflexionado en metodología comparada: la opción de uno o varios casos. Al optar por incluir en el estudio de caso todos los procesos de evaluación existentes en torno a los planes de igualdad autonómicos y centrales en un período determinado (en realidad, tenemos once subcasos), se gana en alcance, ya que se metaevalúan todos los procesos de evaluación que se han realizado, pero se pierde en profundidad ya que queda fuera de los posibles objetivos de esta investigación estudiar los once subcasos de forma detallada. Somos conscientes de que cada una de las once

evaluaciones estudiadas requeriría de un estudio de caso propio en profundidad; sin embargo, la forma idónea de obtener una visión más global y general que contribuyera al estudio de la función de evaluación en las políticas de igualdad de género en los dos niveles gubernamentales escogidos, era la realización de una metaevaluación que incluyera la totalidad de las experiencias de evaluación existentes.

2. La segunda limitación tiene que ver con las dificultades en la recopilación de información. La disponibilidad de información en cada uno de los subcasos estudiados ha sido muy irregular, no sólo porque los procesos de evaluación están diferencialmente documentados sino porque las actitudes respecto a hacer o no disponible la información también varían. En este sentido, creemos que la falta de cultura de evaluación y la propia novedad de estos procesos han jugado en contra; así, por ejemplo, no existe ningún criterio común sobre el carácter público de los informes de evaluación ni sobre la posible facilitación de los mismos para esta u otras investigaciones.

3. Asimismo, hay algunos criterios de la metaevaluación, los relacionados con el impacto de los procesos de evaluación, para los que ha sido más difícil obtener información. Esto es debido a la propia naturaleza de la información requerida. Es decir, obtener una información completa y fiable sobre la utilización real de los resultados de evaluación, la creación de una cultura y el desarrollo de la función de evaluación, el aumento de la conciencia sobre las políticas de igualdad de género, o la mejora real de los sistemas de información en los organismos de igualdad, hubiera requerido de una investigación de mayor alcance, y con una perspectiva temporal más amplia, de lo que este estudio de caso puede ofrecer. Sin embargo, no hemos querido renunciar a obtener una información que de forma más indirecta (por ejemplo, la percepción de responsables políticas y técnicas al respecto, la adquisición creciente de compromisos de evaluación, los planes de evaluaciones futuras, modificaciones introducidas en los sistemas de información, etc.), contribuyera a esclarecer y ofreciera una idea general respecto a dichos aspectos.

4. La última posible limitación se relaciona con algunas dificultades de generalización. Los procesos de evaluación metaevaluados son muy heterogéneos entre sí, tanto que, en muchas ocasiones, era inicialmente difícil el establecimiento de tendencias o conclusiones generales. En este sentido, hemos procurado reflejar la diversidad encontrada y hemos sido prudentes a la hora de concluir si existían o no determinadas regularidades. Por esta razón, la exposición de los datos e información hallada es, en algunas ocasiones, excesivamente minuciosa y puede resultar difícil obtener una visión de conjunto en cada momento. Esta minuciosidad también tiene que ver con nuestra decisión de reflejar, lo más fielmente posible, las diferentes formas de abordar la evaluación de los planes de igualdad.

OBJETIVOS E HIPÓTESIS DE LA INVESTIGACIÓN

Los **objetivos** generales de esta tesis doctoral son:

- Desarrollar el concepto de evaluación de políticas públicas y analizar su aplicación práctica en las políticas públicas de igualdad de género
- Conseguir una mayor elaboración teórica sobre el significado, el papel y las posibles formas de utilización de las políticas públicas de igualdad de género.

Estos objetivos generales se concretan en los siguientes objetivos específicos:

1) Analizar la utilización que desde las administraciones públicas autonómicas y la administración central se está haciendo de la evaluación como un instrumento de análisis de las políticas públicas de igualdad, a través de:

1.1) Análisis comparativo de los procesos de evaluación que se han llevado a cabo de las políticas autonómicas y centrales de igualdad de género en la segunda mitad de la década de los años noventa

1.2) Análisis del discurso político de responsables de los organismos de igualdad sobre el papel de la evaluación en el desarrollo de las políticas de igualdad de género en España.

2) Identificar los factores que han facilitado y dificultado la evaluación de los planes de igualdad en los niveles gubernamentales central y autonómico.

3) Extraer conclusiones sobre posibles líneas de futuro para la evaluación de las políticas de igualdad de género.

Las **hipótesis** que formulamos sobre los resultados previsibles son las siguientes:

1: Existen factores que dificultan la evaluación de los planes de igualdad (abstracción de los mismos, falta de equipos preparados y de experiencia previa, no anticipación de la misma, falta de insitucionalización y cultura evaluativa, etc.) y factores que la facilitan (la propia concepción de planes, institucionalización de compromisos respecto a la evaluación, etc.).

2: Entendida la eficacia de una evaluación como el cumplimiento de los fines para los que se diseña, es decir, si resulta útil para la mejora de los planes y políticas, para el control de responsabilidades y/o para ilustrar las acciones y políticas futuras, entre las evaluaciones realizadas de los planes de igualdad centrales y autonómicos, las más eficaces son las que se producen:

a) Con una mayor amplitud y visión global respecto a los contenidos que se evalúan (si además del grado de cumplimiento de las acciones se evalúa el propio diseño del plan, el proceso de puesta en práctica y sus efectos).

b) En el momento adecuado, es decir, planificándose desde el principio del plan y presentándose los resultados a tiempo para la toma de decisiones.

c) Utilizando agentes externos simultáneamente a una fuerte implicación de agentes internos.

d) Con una mayor interpretación valorativa y orientada a la acción, es decir, “traduciendo” los datos en valoraciones y recomendaciones concretas.

e) Con una mayor comunicación y publicidad de sus resultados.

3: El hecho de que se produzcan procesos de evaluación de los planes de igualdad puede, además de producir resultados útiles para los propios planes, tener los siguientes efectos colaterales beneficiosos³:

a) introducción de una cultura de la evaluación en los organismos de igualdad e incluso que pueda ir más allá de dichos organismos y extenderse a otras unidades administrativas (por la horizontalidad de las políticas de igualdad)

b) mayor conciencia de los procesos y políticas públicas de igualdad de género

c) mejora de los sistemas de información sobre acciones públicas de igualdad de género

METODOLOGÍA

Para la realización de la investigación que aquí se presenta hemos seguido el **método del estudio de caso**. Si, por el contrario, hubiésemos decidido seguir el método comparado, habríamos supuesto que los casos a comparar –en este caso, evaluaciones de otras políticas públicas- se parecían lo suficiente entre sí como para analizarlos –y, por lo tanto, compararlos- en relación a ciertos criterios. No sólo creemos que todavía estaría por demostrar que existen otras políticas públicas (y, por lo tanto, sus evaluaciones) “comparables” con las políticas de igualdad de género, sino que también se perdería la lógica de un caso concreto estudiado en su complejidad.

Escogiendo un caso concreto –las políticas centrales y autonómicas de igualdad de género-, pretendo dar cuenta de un conjunto de resultados (cómo se evalúan dichas políticas) situados en su contexto y momento, componiendo la evidencia o información de manera “sensible” a la cronología histórica y procurando ofrecer generalizaciones históricas limitadas que tengan en cuenta y se fundamenten en dicho contexto (políticas, planes y organismos de igualdad en la segunda mitad de la década de los años noventa).

³ Es decir, la evaluación puede tener efectos no sólo respecto a los resultados a raíz de su aplicación, sino también por su mera aplicación como instrumento de gestión en las administraciones públicas.

Si se tiene en cuenta el valor que normalmente se le otorga a los estudios de caso y los diferentes tipos que se encuentran -ateóricos, interpretativos, generadores de hipótesis, confirmadores de teoría y el estudio de caso discordante- (Lijphart, 1971), el estudio de caso que aquí se presenta podría ser clasificado como *generador de hipótesis*, ya que uno de nuestros objetivos era desarrollar algunas generalizaciones (quizá, de manera más concreta, posibles tendencias) teóricas sobre la función de evaluación de políticas públicas, ilustrando con un caso -la evaluación de políticas públicas de igualdad de género en un contexto y momento determinado-, en áreas en las que no existe un cuerpo teórico ni empírico importante. De esta forma, creemos que la evaluación de la política de igualdad de género permite identificar algunos rasgos teóricos derivados de su especificidad⁴, especialmente la derivada de su carácter transversal e integral, su carácter persuasivo y no sancionador y su carácter simbólico.

De todas formas, y dada la elección del método del estudio de caso, aunque nuestra finalidad es la de establecer relaciones entre diferentes variables, las conclusiones sólo pueden ser en términos de tendencias y no tanto en términos de causalidad. De hecho, en la exposición de los datos nos hemos guiado por la inevitable complejidad, coyunturalidad y contextualidad de los fenómenos estudiados (los procesos de evaluación). La complejidad viene dada por la necesidad de considerar varias variables independientes y la imposibilidad de considerar una única causa simple. La coyunturalidad, porque los fenómenos de evaluación son el efecto de varias actuaciones y de su intersección en el tiempo y el espacio. Y la contextualidad, porque cualquier causa específica puede tener efectos diferentes dependiendo del contexto. Por lo tanto, se analizan los fenómenos de evaluación realizados en torno a las políticas y a los organismos de igualdad de género, en su especificidad, procurando ser fiel a su complejidad, coyunturalidad y contextualidad. Aunque esto no obvia, como ya se ha señalado, la intención de establecer ciertas tendencias, regularidades y generalizaciones.

La **unidad de análisis**, por lo tanto, son las evaluaciones de los planes de igualdad centrales y autonómicos, realizadas durante la segunda mitad de la década de los años noventa en España. Pero también hemos considerado como unidades de análisis, en primer lugar, los propios planes de igualdad de “segunda generación”, tanto los que fueron evaluados como los no evaluados, especialmente en lo que se refiere a su grado de generalidad y su grado de compromiso explícito sobre el seguimiento y la evaluación. En segundo lugar, se considera también unidad de análisis el discurso sobre la evaluación de las responsables políticas de los organismos específicos de igualdad que impulsan y coordinan los planes, hayan o no sido evaluados, incluyendo las percepciones sobre las dificultades para evaluar los planes de igualdad.

Las **técnicas para la recopilación y análisis de información** utilizadas han sido esencialmente tres: recopilación y análisis documental, entrevistas abiertas semiestructuradas a las responsables políticas de los organismos de igualdad y entrevistas telefónicas estructuradas -en algunos casos presenciales- a las responsables técnicas de evaluación de los organismos de igualdad.

⁴ Esta especificidad de las políticas de igualdad de género se desarrolla tanto en la parte II (capítulo tres) como en la parte IV (capítulo ocho y nueve).

Las **fuentes de información** a las que hemos recurrido han sido, esencialmente, los organismos de igualdad responsables de los planes (tanto en lo que refiere a la documentación generada en torno a los mismos, como a las responsables políticas y técnicas de los mismos). En algunos casos nos hemos apoyado también en los responsables técnicos externos que realizaron las evaluaciones, tanto a través de entrevistas personales (telefónicas o presenciales), como a través de la documentación producida por los mismos.

Respecto a la recopilación y el análisis documental, el cuerpo empírico más importante de esta tesis doctoral son los informes de evaluación producidos por cada organismo de igualdad o por los equipos responsables de las evaluaciones en los casos de las evaluaciones externas. Pero también se han consultado y utilizado⁵ los propios planes de igualdad, las leyes, decretos y, en general, la normativa legal en torno a los organismos y a los planes de igualdad, así como otros documentos internos, por ejemplo, memorias de actividades, etc. Sin embargo, no se encuentran muchos documentos institucionales o declaraciones explícitas que traten el tema de la evaluación en los organismos de igualdad. Como ya se ha dicho, la principal documentación hallada han sido los informes de evaluación aunque, como se verá en el capítulo siete, en muchos casos son parcos a la hora de describir los procesos de evaluación pasando directamente a exponer los resultados hallados.

Las entrevistas tenían como objetivo captar el discurso de las responsables políticas sobre la evaluación de los planes y las políticas de igualdad. En ese discurso se incluye la propia valoración que realizan dichas responsables de las evaluaciones y, por extensión, de la función de evaluación de las políticas de igualdad de género en el seno de dichos organismos. Además de recoger información sobre las evaluaciones en sí, pretendíamos captar la perspectiva institucional de dichas evaluaciones, comprender de una forma más global el contexto en el que se generan, así como saber cómo se entiende la función de evaluación y qué papel cumple en los organismos de igualdad. Al no disponer, en la mayoría de dichos organismos, de documentación institucional sobre la evaluación, resulta especialmente importante captar el discurso que sobre la misma tengan sus máximas responsables. En definitiva, el objetivo de las entrevistas era la captación del discurso político en torno a los procesos de evaluación: qué papel cumple la evaluación de las políticas de igualdad en los organismos que las impulsan y para las propias políticas, qué esperan y cómo utilizan la evaluación, cuáles son/han sido sus principales dificultades, etc.

Las entrevistas se realizaron de forma abierta y con un guión semiestructurado, ya que más que una información cerrada y concreta, se quería captar el discurso general de las entrevistadas sobre la evaluación, desde su propia perspectiva. Para esto, se dirigieron las entrevistas con la técnica no directiva, facilitando el discurso de una forma no estructurada. Se realizaron entrevistas a dieciséis directoras de organismos de igualdad, de las cuales once lo eran de los organismos que llevan a cabo los procesos de evaluación analizados. De este modo, entrevistamos a todas las responsables políticas de los organismos de igualdad autonómicos (exceptuando a las directoras del Instituto Navarro de la Mujer y de la Dirección General de la Mujer de Cantabria), y a la

⁵ Véase el anexo correspondiente a la documentación consultada.

directora del Instituto de la Mujer. Posteriormente, se realizó un análisis del discurso sobre las transcripciones de las entrevistas.

Finalmente, otra parte importante de la información recopilada proviene de las entrevistas telefónicas y personales realizadas al personal técnico responsable de los procesos de evaluación en los organismos de igualdad. Gracias a la información que estas personas nos proporcionaron se pudieron solventar muchas de las carencias de información que, sobre los procesos de evaluación, existían en la documentación consultada.

PLAN DE EXPOSICIÓN

La tesis doctoral se estructura en cinco partes diferenciadas. En la primera parte se expone el marco teórico sobre la evaluación de políticas públicas que sustenta la investigación realizada. La segunda consiste en una descripción pormenorizada de las políticas de igualdad de género en España, dado que el estudio de caso seleccionado es la evaluación de dichas políticas. En la tercera parte se desarrolla el cuerpo empírico de la tesis, describiendo las once experiencias de evaluación, según los criterios pre-establecidos para realizar la metaevaluación. La cuarta parte contiene el análisis valorativo global de la información expuesta en la parte anterior, incluyendo el análisis del discurso de las responsables de los organismos de igualdad sobre la evaluación y los factores identificados como elementos que facilitan o dificultan los procesos de evaluación de las políticas de igualdad de género. Por último, en la parte quinta se exponen las principales conclusiones de esta tesis doctoral.

La *parte I*, “*La evaluación de las políticas públicas*”, consta de dos capítulos. En el **primer capítulo**, “**La evaluación en el marco del análisis de políticas públicas**”, se aborda la evaluación desde el marco general del análisis de políticas públicas, exponiendo las referencias teóricas en las que se basa la investigación realizada. Para ello, se desarrolla de forma sucinta el concepto y las características del análisis de políticas públicas y se enmarca la evaluación de políticas públicas en dicho contexto. De este modo, se expone la idea de que la evaluación es la fase del ciclo de vida de una política menos estudiada desde la perspectiva del análisis de las políticas públicas y se fundamenta el hecho de que el desarrollo teórico y práctico de la evaluación como actividad científica diferenciada se ha dado más en un campo afín, conocido como la evaluación de programas. Por lo tanto, en este primer capítulo, se señala la necesidad y conveniencia de desarrollar la evaluación de políticas desde el marco general del análisis, pero integrando los avances teóricos desarrollados en el campo concreto de la evaluación de programas. Así, se discuten el concepto, el sentido, las funciones y los objetivos de la evaluación y, posteriormente, se desarrollan las características específicas de la evaluación de políticas y que la diferencian de otras prácticas científicas y de análisis.

Por lo tanto, después de establecer las diferencias entre investigación y evaluación, se señalan las características que consideramos más diferenciadoras y

específicas de la evaluación: su carácter político y su carácter práctico y aplicado a la acción. El *carácter político* de la evaluación implica un necesario reconocimiento del mismo tanto para la práctica evaluativa como para el desarrollo teórico de dicha actividad científica. Resulta necesario, pues, reconocer e identificar los procesos políticos subyacentes a todo contexto evaluativo: la evaluación, para resultar eficaz, debe partir de dicho contexto y tener capacidad de respuesta ante el mismo. Las implicaciones de este reconocimiento del carácter político de la evaluación van desde la formación de evaluadores más allá de las cuestiones metodológicas, a la necesidad de su credibilidad (y que implica algo más de la credibilidad “técnica”), pasando por la centralidad del concepto de temporalidad y la necesidad de obtener los resultados a tiempo de la evaluación. El *carácter práctico y aplicado* se refleja en dos cuestiones. A) En primer lugar, en el necesario desarrollo de los dos niveles de análisis específicos de la evaluación: el enjuiciamiento sistemático y el establecimiento de recomendaciones. Es decir, en evaluación, además de recopilar y analizar información sistemática, se emiten juicios de valor sobre dichos resultados a la luz de los criterios de evaluación establecidos previamente, y se sugieren líneas de acción (recomendaciones) para la mejora de los programas y políticas evaluadas. B) En segundo lugar, en la preocupación por el uso y la utilización de la evaluación que es lo que confiere y da sentido a la misma. En el último punto del capítulo se recoge esta preocupación y se discuten los diferentes tipos posibles de uso de la evaluación descritos y tipificados en la literatura especializada.

En el **segundo capítulo, “La evaluación de políticas públicas desde una perspectiva histórica”**, se realiza un repaso histórico, tanto de la propia evolución o historia de la evaluación como de los diferentes enfoques o modelos que se han propuesto a lo largo de sus años de desarrollo científico. Así, además de relatar una breve historia de la evaluación, se identifican dos factores que han influido en el desarrollo de la evaluación como actividad científica diferenciada y en la propuesta de modelos y enfoques. Por un lado, el grado de institucionalización y consolidación profesional de la misma, que ha sido diferente en cada país y, por otro, la propia evolución en el estudio y la concepción de los problemas sociales, ya que se ha ido perdiendo paulatinamente la credibilidad en las posibilidades de la ciencia para resolver los problemas sociales, reconociendo la complejidad de los procesos sociales y, por lo tanto, de su evaluación. En este segundo capítulo también se repasan los principales enfoques o modelos propuestos a lo largo de la historia de la evaluación, desde los más tradicionales (evaluación por objetivos y enfoques experimentales), hasta los más innovadores y actuales (los enfoques participativos, como la evaluación de cuarta generación, la pluralista o la *empowerment evaluation*), pasando por los enfoques orientados a la toma de decisiones, al cliente, al consumidor y a la utilización de la evaluación.

Además del enfoque histórico descrito, en el segundo capítulo se definen cuatro clasificaciones para establecer las tipologías básicas de evaluación, según el papel que cumple la evaluación, según el contenido o fase de la política o programa que se evalúa, según el agente evaluador y según el momento de la evaluación, que son las que se utilizan posteriormente para la clasificación de las evaluaciones estudiadas. También se mencionan otros tipos de evaluación y otras clasificaciones utilizadas por otros autores. Por último, se tratan otros tres temas básicos para la teoría y la práctica de la evaluación, y que son foco de atención de los científicos y profesionales: a) los actores y participantes en la evaluación -distinguiendo entre clientes, destinatarios y

agentes críticos (*stakeholders*)- b) el papel del evaluador dependiendo de los diferentes enfoques y estructuras institucionales, y c) las implicaciones éticas para la práctica evaluativa. Finalmente, y a modo de conclusión de esta primera parte, se extraen las cuestiones más importantes que, según nuestro punto de vista, deberían estar presentes en un enfoque integrador para la evaluación de las políticas públicas.

La *parte II*, “*Las políticas públicas de igualdad de género en España*”, consta de un único **capítulo (tercero)** en el que se describen dichas políticas en España en el nivel de gobierno central y en el autonómico. En él, se discute sobre el concepto e historia de las políticas de igualdad de género, poniendo especial énfasis en las difíciles relaciones que siempre han tenido el movimiento feminista y el Estado. También se explica el término de feminismo de Estado o institucional acuñado en la literatura especializada y que se refiere a la maquinaria institucional de organismos e instituciones específicas que se dedican a promover a la mujer y la igualdad de género. En un segundo punto se realiza un repaso histórico de lo que significan las políticas públicas de igualdad de género en España. Sus inicios se remontan formalmente a la creación del Instituto de la Mujer en diciembre del año 1983, por lo que se trata de una historia muy reciente aunque de crecimiento muy rápido. Tanto, que la evolución en la definición del problema que se intenta resolver (la desigualdad o discriminación de género) también ha sido significativa. Asimismo, se explica el marco institucional que sostiene dichas políticas en los niveles de gobierno central y autonómico y se hace una revisión de los actores que contribuyen a su formulación. Finalmente, en un tercer apartado, se tratan específicamente los planes de igualdad de género, ya que éstos son, actualmente, el principal instrumento en el que se articulan las políticas de igualdad de género en España. De este modo, se explica lo que son y significan estos planes, se repasa su evolución en los niveles central y autonómico desde sus inicios y se discuten sus principales fortalezas y debilidades.

La *parte III*, “*La práctica de la evaluación de las políticas públicas de igualdad de género en España*”, consta de cuatro capítulos (del capítulo cuarto al séptimo), y en ella se desarrolla la descripción empírica de los once procesos de evaluación metaevaluados.

En el **capítulo cuatro**, “**La evaluación de las políticas públicas de igualdad de género en España: marco para el análisis y la valoración**”, se explica y fundamenta el término y el concepto de metaevaluación, base metodológica y teórica sobre la que se asienta la descripción y valoración de los procesos de evaluación que son objeto de estudio. También se señala y desarrolla la necesidad y la conveniencia de analizar los procesos de evaluación en torno a las políticas públicas y, en especial, a las políticas de igualdad de género, dadas sus características de políticas públicas innovadoras, integrales y transversales. En segundo lugar, se realiza un panorama general de las evaluaciones realizadas en España en torno a las políticas de igualdad de género en los niveles gubernamentales central y autonómico, distinguiendo tres “generaciones” de evaluaciones y señalando que la metaevaluación objeto de esta tesis doctoral es la referida a la *segunda generación de evaluaciones*, dado que la primera consistió en unos pocos intentos parciales de evaluación y la tercera generación de evaluaciones está todavía por concluir. Finalmente, se establece el marco para la metaevaluación

definiendo los criterios que se utilizan para el análisis y la valoración de las experiencias de evaluación. Estos criterios, los cuales son definidos y explicados uno a uno en este capítulo cuarto, se agrupan en tres grandes dimensiones: el planteamiento de la evaluación y las estrategias empleadas, los elementos clave de los procesos evaluativos y la utilización de los mismos. Estas tres dimensiones son las que, respectivamente, van a guiar los tres siguientes capítulos de esta parte III. Es decir, una vez definidos los criterios en este capítulo, se pasa a la descripción de los once procesos de evaluación a lo largo de los capítulos cinco, seis y siete.

En el **capítulo cinco, “El planteamiento general de los procesos de evaluación: tipos, contexto, propósitos, estructuras y recursos”**, se realiza primero una clasificación de las once experiencias de evaluación según las cuatro tipologías básicas: según el papel de la evaluación, las clasificamos en evaluaciones sumativas o formativas; según el contenido o fases que se evalúan, en evaluaciones de diseño, proceso y/o resultados; según el momento en el que se evalúa, en evaluaciones *ex-ante* o *ex-post-facto*; y según el agente evaluador, en evaluaciones internas, externas o mixtas. Esta última clasificación es la que se selecciona para presentar los datos a lo largo de los tres capítulos. De las once experiencias, seis son externas (las correspondientes a los planes andaluz, canario, castellano-manchego, castellano-leonés, gallego y madrileño), dos son internas (correspondientes a los planes aragonés y riojano) y tres son mixtas (correspondientes al plan central del Instituto de la Mujer y a los planes catalán y vasco). Por lo tanto, para cada criterio de metaevaluación establecido, se distinguen estos tres tipos de evaluaciones. Los criterios que se exponen en este capítulo cinco y que se relacionan con el planteamiento general de los procesos de evaluación son, además de los tipos de evaluación realizados, 1) los contextos evaluativos -teniendo en cuenta las variables del tipo y la antigüedad del organismo de igualdad correspondiente y el área gubernamental a la que está adscrito, el partido político en el gobierno en el momento de la evaluación, los cambios de responsables políticas y si existía experiencia evaluadora previa-, 2) la claridad de los propósitos y los objetivos de las evaluaciones -si éstos han sido explicitados previamente o no-, 3) las estructuras institucionales de la evaluación -es decir, qué tipo de estructuras existen y están implicadas en el seguimiento y evaluación de los planes- y 4) los recursos existentes para la evaluación, tanto de carácter directo -reflejados en el presupuesto- como indirecto - reflejados en el tiempo de dedicación a la evaluación de los equipos técnicos internos de los correspondientes organismos de igualdad-.

En el **capítulo seis, “Los elementos clave de la evaluación: agentes, momentos, criterios y metodología”**, se exponen quiénes han sido los agentes y los participantes activos en los procesos de evaluación estudiados, y cuáles han sido los momentos y el calendario de los mismos -distinguiendo entre el momento del desarrollo del plan en el que se plantea la evaluación, la duración de los procesos de evaluación, el tiempo que pasa desde que finaliza el plan hasta que se entrega el informe final de evaluación, así como los retrasos producidos en las evaluaciones-. También se describe en este capítulo otro de los elementos clave en cualquier evaluación, como es lo que se evalúa, es decir, los criterios y los indicadores de evaluación que se han utilizado en cada proceso metaevaluado, así como cuáles han sido sus fuentes y su proceso de establecimiento. Este es definitivamente uno de los factores clave, no sólo porque define las variables a evaluar de los planes de igualdad, sino porque a través de los mismos se definen las perspectivas de valoración con las

que se enjuician dichos planes. Finalmente, se señalan y analizan también los procedimientos y la metodología empleada en las evaluaciones.

Las cuestiones relacionadas con la utilidad y la utilización real de los procesos y los resultados de las evaluaciones se tratan en el **capítulo siete, “La utilización de las evaluaciones: pertinencia, comunicación e impacto”**. En este capítulo se recogen datos y se analizan, en primer lugar, la *pertinencia y la utilidad de la información* que han producido las evaluaciones, tanto en los que se refiere al alcance y a la globalidad de la información producida, como a la traducción práctica para la aplicación de la misma. Por lo tanto, se refleja, por una lado, la pertinencia de los contenidos de los estudios de evaluación para el plan evaluado, para el organismo de igualdad protagonista de la evaluación, así como para la situación evaluativa concreta. Por otro lado, se estudia el grado de análisis y elaboración de conclusiones, así como la existencia y la calidad de las recomendaciones en los informes de evaluación, y a esto es a lo que denominamos el grado de “traducción práctica”, que tiene que ver con la aplicabilidad de los procesos de evaluación (un alto grado de elaboración práctica es una condición facilitadora para su utilización). En el segundo apartado de este capítulo siete, se trata el *proceso de comunicación de la evaluación* como otro elemento clave para la utilidad y utilización de la misma. Para esto, se analizan los informes de evaluación, su contenido y estructura, la difusión que se realiza de los mismos, así como sus destinatarios y el calendario de dicha difusión. Finalmente, en el último apartado de este capítulo, se analiza el *impacto que han tenido las experiencias de evaluación*, tanto en los propios planes y políticas de igualdad de género, como en los organismos de igualdad impulsores de dichas políticas. El impacto de las evaluaciones depende, en gran medida, del uso que se haga de ellas. Así, además de los usos de los *resultados* de la evaluación, se tratan también los usos de los *procesos* de evaluación, ya que estos últimos pueden servir, potencialmente, para introducir una cultura y una metodología de evaluación y para generar una mayor comprensión de la intervención pública y una mayor conciencia sobre cómo se abordan los problemas públicos. De esta manera, se exponen en este apartado los indicios encontrados sobre los efectos de los mismos procesos de evaluación, si se ha producido una mayor comprensión de las políticas de igualdad de género, como consecuencia de las evaluaciones. Igualmente se estudia si se han realizado acciones o modificaciones a resultados de las mismas.

La *parte IV* de la tesis doctoral, “*La valoración de la evaluación. El reto de la evaluación de las políticas de igualdad de género*”, se compone de los dos últimos capítulos. En el **capítulo ocho (“El discurso sobre la función de evaluación en torno a las políticas de igualdad de género”)** se analiza el discurso que sobre la evaluación de las políticas de igualdad de género tienen las responsables políticas de los organismos de igualdad. Se divide este capítulo en cuatro apartados, extrayendo conclusiones sobre las ideas asociadas y las actitudes ante la evaluación, sobre la utilidad de la evaluación (para qué sirve evaluar), sobre lo que hay que evaluar de las políticas de igualdad de género y, por último sobre las principales dificultades percibidas para realizar evaluación de los planes que impulsan. Estas dificultades están relacionadas o bien con la falta general de una cultura de la evaluación, con la propia complejidad de los planes o bien con la estrategia de transversalidad subyacente a las políticas de igualdad de género, que hace que la evaluación deba tener unas características propias.

El análisis global de la metaevaluación realizada, así como la valoración que del conjunto de las experiencias de evaluación se realiza, se refleja en el **último capítulo, “Los factores que facilitan y dificultan la evaluación de las políticas de igualdad de género. Hacia una evaluación innovadora de dichas políticas”**. En este capítulo se señalan los factores que identificamos como elementos influyentes en los procesos de evaluación estudiados. En primer lugar, se identifican una serie de factores propios de las políticas públicas de igualdad de género y que facilitan o dificultan su evaluación. Como elementos que facilitan la evaluación se señalan, entre otros, la propia existencia de planes, la institucionalización de compromisos respecto a los mismos, la fortaleza y consolidación de los organismos de igualdad, el aval internacional para las políticas de género y el reconocimiento -al menos simbólico- de la evaluación como una actividad necesaria en los últimos años. Sin embargo, otros factores la dificultan, como son la generalidad de los planes, su propio carácter transversal y no sancionador -y que debe utilizar la persuasión como principal estrategia política-, la falta de cultura y profesionales de la evaluación, así como no pensar dicha evaluación a tiempo. En segundo lugar se identifican los factores inherentes a los propios procesos de evaluación y que más han condicionado los mismos: la falta de claridad en los propósitos de la evaluación, de una visión global y de un reconocimiento del carácter político de la evaluación. Así mismo, se analiza la visión parcial de la evaluación y su consideración como una función secundaria, la falta de claridad en el qué evaluar, las dificultades de coordinación y recopilación de información, el déficit de traducción práctica y de comunicación, así como la importancia de los tiempos y los recursos dedicados a la misma. Por último, se discuten dos factores relacionados con el impacto de las evaluaciones: su relativa pertinencia y utilidad, y la imposibilidad de realizar síntesis de evaluaciones por la heterogeneidad de las mismas.

Finalmente, la *parte V* consiste en las *conclusiones* de la tesis. En ellas reflejamos las conclusiones formales de cada una de las partes y los capítulos de los que consta este trabajo doctoral, además de señalar algunas sugerencias y líneas de investigación futuras en torno a la evaluación de políticas públicas en general y de las políticas de igualdad de género en particular.

CAPÍTULO 1

LA EVALUACIÓN EN EL MARCO DEL ANÁLISIS DE POLÍTICAS PÚBLICAS

INTRODUCCIÓN

En este primer capítulo se exponen las referencias teóricas sobre evaluación de políticas públicas que sirven como marco para el análisis y la valoración de las evaluaciones (metaevaluación) que se ha realizado de las políticas públicas de igualdad de género en España, en los niveles de gobierno autonómico y central. Su concepción se basa en una revisión extensa de la literatura sobre evaluación de programas y políticas públicas.

En este trabajo existe la firme voluntad de abordar la evaluación desde el marco más general del análisis de políticas públicas. El papel que la evaluación juega en este marco de referencia es el objeto de estudio del primer punto de este capítulo. Para esto, se analizan diferencialmente los conceptos de análisis de políticas públicas y de evaluación de programas. Se expone el componente descriptivo y el prescriptivo del análisis de políticas públicas ya señalados por Lasswell, y se sitúa la evaluación de programas respecto a estas dos perspectivas. A continuación se reflexiona sobre la evaluación en el ciclo de vida de la política pública, fundamentando su necesidad y posibilidades, aunque se reconoce que ha sido la fase eternamente abandonada en el estudio de las políticas públicas. También se realiza una crítica a la concepción tradicional de la evaluación como una fase exclusivamente finalista y que se limita al análisis y valoración de los resultados de una política. A tal efecto, se expone un panorama general sobre cómo se ha ido ampliando el objeto de estudio de la evaluación de políticas públicas, poniendo en cuestión, incluso, la idea de que la evaluación tiene un exclusivo carácter retrospectivo y esto sea lo que la diferencia del análisis de políticas públicas. Por último se discute una visión más integrada del análisis de políticas públicas y se señala el papel que la evaluación tiene en la misma.

En una segunda parte del capítulo, se revisa, elabora y discute sobre el concepto y el sentido de la evaluación de programas y políticas públicas. Una vez discutido el concepto, se exponen y elaboran las funciones de la evaluación, diferenciándolas de sus propósitos y objetivos. En la literatura especializada y los manuales sobre evaluación se suele hablar indistintamente de la evaluación de proyectos, programas, planes y políticas; sin embargo, se han querido señalar aquí las diferencias que existen, en nuestro parecer, entre el concepto de evaluación de programas (y que incluye proyectos, planes, medidas, servicios públicos, etc.) y el de evaluación de políticas públicas.

Una vez que se ha explicado el concepto de evaluación en sus aspectos más importantes, se define la evaluación **diferencialmente**, poniendo énfasis en lo que la evaluación tiene de específico respecto a otras formas de análisis e indagación. Para esto, se presentan, en primer lugar, las diferencias existentes entre evaluación e investigación, con la idea de reforzar la identidad propia de la evaluación de programas y porque en la todavía incipiente práctica de la evaluación de programas en las administraciones públicas españolas tienden, en algunas ocasiones, a confundirse esos dos extremos.

También en esta tercera y última parte del capítulo, se insiste en el elemento valorativo de la evaluación y en la importancia del concepto de “enjuiciamiento sistemático” como una de las claves de la identidad de la evaluación de programas y políticas públicas. Paralelamente, se discute sobre el eminente carácter práctico y aplicado de la evaluación, y se elabora sobre el cuarto nivel de análisis en la evaluación –el establecimiento de recomendaciones–, exponiendo el debate existente sobre si deben los evaluadores o no establecer dichas recomendaciones. Finalmente, se trata la preocupación por el uso y la utilización que tradicionalmente ha ocupado a los profesionales y académicos de la evaluación de programas y que se concreta, entre otros aspectos, en la importancia creciente en la literatura especializada de la fase de presentación de la información y comunicación de los resultados de la evaluación.

1. EL ANÁLISIS DE POLÍTICAS PÚBLICAS: EL PAPEL DE LA EVALUACIÓN EN ESTE MARCO DE REFERENCIA

1.1. Concepto y características del análisis de políticas públicas

El análisis de políticas públicas se ocupa, en general, de los problemas públicos y de cómo los poderes públicos los abordan. La evaluación de las políticas públicas es una parte de ese análisis general que se ocupa de valorar la actuación de los poderes públicos en su intento de solucionar problemas.

Efectivamente, el objeto de estudio del análisis de políticas públicas consiste tanto en los problemas públicos y cómo estos son definidos como, precisamente, en cómo los poderes públicos intentan dar respuesta a dichos problemas. “Por lo cual, no se trata tanto de preocuparse de las características propias de las instituciones, de la normativa, o de los grupos presentes, como de considerar aquellos aspectos que resulten relevantes o significativos con respecto a la adopción y puesta en práctica de una política pública” (Subirats, 1992: 13).

Así, el análisis de políticas públicas tiene un carácter práctico y aplicado. Ni siquiera se plantea el estudio de los problemas públicos en abstracto, sino cómo estos son definidos para ser abordados. Este carácter aparece en el mismo concepto y, por ejemplo, William N. Dunn define el análisis de políticas como “una disciplina de ciencia social aplicada que utiliza métodos múltiples de indagación y discusión para producir y transformar información relevante para las políticas y que puede ser utilizada en contextos políticos para resolver problemas relacionados con las políticas” (Dunn, 1981: 35). Precisamente, esa capacidad e intención del análisis de políticas públicas de estudiar y explicar los procesos *concretos* de actuación de los poderes públicos dando repuesta a los problemas sociales, es lo que hace que se ponga énfasis en su aplicabilidad y utilidad para un mejor gobierno y tratamiento de los asuntos públicos.

Además, el análisis de políticas públicas ha ido ampliado su objeto de estudio, sobre todo a partir de la década de los años setenta. Esta, entre otras, es una forma en la que se ha ido dando respuesta a las cada vez mayores dificultades de gobierno y al abordaje de los múltiples problemas sociales en sociedades cada vez más complejas. Se ha pasado del simple análisis de la decisión pública y sus implicaciones, al estudio de las complejas acciones que se derivan de la misma y sus efectos sobre el planteamiento del problema original (Subirats, 1992: 14), o en términos de Wayne Parsons, se ha pasado del exclusivo interés en el análisis de la decisión (decision analysis), es decir, el análisis del proceso de toma de decisiones y el análisis de políticas para la toma de decisiones, al análisis de la entrega o reparto (delivery analysis), es decir, al análisis de la ejecución o implementación, la evaluación, el cambio y el impacto (Parsons, 1995).

1.2. El componente descriptivo y prescriptivo del análisis de las políticas públicas

Harold D. Lasswell, padre de la perspectiva de las políticas públicas, en su primer trabajo publicado en 1951 (“La orientación hacia las políticas”), ya hace una importante distinción entre el conocimiento del proceso de la política (*knowledge of*) y el conocimiento en el proceso de la política (*knowledge in*). La “ciencia de políticas” que él propone debe tener estos dos marcos de referencia. En el primero (conocimiento del o *análisis de la política*) se pretende conocer el proceso de decisión de la política así como de hecho sucede. En realidad se trataría de describir por qué determinados problemas son calificados como públicos, entran en la agenda, se definen, cómo y por qué se eligen ciertos cursos de acción y ejecución y se evalúan a través de ciertos criterios y metodologías de evaluación. En el segundo (conocimiento en o *análisis para la política*) se trata de incorporar los datos y los teoremas de las ciencias en el proceso de deliberación y decisión de la política, con el propósito de corregir y mejorar la decisión pública. Este marco de referencia tiene, entonces, un componente prescriptivo, tal y como lo define inicialmente Lasswell. Se trataba de ofrecer métodos e información pertinente para incrementar la racionalidad de la construcción, examen, selección y desarrollo de las opciones de la política. Como dice Aguilar (1996: 53) se trata de una teoría normativa de carácter tecnológico.

En realidad, para Lasswell, estas dos actividades de conocimiento (que corresponderían a los componentes descriptivo y prescriptivo del análisis de las políticas públicas) son interdependientes y complementarias, además de interdisciplinarias y contextuales. Sin embargo, en el desarrollo posterior de los estudios de políticas públicas hubo una escisión en dos corrientes: por un lado, de los interesados exclusivamente en el estudio positivo de la elaboración de la política (más ligado a los estudios de las *policy sciences* o el *policy making*), más contextualizadores y relacionados con lo político-administrativo y, por otro, los normativo-técnicos (más ligado a los estudios de *policy analysis*), más dados a las técnicas cuantitativas y de análisis económico (Aguilar, 1996: 54 y ss).

Estas dos tradiciones se van desarrollando y, a partir de la década de los años 70, y paradójicamente por los resultados desalentadores para las dos corrientes, hubo una expansión en EE.UU. de escuelas y programas de políticas públicas (Bañón, 1997: 41). De todas formas, aunque haya habido tradiciones separadas de estudio, también ha habido posturas eclécticas e integradoras. Hogwood y Gunn (1984) plantean el doble componente descriptivo y prescriptivo desde la misma definición del análisis de políticas públicas. Consideran ambos componentes necesarios, aunque consideran como inapelable y definitorio el carácter prescriptivo, por su aspecto novedoso y valioso respecto a la aplicación. Para estos autores, con este carácter, el análisis de políticas públicas tiene una naturaleza aplicada, relevante socialmente, multidisciplinaria, integral y dirigida hacia los problemas (Hogwood y Gunn, 1984:3).

Efectivamente, la Administración Pública y los subcampos de políticas públicas y de gestión pública tienen una tendencia a la acción o vocación profesional y una tendencia o vocación a prescribir. “Hay que evitar, sin embargo, la tendencia a extraer enseñanzas prescriptivas de enseñanzas insuficientes” (Bañón, 1997: 44). Es decir, la prescripción debe basarse necesariamente también en la descripción y ambos componentes, utilizados de forma conjunta, pueden tener una mayor potencia, utilizándose de forma sinérgica. Este carácter integrador también se detecta en otros autores, como por ejemplo Dunn, que veremos más adelante.

Desde nuestro punto de vista y para enmarcar la evaluación, conviene asumir una visión integradora de estos dos componentes. La evaluación, al igual que el análisis en general, se ve también en la disyuntiva de si el análisis de la política debe mantenerse al margen del análisis para la política. Sin embargo, la evaluación está más directamente vinculada a la doble exigencia de conocimiento y acción, ya que la calidad del razonamiento no es algo separable de la pregunta sobre su pertinencia práctica (Monnier, 1995: 259). Además, la fase de la evaluación de las políticas, dado el reconocimiento formal con el que puede contar el evaluador, y con el que a veces no cuenta el analista general (hay algo más de “mercado” para evaluadores que para analistas), se ha trabajado más en su aspecto prescriptivo, que en el descriptivo. Es decir, el evaluador ha intervenido estableciendo cómo debe hacerse una evaluación o, de hecho, haciéndola, más que estudiando o describiendo los procesos de evaluación. Sin embargo, esta segunda faceta descriptiva, que incluiría la realización de estudios de metaevaluación (evaluación de los procesos de evaluación), puede resultar especialmente interesante en contextos en los que no existe una cultura evaluativa afianzada, pero en los que, de hecho, están emergiendo con cierta fuerza diferentes prácticas evaluativas (como es el caso español). En este sentido, insistimos, de acuerdo con Hogwood y Gunn (1984), en la necesidad de incluir ambas perspectivas (descriptiva y prescriptiva) en el marco de análisis y, también, en la consideración de que la perspectiva prescriptiva es lo que confiere su carácter más específico y aplicado, tanto al análisis en general como a la evaluación de políticas públicas en particular. Así, del análisis descriptivo de los procesos de evaluación debe surgir, además de una comprensión general sobre las políticas públicas, iluminación sobre cómo hacer evaluación de dichas políticas en determinados contextos.

1.3. La necesidad de la evaluación en el ciclo de vida de la política

Además de los actores, el estilo de los procesos decisorios, la estructura de los problemas en juego y las propias reglas del juego (procedimientos formales seguidos en los procesos de adopción de las políticas públicas), una de las categorías analíticas más utilizadas en el estudio de las políticas públicas es la del ciclo de vida de una política (es decir, la dinámica que caracteriza las distintas fases en que se articula dicho ciclo de vida). Bajo esta perspectiva, se propone la posibilidad de estudiar una política pública descomponiendo en diversas fases su ciclo de vida: la definición del problema y la

entrada en la agenda de los poderes públicos, la formulación de propuestas y la adopción de decisiones, su implementación, la valoración de los resultados y la eventual sucesión o terminación de la política iniciada. De hecho, esta categoría analítica puede incluir las anteriores ya que cada una de las fases descritas puede distinguirse de las demás por el tipo de actores dominantes, el contexto institucional y los estilos de decisión prevalentes. Por ejemplo, mientras en la fase de entrada en la agenda el papel de los medios de comunicación suele ser crucial (en el intento de captación de la atención pública), éste se ve atenuado durante la fase de implementación en la que los colectivos beneficiados o perjudicados son los que siguen directamente interesados en su evolución (Regonini, 1991:80). Es decir, existe una redefinición de los actores de una fase a otra y, esto hace que esta perspectiva de análisis adquiera un interés especial: en lugar de analizar, de forma tradicional, las características de cada fase por separado, se intenta reconstruir la dinámica interna del ciclo de vida de una política.

Este marco de análisis, representado a través de las fases del ciclo de una política, y en sus diferentes versiones, es válido además para la doble faceta descriptiva y prescriptiva del análisis de políticas públicas (Hogwood y Gunn, 1984: 4).

Bajo esta perspectiva del ciclo de vida de una política pública, la evaluación se convierte en una de las fases de desarrollo. Si se estudian y analizan todas las fases, la evaluación tiende a ser concebida como una de las fases “finales”, aunque la mayoría de los autores reconocen la “circularidad” de las políticas públicas, y se suele limitar a la evaluación final de los resultados de la política. Sin embargo, el campo de la evaluación de programas y políticas se ha ido desarrollando en los últimos años hacia una clara expansión del ámbito que abarca, tanto en lo que se refiere a la metodología como a las preguntas que responde. Inicialmente, la evaluación se concebía exclusivamente como un medio para medir los resultados y los efectos de una política o programa. En la actualidad, nadie discute que la evaluación puede abarcar las diferentes fases de la política o programa, desde su diseño o concepción hasta su impacto.

Sin embargo, la evaluación en el marco de las políticas públicas, es la fase menos estudiada desde el punto de vista de su análisis. Un buen indicador de esta tendencia es el hecho de que de las cuatro excelentes antologías compiladas por Aguilar Villanueva (“El estudio de las políticas públicas”, “Problemas públicos y entrada en agenda de gobierno”, “La hechura de las políticas” y “La implementación de las políticas”) no hay ninguna que aborde los temas de evaluación. Se puede decir que una buena parte del desarrollo teórico sobre la evaluación de programas y políticas se ha realizado fuera del marco de la disciplina del análisis de políticas públicas, aunque hay toda una serie de autores que bien debido a su disciplina inicial (por ejemplo, Wholey, Vedung, Monnier), bien a su especial interés por (por ejemplo, Weiss) o su inmersión profesional (por ejemplo, Chelimsky) en los procesos políticos, se han encargado de hacer el engarce entre ambas (sub)disciplinas.

1.4. El carácter retrospectivo y prospectivo del análisis de políticas públicas

Algunos autores (Rist, 1990; Ballart, 1993; Vedung, 1997) señalan el carácter retrospectivo de la evaluación, dado que pretende ofrecer información sobre lo que ha ocurrido o está ocurriendo. Además, en la primera acepción de la evaluación como medio para evaluar los resultados y efectos de un programa o política, resultaba lógico e inevitable considerarla de forma exclusivamente retrospectiva. A medida que se va ampliando el foco de la evaluación, sin embargo, este carácter exclusivamente retrospectivo no queda tan claro. De hecho, en el marco de algunos organismos internacionales y algunos países, como Suecia, se desarrollan modalidades de evaluación "*ex-ante*", que hacen difícil su clasificación como actividad retrospectiva. Este carácter retrospectivo puede implicar, además, un cierto riesgo de considerar la evaluación exclusivamente hacia el final de los procesos, perdiendo una gran parte de su potencial como elemento integrador, y desde un principio, del ciclo de vida de un programa o política.

Precisamente, es ese carácter retrospectivo el elemento que Rist (1990) considera diferenciador entre la evaluación de programas y el análisis de políticas. Mientras la evaluación de programas valora las políticas públicas de forma retrospectiva, el análisis de políticas lo hace de forma prospectiva, es decir, busca informar decisiones que todavía tiene que tomarse. El foco del análisis de políticas es, pues, el análisis prospectivo y el de la evaluación de programas el retrospectivo. También para Chelimsky (1985), esta diferencia afecta al tipo de preguntas que cada uno de los tipos de indagación puede hacerse: el análisis de políticas pone énfasis en los efectos *probables*, mientras la evaluación de programas lo hará en los efectos *reales*. Efectivamente, ésta es una diferencia general entre evaluación y análisis, pero el propio Rist reconoce que "lo que aparece a primera vista como una dicotomía clara - retrospectivo y prospectivo- se convierte, si miramos más de cerca, en dos pasos de un proceso interactivo: se toman decisiones, se recoge información sobre los efectos de dichas decisiones, se toman más decisiones con los datos disponibles de los resultados de decisiones previas, etc., etc." (Rist, 1990: 4).

En definitiva, aunque el carácter retrospectivo sirva como elemento diferenciador del análisis de políticas públicas en general, que tiene un carácter más prospectivo, hay que reconocer la amplitud de miras cada vez mayor de la evaluación. Es esta amplitud, precisamente, la que permite que la evaluación sea parte integrante de la vida de una política y que sirva como herramienta metodológica útil para la totalidad de dicho proceso.

1.5. Una visión integrada del análisis de políticas públicas: el papel de la evaluación

Como ya se ha comentado anteriormente, algunos autores han hecho un esfuerzo de integración de las diferentes perspectivas y componentes del análisis de políticas públicas. Según Dunn (1981: 36 y ss.), por ejemplo, el análisis de políticas públicas debe generar información sobre tres tipos de cuestiones:

- 1) sobre *valores*, cuyo logro es la prueba principal sobre si un problema se ha resuelto,
- 2) sobre *hechos*, cuya presencia puede limitar o mejorar el logro de los valores, y
- 3) sobre *acciones*, cuya adopción puede resultar en el logro de los valores y la resolución de los problemas.

El analista al generar información sobre las anteriores tres cuestiones puede emplear una o más de estas tres perspectivas de análisis:

- 1) *Empírica*, que describe las causas y los efectos de una política pública dada. Las cuestiones se refieren a los hechos y el tipo de información que se genera es denominativo o designativo. Un posible ejemplo sería la descripción, explicación o predicción del gasto público en sanidad o educación.
- 2) *Evaluativa*, que determina el valor de una política pública. Las cuestiones son de valor y el tipo de información que se genera es de carácter evaluativo. Como ejemplo, la evaluación de las diferentes maneras de distribución de la carga fiscal según sus consecuencias éticas y morales.
- 3) *Normativa*, que recomienda futuras acciones que pueden resolver los problemas públicos. Las cuestiones están relacionadas con la acción y el tipo de información que se genera es de recomendación o consejo (*advocative*). Así, se puede recomendar, por ejemplo, una política de ingresos mínimos garantizados como una forma de paliar los problemas de desigualdad social.

Tabla 1.1:
TRES PERSPECTIVAS PARA EL ANÁLISIS DE POLÍTICAS PUBLICAS

<i>Perspectiva</i>	<i>Cuestión principal</i>	<i>Tipo de información</i>
EMPÍRICA	¿Existe? (<i>hechos</i>)	Designativa
EVALUATIVA	¿Qué valor tiene? (<i>valores</i>)	Evaluativa
NORMATIVA	¿Qué debe hacerse? (<i>acciones</i>)	De recomendación (<i>advocative</i>)

Fuente: Dunn, 1981: 37

Además, la producción de información, sea del tipo que sea, se lleva a cabo a través de diferentes técnicas analíticas. Para Dunn, el análisis de políticas utiliza los mismos procedimientos analíticos generales que se utilizan en otras ciencias sociales, mencionando la descripción, la predicción, la evaluación y la prescripción. La descripción y la predicción ayudan a contestar cuestiones designativas, la evaluación las evaluativas y la prescripción las normativas. La predicción y la recomendación se realizan típicamente antes de que se adopte la acción (*ex-ante*), es decir, están relacionadas con el futuro; la descripción y la evaluación lo hacen después (*ex post*) y están relacionadas con el pasado.

Tabla 1.2:
PROCEDIMIENTOS ANALÍTICOS GENERALES SEGÚN EL MOMENTO Y EL TIPO DE CUESTIÓN

Tipo de cuestión/ Momento	<i>Designativa</i>	<i>Evaluativa</i>	<i>De recomendación</i>
Antes de la acción	PREDICCIÓN	EVALUACIÓN	PRESCRIPCIÓN
Después de la acción	DESCRIPCIÓN	EVALUACIÓN	

Nota: La prescripción puede ocurrir después de que la acción se haya llevado a cabo, aunque esta forma retrospectiva de análisis de políticas (“El gobierno debería haber hecho esto o aquello”) es similar a un médico que receta la medicina cuando el paciente ha muerto.

Fuente: Dunn, 1981: 39

Estos procedimientos analíticos generales se corresponden con cuatro métodos para analizar las políticas:

- 1) *Seguimiento* (descripción): produce información sobre las causas y consecuencias pasadas de las políticas.
- 2) *Pronóstico* (predicción): produce información sobre las consecuencias futuras de las políticas.
- 3) *Evaluación* (evaluación): produce información sobre el valor de políticas pasadas o futuras.
- 4) *Recomendación* (prescripción): produce información sobre la probabilidad de que acciones futuras produzcan consecuencias valiosas.

Además hay otros dos métodos que no tienen correspondencia directa con ningún procedimiento analítico general:

5) Estructuración de problemas: basándose en procedimientos de clasificación, produce información especulativa sobre las soluciones potenciales de un problema. Implica todo el resto de métodos, siendo un *metamétodo* que funciona como una especie de regulador central de todo el proceso de análisis.

6) Deducción práctica: permite llegar a conclusiones sobre el grado en el que los problemas se han solucionado.

Estos métodos analíticos están relacionados jerárquicamente y son interdependientes. Algunos, como por ejemplo, el seguimiento se pueden utilizar por sí solos mientras otros, como la evaluación, requieren de la utilización previa de otro método. La recomendación requiere el uso previo del seguimiento, la evaluación y el pronóstico.

Para Dunn, finalmente, esto se combina con cinco tipos de información relevante para las políticas: problemas, alternativas, acciones, resultados y desempeño de la política. El análisis de políticas públicas integrado que él propone es un proceso de indagación que implica esos cinco componentes de información y que se transforman el uno en el otro a través de los seis métodos analíticos (estructuración de problemas, pronóstico, seguimiento, evaluación, recomendación y deducción práctica). La información y los métodos son interdependientes y forman parte de un proceso dinámico en el que ocurren las transformaciones correspondientes: los problemas se transforman en alternativas y éstas en acciones y éstas en resultados, y así sucesivamente.

En esta visión integrada, es lógico que la evaluación tenga un papel acotado y definido. Sin embargo, insistimos en la dificultad de encontrar políticas públicas analizadas de forma integral y continuada o, si se quiere, de la posibilidad de analizar una política integralmente en un momento determinado o por parte de un único equipo de analistas. Esta es una de las razones por las cuales la evaluación necesariamente amplía su visión y, a su vez, ofrece una valoración que asume los diferentes componentes, perspectivas y fases del análisis de políticas públicas.

2. LA EVALUACIÓN DE POLÍTICAS PÚBLICAS

2.1. El concepto y el sentido de la evaluación

A lo largo de los años de desarrollo de la disciplina de la evaluación, se han dado múltiples definiciones del término, dependiendo de la amplitud del enfoque y de cada uno de los autores.

En la definición de Chelimsky (1985), "la evaluación de programas es la aplicación de métodos de investigación sistemáticos para la valoración del diseño, la implementación y la efectividad de un programa". También Weiss (1998), que utiliza el término de investigación evaluativa (*evaluation research*), dice que ésta consiste en la utilización de la metodología propia de las ciencias sociales para hacer que el proceso evaluador sea más exacto y sistemático. En la misma línea, Rossi y Freeman (1989:14) también definen la investigación evaluativa como "la aplicación sistemática de los procedimientos de investigación social, para valorar la conceptualización y el diseño, la ejecución y la utilidad de los programas de intervención social". En estas definiciones se refleja ya su expansión a todo el ciclo de vida del programa o política. Sin embargo, se hace hincapié en la utilización las metodologías de investigación social. En otras palabras, lo que define la evaluación es su objeto de estudio y no tanto sus procedimientos, que serían los mismos que los de la investigación social.

Otra de las definiciones de evaluación, la cual es citada por parte de muchos autores, es la definición adoptada por el *Joint Committee on Standards for Educational Evaluation* (a partir de ahora *Joint Committee*), que como se verá más adelante fue el comité encargado de elaborar el primer código profesional, que se convierte en referencia en el ámbito de la evaluación. Tanto en la primera (1981) como en la segunda edición (1994) de las normas del *Joint Committee*, la evaluación queda definida, y de forma muy amplia, como "la investigación sistemática del valor o mérito de un objeto". En este caso, la definición hace más énfasis en el carácter específico de valoración de la evaluación, es decir, implica que la evaluación siempre supone un juicio. Esto, lógicamente, debe reflejarse en los procedimientos con los que se realice. Por lo tanto, se define diferencialmente la evaluación por sus procedimientos y no por su objeto, que se deja sin acotar.

Creemos que una definición completa del concepto de evaluación debe reflejar ambos extremos, tanto lo referido al **objeto** como lo referido a los **procedimientos**. Argumentaremos más adelante la necesidad también de incluir en el mismo concepto de evaluación cuestiones referidas a las **funciones** de la misma.

En primer lugar, parece claro que si nos referimos a la evaluación de programas o políticas, tenemos ya una primera acotación del **objeto** de estudio. Se evalúan programas, realidad más definida y acotada, y políticas en general (más adelante diferenciamos entre estas dos realidades). En general, se puede hablar de la evaluación de cualquier intervención social que pretenda abordar un problema público, bien sea globalmente (políticas) o bien de forma particular a través de proyectos, programas, planes, medidas, servicios, materiales, etc. Parece claro, además, que las intervenciones (políticas, programas, etc.) no son estáticas y tienen un ciclo de vida, desde su diseño hasta los impactos a largo plazo que puedan producir. Como ya hemos visto, aunque en las primeras etapas del desarrollo de la disciplina de la evaluación se tendía a considerar como objeto de evaluación únicamente los resultados y efectos de los programas, en la actualidad, el objeto de la evaluación de programas se extiende a todo su ciclo de vida.

Respecto a los **procedimientos**, ya hemos mencionado el carácter específico de enjuiciamiento de la evaluación. Es decir, que además de la recopilación y el análisis sistemático de información, se realiza un enjuiciamiento también sistemático. O mejor dicho, que dicha recopilación y análisis de información se realiza con el objetivo de valorar el objeto a evaluar. El carácter sistemático viene determinado al menos por tres factores: en primer lugar, la conciencia de la actividad, es decir, saber que se está evaluando y se quiere evaluar con carácter previo y respondiendo a unas determinadas necesidades de información. En segundo lugar, porque se sigue un determinado orden y método al hacerlo (aquí es donde la evaluación toma prestado elementos de la metodología de investigación social). Y en tercer lugar, porque estructura de forma lógica el camino hacia la valoración, es decir, se recopila información, se interpreta dicha información y se valora con la base de dichas interpretaciones que, a su vez, están basadas en la información recopilada. Patton (1982: 273) señala cuatro niveles de análisis en evaluación, que aparecen en la tabla 1.3.: resultados o hallazgos, interpretaciones, juicios y recomendaciones.

Tabla 1.3. EL ANÁLISIS VALORATIVO EN EVALUACIÓN Y SUS CUATRO NIVELES

Es importante distinguir entre estos cuatro niveles:

RESULTADOS o HALLAZGOS: son los hechos, datos e información que se recoge, la presentación básica de los datos, los hallazgos empíricos.

INTERPRETACIONES: son las explicaciones sobre los hechos, especulaciones entre las interrelaciones, causas, razones y significados que se dan a los datos.

JUICIOS o VALORACIONES: son las valoraciones acerca de los resultados y sus interpretaciones, criterios específicos aplicados a los hallazgos. Estableciendo si son "buenos" o "malos", "positivos" o "negativos", "en la dirección deseada" o "en la no deseada", "por encima, acorde o por debajo de las expectativas", etc.

RECOMENDACIONES: son sugerencias sobre posibles líneas de acción; cambios propuestos en el programa o cosas a mantener en el mismo; consejos para los patrocinadores, responsables y técnicos del programa o política que se evalúa, y que harán que mejore.

(A partir de PATTON, M.Q. 1986, Practical Evaluation, Beverly Hills, Sage)

Según la definición apuntada, se pretende emitir juicios sobre el valor y el mérito del objeto (programas, políticas, etc.) evaluado. Conviene aquí definir y distinguir entre los dos términos señalados: mérito y valor (o valía). Quien primero distingue estos términos es Michael Scriven, cuya aportación teórica al campo de la evaluación es calificada, precisamente, como "la ciencia de la valoración" (Shadish, Cook y Levinton, 1991), pero también se recoge en la obra de Stufflebeam y Shinkfield (1987), y en una propuesta tan diferente como es la de Guba y Lincoln (1989). Para estos autores, el mérito depende de las características intrínsecas del evaluando y tiene que ver con si hace bien lo que se supone que debe hacer, el valor depende de lo extrínseco, del uso que se hace en el contexto. Es decir, mientras el mérito podría ser valorado aisladamente, el valor sólo se podría ser considerado respecto a un contexto y a unas necesidades particulares. Es decir, el evaluando podría ser meritorio aunque no válido (puede que algo sea bueno en sí mismo pero que no sea adecuado para un contexto determinado). Esta circunstancia de estar interesados no sólo en el mérito sino también en el valor de lo que se evalúa abunda en el carácter práctico y aplicado de la evaluación.

Precisamente es este carácter aplicado y práctico, así como el cuarto nivel de análisis señalado por Patton (el de las recomendaciones), lo que recuerda la necesidad de incluir en la definición de evaluación su **porqué** o razón de ser, es decir, sus **funciones**. Efectivamente, la evaluación no se hace en abstracto sino con la intención de mejorar los programas y políticas evaluadas, rendir cuentas y/o ejemplificar o ilustrar acciones futuras (más adelante, tratamos las funciones de la evaluación). Es decir, la información que se pretende arroje la evaluación tiene vocación de ser utilizada para dichas funciones. Esta es la razón por la cual el evaluador extrae también recomendaciones para la acción a la luz de los hallazgos, interpretaciones y juicios establecidos.

Así, el sentido último de la evaluación es su **utilidad** para el proceso de intervención social. Esta utilidad debe ser concebida, además, en dos sentidos. Por un lado, la evaluación debe ser capaz de generar una información útil para el programa o política que esta evaluando, es decir, la información ha de ser pertinente al objeto evaluado y a su contexto determinado. Por otro, debe ser una información con carácter práctico, es decir, debe estar dirigida a la acción, proponiendo cursos de acción futura (que pueden ir desde modificaciones concretas a continuar o terminar el programa o política evaluada). Por esto, creemos que la evaluación puede y debe tener también, en este sentido, un carácter prospectivo. Prácticamente todos los autores de la disciplina coinciden en la consideración de la utilidad de la evaluación como eje de la misma, aunque ponen énfasis en elementos diferentes. Para Monnier (1995) la utilidad social es el elemento clave de la evaluación y aquélla se consigue si se reconoce e implica a los diferentes actores en el proceso evaluativo. Para Vedung (1997), la evaluación debería ser útil porque tiene una orientación claramente práctica: es una valoración destinada a desempeñar un papel en futuras acciones. Para Weiss (1998), la evaluación está concebida para su uso, y eso es lo que la diferencia, precisamente, de otras formas de investigación social. Incluso cuando su utilización es menos directa e inmediata, la utilidad sigue siendo el fundamento de la evaluación. Patton (1997) estructura su enfoque teórico-práctico, precisamente, alrededor de la utilización, realizando una propuesta que tiene en cuenta este aspecto desde el principio de los procesos evaluativos.

2.2. Las funciones de la evaluación

Ya se ha comentado en el punto anterior la necesidad de incluir las funciones de la evaluación en la propia definición del concepto, dado su carácter práctico y aplicado. La elaboración sobre dichas funciones, además, va a contribuir a comprender el carácter específico de la evaluación sobre otro tipo de indagaciones. Las funciones tienen que ver, lógicamente, con los usos que se pretende dar a la evaluación, de manera que la definición de las mismas se encuentra en el mismo corazón del concepto.

Stufflebeam y Shinkfield (1987: 23) señalan tres funciones según el uso que se quiera hacer de la evaluación: perfeccionamiento o mejora (*improvement*), recapitulación¹, rendimiento de cuentas o responsabilidad (*accountability*) y ejemplificación o iluminación para acciones futuras (*enlightenment*). En sus diferentes versiones, creemos que éstas son las tres funciones básicas de la evaluación.

La primera función, la de **mejora** del programa o política evaluada, significa que la evaluación se concibe como un artefacto metodológico que permite la retroalimentación y el aprendizaje sobre la propia práctica. En este sentido, la evaluación resulta imprescindible como instrumento de mejora y perfeccionamiento y como garante de la calidad de un servicio o programa. Lógicamente esta función refuerza la atención al contexto concreto y determinado de lo que se está evaluando, a las necesidades y naturaleza de las personas destinatarias de dicho servicio y de todos los agentes críticos en el mismo, especialmente el personal que lo atiende y la organización que lo promueve.

El **rendimiento de cuentas** o responsabilidad sería la segunda función de la evaluación. La evaluación debe servir como instrumento que permita, en los diferentes niveles de responsabilidad, rendir cuentas sobre la gestión y los resultados de un programa. Esto adquiere una mayor sentido cuando se refiere a programas y políticas públicas, ya que el rendimiento de cuentas debe formar parte indiscutible del sistema democrático. Este rendimiento de cuentas debe darse a varios niveles, incluyendo de forma especial a la ciudadanía: el ciudadano/a tiene derecho a saber no sólo en qué se están empleando los fondos públicos sino también con qué grado de idoneidad, eficacia y eficiencia se están asignando, gestionando y empleando dichos fondos.

Además de estas dos funciones, la evaluación puede cumplir una tercera que es la de dar luz o **iluminar acciones futuras**. Aunque el sentido último de la evaluación tiene que ver, esencialmente, con el programa concreto evaluado en un contexto, situación y momento determinado de forma que se contribuya a su mejora y se rindan cuentas sobre el mismo, las evaluaciones arrojan una información sistemática que contribuye al "acervo científico" no sólo de ese programa, servicio o política concreta sino también del abordaje general de determinados problemas públicos. Como dice Carol Weiss, "el estudio de las políticas y programas en acción es una oportunidad para estudiar los estímulos que conducen al cambio y las condiciones ambientales bajo las cuales es más probable que ocurran cambios" (Weiss, 1998a: 28). Es claro que si se evalúan sistemáticamente los diferentes programas de prevención de drogodependencias, por ejemplo, dichas evaluaciones contribuirán en su conjunto, a construir conocimiento sobre las diferentes formas de hacer prevención, cuáles son los contextos y grupos adecuados, quiénes son los agentes intervinientes necesarios, qué respuestas han dado los poderes públicos en diferentes momentos y por qué, etc. En definitiva, se sabrá más sobre cómo hacer prevención y esto servirá de orientación general a los poderes públicos sobre cómo abordar el problema de las drogodependencias. En este sentido, en la actualidad, en la comunidad científica internacional algunos evaluadores reclaman más experiencias de síntesis de

¹ Aquí respetamos la traducción realizada en la versión castellana del término *accountability* aunque nos parece más adecuada la expresión "rendimiento de cuentas" o incluso de "responsabilidad" como se hace en la traducción de este término en otras obras (Rossi y Freeman, 89; Vedung, 97).

evaluaciones, meta-análisis y metaevaluaciones para aprovechar todo el cúmulo de estudios de evaluación realizados en determinadas áreas en los últimos años (Cook, 1997). Desgraciadamente, esta petición no puede ser trasladada a muchos países como España, en donde lo primero que tiene que ocurrir es que se hagan suficientes estudios como para poder ser sintetizados.

Vedung (1997) también se centra en estas tres funciones básicas de la evaluación: mejora (o perfeccionamiento), responsabilidad (o rendimiento de cuentas) y conocimiento básico (que equivaldría a la función de *enlightenment* propuesta por Stufflebeam y Shinkfield). Dos cuestiones nos parecen especialmente destacables y novedosas en su discurso. Por un lado, el autor parte de la consideración del conocimiento básico como objetivo subordinado a los otros dos. Efectivamente, la mejora de las políticas evaluadas y el rendimiento de cuentas sobre las mismas "destacan como las razones fundamentales para llevar a cabo una evaluación, ya que los conocimientos básicos deberían considerarse más bien como un efecto colateral de los dos anteriores" (Vedung, 1997:131). Esto nos parece un punto especialmente importante en la diferenciación de evaluación e investigación, como se verá más adelante.

En segundo lugar, Vedung analiza con profundidad la función de la **responsabilidad**, cuestión que normalmente se obvia en la literatura al uso. En su análisis, distingue cuatro perspectivas distintas respecto a la responsabilidad. Desde la *perspectiva política*, la evaluación es un instrumento clave para que los representantes políticos comprueben que la administración está realmente ejecutando las tareas que se le han asignado. Desde una *perspectiva técnico-organizativa*, la evaluación es un instrumento para que los directivos de los organismos públicos tengan información sobre si los subordinados están cumpliendo correctamente su tarea. Aunque Vedung reconoce que en el discurso sobre la evaluación rara vez se escucha el punto de vista de los *ciudadanos*, éstos configurarían una tercera perspectiva. La ciudadanía necesita la evaluación para valorar cómo los políticos elegidos y sus agentes a diferentes niveles están desarrollando sus tareas, y para exigirles responsabilidades en las próximas elecciones. Por último, desde la *perspectiva del cliente* (categoría más limitada que la de los ciudadanos), la evaluación puede producir información sobre si los servicios prestados son satisfactorios y se están distribuyendo equitativamente. Estas cuatro perspectivas hacen que se den dobles lealtades en la cadena de representación pública (los representantes políticos son a la vez superiores y subordinados, la burocracia nacional es agente del gobierno y "jefe" de otros organismos públicos, etc.), siendo evaluados los que a su vez tienen que evaluar, lo que contribuye, según el autor, a la tibieza en el interés mostrado por la evaluación en la vida pública.

2.3. Propósitos y objetivos de la evaluación de las políticas públicas

En la literatura sobre evaluación no se suele distinguir entre cuáles son las funciones y cuáles los propósitos y objetivos de la misma. Nosotros, sin embargo, queremos distinguir entre funciones básicas de la evaluación y objetivos o propósitos, basándonos en la distinción de los dos conceptos. Según la Real Academia de la Lengua Española la **función** está definida como "capacidad de acción o acción de un ser apropiada a su condición natural (para lo que existe) o al destino dado por el hombre (para lo que se usa)". Claramente, las funciones de la evaluación han sido definidas por el destino que se le ha dado a tal actividad (para lo que se usa) en términos genéricos, de forma que pueden quedar definidas como su condición natural (para lo que existe). Nosotros queremos recalcar este factor de condición natural, es decir, la evaluación existe para mejorar, rendir cuentas y/o iluminar acciones futuras. Si cumple una función distinta, no es evaluación, es otra cosa.

Sin embargo, los **objetivos** ("fin o intento a que se dirige o encamina una acción u operación"), **finés** ("motivo con que se ejecuta una cosa") y **propósitos** ("ánimo o intención de hacer o de no hacer una cosa"), se distinguen de las funciones en que pueden ser mucho más variados y no definen en sí mismos, la naturaleza, en este caso, de la evaluación.

Carol Weiss (1998a: 25) señala que, tradicionalmente, la evaluación ha hecho hincapié en el propósito de ofrecer información para **facilitar el proceso de toma de decisiones**. El tipo de decisiones pueden ser:

- hacer correcciones intermedias en el proceso de ejecución de un programa
- continuar, expandir o institucionalizar un programa o cortarlo, acabarlo o abandonarlo
- examinar nuevas ideas y/o programas
- elegir la mejor entre varias alternativas
- decidir si se continúa con la financiación

En esta misma línea, Chelimsky (1987: 74-75) liga los propósitos de la evaluación al ciclo de las políticas públicas y a los procesos de decisión en cada una de esas fases. Estos diferentes propósitos tendrán implicaciones en los tipos de evaluación que deberán aplicarse para cubrir las diferentes necesidades de información, como se verá más adelante. Ella habla, entonces, de tres grandes propósitos de la evaluación, que son:

- Para la *formulación de políticas*, esto es, por ejemplo, para valorar y/o justificar la necesidad de un nuevo programa. Requiere información de la evaluación de, al menos, tres grandes áreas: 1) información del problema que se trata, 2) información sobre anteriores programas y esfuerzos similares o relacionados, e 3) información que permita la selección de un programa alternativo sobre otro.

- Para la *ejecución de políticas*, esto es, por ejemplo, para asegurar que el programa se está ejecutando de la forma más efectiva y eficiente. Se aplica igualmente a nuevos programas o programas ya existentes. La información que se requiere para este propósito tiene que ver con: 1) información sobre la ejecución o implementación del programa (si es operativo, cuánto cuesta, si es similar en las diferentes localizaciones, si se ajusta a lo formulado, si hay problemas en la puesta en práctica, etc.) , 2) información sobre la gestión del programa (el grado de control del gasto, la cualificación y credenciales del personal, la distribución de recursos, el uso de la información del programa en el proceso de toma de decisiones, etc.), e 3) información sobre el estado actual del problema abordado en el programa (si está creciendo o aumentando, si está cambiando, etc.).

- Para la *responsabilidad o rendimiento de cuentas (accountability)*, esto es, por ejemplo, para determinar la efectividad de un programa y la necesidad de su continuación, modificación o terminación. También se aplica a programas nuevos y ya existentes y requiere información sobre: 1) información sobre los resultados del programa, 2) información sobre si el programa ha logrado sus objetivos, e 3) información sobre los efectos (esperados e inesperados) del programa.

Sin embargo, en los últimos años ha aumentado la conciencia de que los resultados de una evaluación no determinan, por sí solos, las decisiones acerca de un programa, servicio o política. En muchas ocasiones, las organizaciones y agentes en torno a un programa o política no aplican de forma directa e inmediata los resultados de una evaluación. Más frecuentemente aprenden de la evaluación y este aprendizaje afecta a las decisiones de forma sutil y gradual. Por esta razón, en los últimos tiempos se ha ido reconociendo el importante papel que puede tener la evaluación como **facilitadora del aprendizaje organizativo** (Weiss, 1998a: 26). En este caso, los propósitos están menos orientados a decisiones concretas sobre los programas o políticas evaluadas, como son:

- registrar la historia del programa o política
- ofrecer *feedback* o información sobre su práctica a los profesionales que intervengan en dichas políticas o programas
- enfatizar o reforzar los objetivos de un programa
- rendir cuentas (la cultura del rendimiento de cuentas puede hacer a los profesionales de una organización estar más atentos a las expectativas del público)
- entender (saber más sobre) la intervención social

Además de estos nobles objetivos, la literatura sobre evaluación, en su evolución hacia la incorporación de la práctica a la teoría evaluativa y su afán realista, ha ido también identificando otros posibles propósitos con los que se llevan a cabo ciertos estudios de evaluación. Por ejemplo, en algunos casos se pueden realizar evaluaciones **a requerimiento de la subvención, becas o proyectos** obtenidos, no tanto porque exista una conciencia sobre la conveniencia o necesidad de la evaluación; en estos casos, la evaluación puede convertirse en un mero requisito formal.

Carol Weiss (1998a) señala que los evaluadores deben ser conscientes de la existencia de **propósitos encubiertos** y Vedung (1997) utiliza el término **propósitos "estratégicos"** ya que normalmente están orientados a un determinado plan (aunque la mayoría de las veces permanecen encubiertos); en su identificación se basa, en parte, en un listado de **abusos evaluativos** identificado por Suchman (67). Stufflebeam y Shinkfield (1987) hablan de **pseudoevaluaciones**. Estos propósitos se podrían resumir en los siguientes:

- **Aplazamiento.** Se trata de llevar a cabo una evaluación con el propósito de demorar una decisión. En algunos casos se puede esperar, incluso, que una determinada cuestión candente se vaya debilitando con el paso del tiempo y disminuya la atención y presión pública. La evaluación será, en cualquier caso, un elemento tranquilizador.
- **Evasión y/o evitación de responsabilidades.** En este caso las evaluaciones se realizan para obligar a los oponentes a compartir responsabilidades por decisiones impopulares. O puede ser que se encargue un estudio de evaluación para evitar toda responsabilidad, pretendiendo que la evaluación sea la que adopte las decisiones.
- **Legitimación de una decisión ya tomada.** La evaluación puede ser un medio para legitimar una decisión ya tomada. Weiss (1998a: 22) utiliza la expresión "*window dressing*" ("desnudarse delante de la ventana") para referirse a este propósito.
- **Relaciones Públicas.** Las evaluaciones también pueden llevarse a cabo por pose o imagen, o como un medio de hacer visible o darle publicidad al programa o política evaluada. Se puede utilizar la evaluación como un gesto de racionalidad y asumiendo una posición de investigación objetiva y científica, ofreciendo así una buena apariencia externa.
- **Ofrecimiento de una imagen determinada.** En el propósito anterior, el mero hecho de encargar una evaluación es suficiente. En este, la evaluación se centra en una selección sesgada y parcial de sus resultados sustantivos. Pueden darse dos movimientos relacionados: uno, la justificación de un programa débil a través de la selección de aquellos aspectos positivos (*Eyewash* para Suchman, 1967); otro, el encubrimiento de un fallo o fracaso evitando la información relativa al mismo (*whitewash*).

La postura de los diferentes autores ante estos últimos propósitos difiere ligeramente. Vedung (1997) parece más comprensivo con estos propósitos denominándolos "estratégicos", ya que responden a un determinado plan o estrategia. En el caso de Suchman (su publicación data del año 1967 y el campo de la evaluación estaba muy poco desarrollado) y el de Stufflebeam y Shinkfield (1987) su posición es claramente condenatoria. Weiss (1998a) también habla de estos propósitos como razones "no tan legítimas" de llevar a cabo evaluaciones y advierte a los evaluadores de las mismas. Su posición al respecto se ilustra en la siguiente cita:

"La evaluación es, entonces, una empresa racional que a menudo se lleva a cabo por razones no racionales o, al menos, no informativas (...), pero lo

importante es que tales motivos tienen consecuencias en la evaluación que pueden ser serias y nada prometedoras. (...) La gracia de las evaluaciones es que a menudo se comprueba que son útiles aún cuando se comenzaron sin ninguna intención de utilizar los resultados. (...) El potencial de ejemplificación general de la evaluación ayuda a sobreponer algunos de las señales no favorables emitidas por los patrocinadores desinteresados." (Weiss, 1998a: 23-24)

En general, pues, se señala la importancia de reconocer la existencia de estos “otros” propósitos de las evaluaciones. Tanto evaluadores como analistas deben estar al tanto de estos posibles propósitos y saberlos reconocer y analizar. Además, los evaluadores deben señalar y hacer explícitos estos propósitos en caso de que haya indicios de los mismos, y clarificar y “negociar” con las personas que encargan la evaluación los propósitos reales y legítimos de un proceso de evaluación. Ante situaciones extremas, los evaluadores deben saber identificar la situación evaluativa de la forma más realista posible, con la intención de poder decidir y si viene al caso (apoyándose en criterios deontológicos éticos y profesionales), no emprender la evaluación encargada.

2.4. Diferencias y similitudes entre la evaluación de programas y de políticas públicas

La literatura en evaluación habla en general de programas, servicios y políticas públicas, no distinguiendo, en la mayoría de los casos, procedimientos distintos para objetos diferentes. Es más, se suele hablar de un objeto genérico de la evaluación, al que suele asignarse el concepto de “programa”, pero que pretende designar, en general, políticas, planes, programas, proyectos, servicios, medidas y materiales, dejando fuera la evaluación institucional y de personal (*Joint Committee*, 1994). En otros casos el campo académico y profesional de la evaluación pretende abarcar aún más objetos: tomamos como ejemplo los Principios Orientadores para Evaluadores elaborados por la Asociación Americana de Evaluación (AEA, 1994), en los que se explicita que se aplicarían, además de a los campos mencionados anteriormente, a la evaluación institucional, de personal, del desempeño, de la tecnología, de proyectos de investigación, de teorías, etc. De hecho, a lo largo de este trabajo se utiliza el término evaluación de programas y políticas en sentido amplio. Sin embargo, nos parece conveniente realizar una serie de distinciones conceptuales y terminológicas en este sentido.

Una primera distinción sería la relacionada con el **objeto** de la evaluación. Sí parece evidente que es diferente evaluar programas que evaluar políticas públicas; y efectivamente, existe una diferencia de nivel. Algunos autores mencionan que existen diferentes **niveles** de evaluación, aunque esta categorización analítica no es frecuente en la literatura de evaluación y, cuando aparece, no se le da una especial importancia.

Así Carol Weiss (Weiss, 1998a: 7), habla de programas nacionales (pone como ejemplo el *Head Start*), de proyectos locales (pone como ejemplo el programa concreto de *Head Start* que se lleva a cabo en el Centro Comunitario de Brigham), y de componentes de los proyectos (pone como ejemplo un elemento del proyecto local mencionado como es la implicación de los padres a través de encuentros semanales). También Weiss habla de que se pueden evaluar políticas, aunque la definición que ofrece de las mismas sea algo limitada (“una política es una declaración de objetivos, aceptada oficialmente, ligada a una serie de actividades que se supone cumplirán dichos objetivos en una determinada jurisdicción”). Sin embargo, la conclusión de Weiss al respecto es que la mayoría de las estrategias evaluativas que se usan en evaluación de políticas son las mismas que en la evaluación de programas.

Sin embargo, nosotros creemos importante hacer un esfuerzo de distinción conceptual entre lo que es evaluación de políticas y evaluación de programas, aunque estemos en general de acuerdo con Weiss en que la mayoría de las estrategias evaluativas son las mismas. En la tabla siguiente se exponen las principales diferencias en una serie de dimensiones. Según el *objeto* de estudio, por evaluación de programas se entiende, en sentido estricto, la evaluación de todo tipo de unidades: servicios, medidas, proyectos, programas y planes. Estas unidades se caracterizan, con un mayor o menor grado de amplitud y generalidad, por ser realidades acotadas, con unos objetivos, unas actividades, unos recursos y un calendario más o menos definidos. Sin embargo, las políticas públicas son unidades menos definidas “a priori”, realidades más “amorfas”, y consisten en procesos, decisiones y cursos de acción en los que intervienen los poderes públicos para abordar un determinado problema público en un cierto período de tiempo. El término de políticas públicas ha sido utilizado de muy diferentes maneras (Hogwood y Gunn, 1984: 13-19) e incluso “una política puede consistir en no hacer nada” (Hecló, 1972). Son más difícilmente acotables y requieren de la redefinición y reconstrucción previa por parte del analista, ya que “toda política pública es definida subjetivamente por el observador, comprendiendo normalmente un conjunto de decisiones relacionadas con una variedad de circunstancias, personas, grupos y organizaciones” (Subirats, 1989: 42). Necesariamente, esta diferencia en el objeto debe marcar, de alguna manera, distintas formas de abordar las evaluaciones. Entre otras, una mayor probabilidad de que, para la evaluación de políticas públicas, el *agente evaluador* sea un analista de políticas públicas, y para la evaluación de programas, sean bien consultores externos (no necesariamente analistas), bien responsables internos de dichos programas.

Respecto a la *aplicación*, mientras la evaluación de programas es probable que enfatice la mejora y el control de responsabilidades internas, la evaluación de políticas puede que lo haga haciendo hincapié en el conocimiento general o iluminación para acciones futuras, o en una información general que permita tomar decisiones fundamentadas sobre la continuación o terminación de tal política o programa. Por último respecto a la aplicación, la evaluación de programas definitivamente es la base y contribuye a la evaluación de las políticas públicas.

Sin embargo, la *metodología* utilizada es esencialmente la misma, aunque se puede decir que la evaluación de políticas siempre se apoya más en el análisis secundario de datos, entre otros, de la información generada por los propios programas. Finalmente, el calendario de la evaluación o la *temporalidad*, en general, suele ser más

ajustada en la evaluación de programas ya que los ciclos temporales suelen ser más cortos y estar más definidos.

Tabla 1.4
DIFERENCIAS ENTRE LA EVALUACION DE PROGRAMAS
Y LA EVALUACION DE POLITICAS

Dimensiones	EVALUACION DE PROGRAMAS	EVALUACION DE POLITICAS
OBJETO DE ESTUDIO	Planes, programas, proyectos Servicios, medidas	Políticas Públicas
AGENTE EVALUADOR	Consultores externos Responsables internos	Analistas
APLICACION	Mejora del programa o servicio Control de responsabilidades Contribución a la evaluación de una política	Iluminación para acciones futuras Ayuda a las decisiones sobre continuación o terminación
METODOLOGIA	Flexibilidad metodológica	Flexibilidad metodológica (relativo mayor apoyo en análisis secundario de datos)
TIEMPOS	Temporalidad ajustada	Puede tomarse más tiempo aunque debe reflejar el carácter dinámico de la política
EJEMPLOS	Programa “Deporte y droga” Red de Albergues Juveniles	Política de Juventud en la Comunidad de Madrid

3. LO ESPECÍFICO DE LA EVALUACIÓN

3.1. Las diferencias con otras prácticas de análisis: evaluación *versus* investigación

Durante los primeros años del desarrollo de la disciplina de la evaluación de programas en los años sesenta y setenta, no se encuentran en la literatura especializada distinciones entre investigación y evaluación. Entonces, se concebía la evaluación como un tipo de investigación de carácter aplicado y que debe guiarse por los estándares de la misma. En realidad, se consideraba la evaluación como una hermana “menor” de la investigación, siendo a veces desprestigiada en la comunidad científica de entonces. La utilización de términos como “investigación evaluativa” (Rossi y Freeman, 1989; Weiss, 1975, etc.), aunque contribuían a dar una mayor “seriedad” y a obtener legitimidad para el nuevo campo, no ayudaban al esclarecimiento o diferenciación de los dos conceptos.

Respecto a la **naturaleza** de la investigación y de la evaluación, es evidente que ambas son modos de indagación sistemática. Esto implica que, como requisito mínimo, son actividades que se hacen de forma consciente y con un sistema, es decir, siguiendo unos pasos o aplicando algún método o forma sistemática de actuación. Además, para Lincoln y Guba (1986: 547) se trata de tipos de indagación “disciplinadas”. Por indagación disciplinada (*disciplined inquiry*), estos autores entienden algo que es confirmable públicamente y que permite explicitar la naturaleza de los datos, las fuentes y el contexto en el que fueron recogidos, así como los procesos de transformación de los datos en información (interpretaciones, conclusiones, extrapolaciones, recomendaciones).

A caballo entre el criterio de la **naturaleza** de cada actividad y el criterio de su **finalidad**, estaría su posible carácter básico o aplicado. Ambos tipos de actividad – investigación y evaluación– pueden ser de carácter aplicado; pero la evaluación es siempre de carácter aplicado, nunca básico. La investigación, por el contrario, puede ser tanto de carácter básico como aplicado. Sin embargo, el hecho de que se dé una investigación aplicada (cuyo fin primario sea la solución de un problema o la respuesta a una necesidad concreta, más que el desarrollo de teoría en abstracto) no la convierte automáticamente en una actividad evaluativa. Tampoco el hecho de ser aplicada no significa que sea una investigación de categoría inferior; muy al contrario, el carácter aplicado de la investigación dota de contenido y utilidad social a la actividad científica. Como señalan Lincoln y Guba “es posible afirmar que muchos avances *básicos* en las ciencias han sido ocasionados por esfuerzos para resolver problemas *prácticos*” (1986: 548). Lo que sí parece claro es, como se ha dicho antes, que la evaluación es siempre de carácter aplicado ya que se hace para valorar el mérito y/o valor de un evaluando en su

contexto, para resolver determinadas necesidades de información y con la intención de mejorar el evaluando, rendir cuentas sobre el mismo o ilustrar sobre acciones futuras. Es más, se puede afirmar que la evaluación es un tipo de indagación sistemática *especialmente* aplicada.

El criterio de la diferente **finalidad** de la investigación y la evaluación es el que proporciona una base más amplia para su distinción. En términos generales, la investigación se hace con el fin último de construir conocimiento. La evaluación, sin embargo, se hace para mejorar los programas evaluados, rendir cuentas sobre los mismos y generar una información que permita ilustrar posibles acciones futuras. Es cierto que, de alguna manera, la evaluación también construye conocimiento, pero éste es siempre de carácter más aplicado o concreto. La realización de evaluaciones sistemáticas en un determinado sector o de un determinado tipo de políticas (por ejemplo, de cooperación al desarrollo) es evidente que contribuye a la acumulación de conocimiento sobre el problema que se trata de resolver y la forma de hacerlo. Pero también es claro que ésta no es su función principal. La evaluación nunca se hace con el fin de comprobar proposiciones teóricas, aunque sí sea, en cierto modo, una aplicación de la teoría ya que las evaluaciones siempre tienen como telón de fondo y referencia, además del contexto específico del evaluando, el marco teórico en torno al tipo de intervención evaluado. Esta diferencia de propósito es la primordial a la hora de distinguir entre investigación y evaluación y la que hace que los dos tipos de indagación sean distintos respecto a otras dimensiones y características. Así, la investigación es un modo de indagación de carácter académico, mientras la evaluación lo es de carácter práctico (Cordray y Lipsey, 1987). Esto quiere decir que la evaluación de programas pone énfasis en la valoración y el enjuiciamiento del evaluando, su significado para los destinatarios, patrocinadores y responsables, en ser relevante y sensible ante el contexto y, sobre todo, en ofrecer una información oportuna en el tiempo y *útil* para la toma de decisiones relacionada con dicho evaluando.

Por lo tanto la *utilidad* es el sentido último o *leit-motiv* de la evaluación de programas. La evaluación se hace con la intención de ser utilizada. “En su forma más simple, la evaluación se lleva a cabo por parte de un cliente que tiene decisiones que tomar y que busca en una evaluación una información en la que basar sus decisiones. Incluso cuando el uso es menos directo e inmediato, la utilidad de algún tipo es la fundamentación de la evaluación” (Weiss, 1998: 15).

El hecho de que la utilidad sea una característica central de la evaluación implica una serie de condicionantes que deben darse en este tipo de actividad científica. La existencia de un cliente o clientes que han demandado la evaluación y la necesidad de responder ante un contexto específico hacen que el evaluador, además de la *lealtad* hacia su campo de estudio y como científico social, tenga, sobre todo, obligaciones con el cliente, los formuladores de la política, los directivos, los profesionales y los participantes y beneficiarios del programa evaluado. Esto significa que la evaluación debe tener mayor *capacidad de respuesta* que la investigación, especialmente ante el contexto (aunque la investigación debe tener también en cuenta el contexto, para la evaluación la respuesta al mismo es crucial) y ante los responsables y agentes críticos de los programas.

En conexión con esta capacidad de respuesta y, en general, con la utilidad como sentido último de la evaluación, está la necesaria *oportunidad en el tiempo* que debe tener la evaluación. Los ritmos de la investigación pueden ser distintos, mucho más lentos. Sin embargo, para la evaluación, la capacidad de respuesta está claramente ligada al calendario. La utilidad de la información que arroja una evaluación depende, en gran medida, de que ésta llegue a tiempo para la toma de decisiones sobre las mejoras, las posibles modificaciones y el rendimiento de cuentas. La entrega tardía de un informe de evaluación puede significar la práctica inutilidad del mismo. Eleanor Chelimsky reflexiona sobre las aportaciones que se pueden hacer a la teoría de la evaluación desde la experiencia de la práctica evaluativa, y es precisamente éste uno de los puntos sobre el que llama la atención; para ella, a la necesidad de oportunidad en el tiempo de la evaluación no se le ha dado la suficiente importancia en el nivel teórico (Chelimsky, 1998). Cuanto más inestable e incierto políticamente sea el contexto en el que se encuentra el programa o política a evaluar, más acuciante es la necesidad de que la evaluación cumpla los criterios de temporalidad. Si se tarda mucho tiempo en contestar las preguntas de evaluación, por muy bien contestadas que estén, es posible que el debate que las generó ya no sea el mismo, que se hayan quedado obsoletas y sean irrelevantes o que los destinatarios a los que se había dirigido la evaluación hayan desaparecido. Chelimsky también señala que el incumplimiento de plazos en un proceso de evaluación suele dañar la reputación y credibilidad de los evaluadores a largo plazo; además del incumplimiento de compromisos, los retrasos suelen entenderse como señal de incompetencia o de que hay algo que ocultar (Chelimsky, 1998: 41).

La investigación y la evaluación se suelen producir en **contextos** distintos y que pueden ser caracterizados de forma diferente, según se trate de una investigación o de una evaluación. Algunos autores como Weiss (1998), quien considera que la evaluación se desarrolla en un escenario de acción, han utilizado este criterio para distinguir entre los dos tipos de actividad. La evaluación se desarrolla en un contexto específico que es al que dicha evaluación debe responder y que debe orientar y “enfocar” el estudio de evaluación. Además, los procesos de evaluación siempre se llevan a cabo en un escenario de acción, en el que lo más importante que ocurre es el programa. Weiss insiste en que en este tipo de escenarios es probable que se den conflictos entre el programa y la evaluación del mismo; lógicamente, cuando esto ocurre, siempre se da prioridad al programa y no a la evaluación (Weiss, 1998: 15). Al fin y al cabo, si la evaluación se realiza para mejorar el programa, es lógico que aquella dependa de y se subordine a éste. Por lo tanto, los procesos de evaluación dependen más del contexto que los de investigación. Esto, en términos prácticos, significa que las estrategias de selección, recopilación y análisis de la información también se deben plantear de forma diferente en los dos tipos de actividad, siendo más cautelosos e inclusivos (ofreciendo una mayor participación a responsables y a agentes críticos) en los procesos de evaluación.

Esta mayor dependencia del contexto de la evaluación está íntimamente ligada al necesario reconocimiento del carácter político de los procesos evaluativos. Esta es otra diferencia clara con la investigación, ya que esta última se ve mucho menos afectada por el contexto político.

De forma especial, las diferencias en el **objeto** de estudio han sido poco tratadas en la literatura sobre evaluación, aunque desde nuestro punto de vista, también se

pueden identificar algunas diferencias. La diferencia básica inicial estriba en que la investigación suele tener por objeto el estudio del *problema* que los programas intentan abordar, es decir *qué problema público se intenta solucionar y sus características* (por ejemplo, la desigualdad de oportunidades entre mujeres y hombres) y la evaluación, sin embargo, se dedica al estudio de las *intervenciones públicas* que intentan solucionar dicho problema, es decir, *cómo se aborda el problema* en un contexto y momento determinado (por ejemplo, la política de igualdad de oportunidades entre mujeres y hombres). Afinando algo más en el análisis, mientras la investigación puede dedicarse a la evolución del problema en abstracto, la evaluación pretende valorar las intervenciones para abordar dicho problema y sus efectos en esa evolución, o, dicho de otra manera, cómo ha ido evolucionando el problema por efecto de la intervención a través de un programa. Esto, en un sentido estricto, es a lo que se llama el “efecto neto” de un programa o política, es decir, lo que ha producido un programa y que no se hubiera producido en ausencia del mismo. Esta diferenciación es importante porque puede dar lugar a equívocos. Por ejemplo, el estudio de la situación social de la mujer y su evolución durante tres años en un determinado territorio no es equivalente a la evaluación del Plan de Igualdad de Oportunidades entre Mujeres y Hombres que se haya ejecutado en ese territorio durante el mismo período. Es decir, una cosa es el análisis periódico de la situación de la mujer, y otra muy diferente, la valoración de las acciones que se han llevado a cabo por parte de los poderes públicos para paliar las situaciones de desigualdad de género.

Además de la valoración de este efecto neto, que en muchas ocasiones es sumamente difícil y costoso de medir, la evaluación de la forma en la que se aborda un problema público debe incluir la valoración del diseño de la intervención, su puesta en práctica y sus resultados. Es decir, *el foco de la evaluación es la valoración de las acciones e intervenciones concretas que se realizan para abordar un problema determinado, en todas sus dimensiones*. Stufflebeam y Shinkfield (1987: 23) consideran que, mientras la investigación se limita al estudio de ciertas variables que tienen un interés teórico, la evaluación se dedica a todos los criterios que se utilizan en la determinación de un valor. Esto significa que la investigación puede centrarse en ciertos aspectos concretos o en algunas variables específicas; es decir, su foco de atención puede ser pequeño o más acotado. Por el contrario, la evaluación, por el hecho de determinar el valor de una intervención, suele tender a considerar ésta de forma global, más que centrarse en algunos aspectos concretos de dicha intervención.

Sin embargo, aunque existen diferencias genéricas entre el objeto de estudio de la investigación y la evaluación y el reconocimiento de las mismas es también importante para establecer una identidad para la evaluación como actividad científica diferenciada, no basta con distinguir los objetos en términos abstractos. Por un lado, es evidente que el abanico de objetos de investigación (cualquier tipo de problema) es mucho más amplio que el de la evaluación, que está mucho más acotado (abordaje de un problema a través de algún tipo de intervención). Pero por otro lado, para diferenciar entre objetos, puede hacer falta el cruce del propósito con el que se lleva a cabo la indagación. Así, el objeto de la evaluación puede ser “invadido” por la investigación, si ésta se lleva a cabo con un propósito diferente al de la evaluación de programas. Por ejemplo, el estudio comparado de las políticas de cooperación al desarrollo en tres países, aunque se trata de un objeto que tiene que ver con el abordaje de un problema, más que con el problema en sí mismo, sería más investigación que evaluación porque,

seguramente, no se llevaría a cabo con la intención de mejorar las políticas o controlar responsabilidades, sino con un propósito más general de construcción de conocimiento.

Como ya se ha mencionado, los criterios más utilizados para resaltar diferencias y semejanzas entre investigación y evaluación, junto con los de finalidad, son los de tipo **epistemológico** y **metodológico**. Algunas de las primeras distinciones que se encuentran en la literatura, tendían a pensar en la investigación en términos positivistas, dejándole a la evaluación mayores “grados de libertad” metodológicos. “Las evaluaciones suelen utilizar métodos subjetivos y no están tan estrechamente controladas ni sujetas a manipulación como las investigaciones” (Stufflebeam y Shinkfield, 1987: 23). Incluso se llega a identificar a cada tipo de actividad con un paradigma diferente desde el punto de vista epistemológico. Para Cordray y Lipsey (1987: 29), la evaluación se inscribiría en una perspectiva más constructivista (examinando diferentes ‘hechos’, vistos desde muchos puntos de vista) mientras la investigación lo haría en una más positivista (examinando un conjunto de ‘hechos’ sobre los aspectos tangibles de una intervención).

En una línea de inclusión de todas las perspectivas paradigmáticas, distinguen investigación y evaluación Lincoln y Guba (1986). Estos autores proponen una definición de investigación que sea válida para cualquier orientación epistemológica y para cualquier tipo de investigación (básica o aplicada): “es un tipo de indagación disciplinada que se lleva a cabo para resolver algún problema de forma que se consiga entendimiento o se facilite la acción”. Sin embargo, “la evaluación es un tipo de indagación disciplinada que se lleva a cabo para determinar el valor (mérito y/o valía) de alguna entidad -el evaluando- tal como un tratamiento, un programa, unos recursos o una actuación, de forma que se mejore o refine el evaluando (evaluación formativa) o se valore su impacto (evaluación sumativa)” (Lincoln y Guba, 1986: 549-550). Como se puede observar, estas definiciones vuelven a hacer hincapié en las diferentes finalidades.

Una de las semejanzas que siempre se aducen al hablar de investigación y evaluación es la utilización de la misma **metodología** (Weiss, 1998). Es más, en ocasiones se ha definido la evaluación como la utilización de los métodos de la investigación para valorar programas. Para Weiss, la utilización de los métodos de las ciencias sociales tiene como fin hacer que el proceso valorativo sea más exacto y sistemático. De nuevo, como muchos otros autores, Weiss señala que lo que distingue la investigación evaluativa no es el método o la materia de la que se ocupa, sino la intención o el propósito para la que se hace (Weiss, 1998: 15). Como ya se ha visto, ni siquiera Lincoln y Guba, quienes enmarcan su discusión con pautas de carácter epistemológico, insisten en las diferencias metodológicas, sino que se centran en las diferencias de propósito, y mencionan algunas cuestiones respecto a los destinatarios y los resultados o productos. Más adelante, se argumenta que también existen diferencias de método, aunque ambos tipos de actividad utilicen las mismas técnicas de recopilación y análisis de información e incluso las mismas metodologías en ocasiones. De esta manera, hacemos eco del símil utilizado por Lincoln y Guba: el hecho de que utilicen las mismas herramientas (martillos, destornilladores, sierras y llaves) no hace iguales a carpinteros, electricistas y fontaneros (Lincoln y Guba, 1986: 537).

Carol H. Weiss, además de la utilización de las mismas técnicas y métodos, e incluso los mismos diseños de investigación, menciona otra semejanza relacionada con el criterio metodológico. Para Weiss, la evaluación pretende, como cualquier otra investigación, describir, entender relaciones entre variables y seguir la secuencia causal de una variable con otra. Sin embargo, para la autora, no siempre se pueden aplicar los métodos con el mismo rigor que en la investigación ya que las dependencias del tiempo, el espacio, la gente y los recursos disponibles son más estrechas y evidentes en situaciones de evaluación. Así, "los métodos de evaluación a menudo representan un término medio entre lo ideal y lo posible" (Weiss, 1998: 18). En definitiva, lo que las diferencia, pues, no es la metodología en sí misma sino las circunstancias y condicionantes con los que se aplica.

Según nuestro punto de vista, la idea de "metodología" de Weiss tiene que ver con una visión tradicional y acotada del método y que se limita a la forma en la que se recopila y analiza la información. Esto también tiene que ver con la necesaria diferencia que se debe establecer entre los niveles metodológico y tecnológico (referido a las técnicas para la recopilación y análisis de la información). La metodología se refiere al razonamiento lógico que vincula la información que se recopila al conocimiento general. Las técnicas son las que permiten recopilar la información. Como veremos más adelante, si consideramos que la metodología de evaluación comprende algo más que la recopilación y el análisis de la información, se pueden también encontrar diferencias de método entre los dos tipos de indagación.

En lo que se refiere a la distinción entre la metodología de investigación y la de evaluación, es interesante insistir en la *diferencia entre los niveles metodológico y tecnológico*. De esta manera, la utilización de las mismas técnicas de recopilación y análisis de información no significa la utilización de una misma metodología. Como se ha comentado antes, Guba (1990) dice que la *metodología* se refiere a la forma en la que se busca el conocimiento; nosotros añadiríamos en que se *busca* y se *ofrece* el conocimiento. Así, creemos que cuando se habla de metodología de evaluación estamos hablando de algo más que la forma en la que se recopila y analiza la información; hablamos, al menos, de 1) cómo se selecciona la información que se quiere recopilar, teniendo en cuenta y respondiendo al contexto específico del programa que se pretende evaluar (diagnóstico de la situación evaluativa, diseño de la estrategia general de evaluación y establecimiento y definición de los criterios o preguntas de evaluación), 2) cómo se recopila y trata la información que nos va a permitir contestarnos a las preguntas de evaluación, y 3) cómo se valora dicha información y, por lo tanto, se juzga el mérito y/o valor del evaluando. En una visión amplia de la metodología podríamos añadir, además, otros dos pasos: 4) cómo se traducen dichas valoraciones en recomendaciones para la acción y 5) cómo se transmiten y comunican los resultados de la evaluación para que éstos sean lo más útiles y utilizados. Aunque estos dos últimos pasos pueden ser también considerados como resultados esperados de los procesos de evaluación. Por lo tanto, hay varias características que distinguen la metodología de evaluación de la de investigación y la hacen, de alguna forma, específica.

En primer lugar, esta especificidad tiene que ver con la definición del problema a resolver y con el primer paso sobre cómo se **seleccionan los criterios o preguntas de evaluación**. Weiss (1998: 15) señala que, en evaluación, las preguntas que se plantean se derivan siempre del propio programa a evaluar. Independientemente de que se tenga

en cuenta el estado de la cuestión, las fuentes más importantes para la elaboración de los criterios o preguntas de evaluación son los objetivos del programa a evaluar y, sobre todo, las preocupaciones de los agentes críticos (de la “comunidad de política o programa”). Mientras los investigadores pueden formular sus propias hipótesis, el evaluador debe convertir en preguntas de evaluación las preocupaciones del programa. Aunque el evaluador pueda influir en lo que él considera importante, el corazón de los estudios de evaluación son las cuestiones de interés programático.

En segundo lugar, en evaluación siempre se da un elemento de **valoración**. Para Weiss (1998:15) una de las diferencias entre investigación y evaluación es lo que ella llama la “calidad de juicio”. En evaluación se compara entre lo que es y lo que debería ser; se compara respecto a normas implícitas o explícitas, relativas o absolutas. El elemento de valoración respecto a una serie de criterios es lo que diferencia la evaluación respecto a otros tipos de investigación. También para Alvira, aunque los procesos de evaluación son muy similares a los de una investigación aplicada, la característica que principalmente diferencia a los dos tipos de actividad es que en evaluación se deben emitir juicios de valor sobre el objeto evaluado (Alvira, 1991: 7). Este debe ser siempre un enjuiciamiento sistemático, que tenga en cuenta los criterios y normas previamente establecidos y que se base, sobre todo, en los datos recopilados y en las interpretaciones que de ellos se han hecho.

Sin embargo, investigación y evaluación no se diferencian por las **técnicas de recopilación y análisis de información** que son, básicamente, las mismas. Aunque, afinando en el análisis, se puede señalar que, en general, las evaluaciones tienden a utilizar técnicas menos sofisticadas que las investigaciones. Las constricciones de tiempo, recursos y contexto que suelen darse en las situaciones de evaluación hacen que, en muchas ocasiones, la utilización de técnicas complejas y costosas sea menos posible y apropiado.

Además de los criterios mencionados, se pueden también señalar diferencias entre investigación y evaluación según los **destinatarios** o usuarios de los dos tipos de actividad, ya que éstos son potencialmente distintos, siendo los destinatarios de la investigación la comunidad científica en términos abstractos, y de la evaluación, una comunidad de política o programa concreta. En el caso de la investigación, el grupo destinatario principal es la comunidad científica y en un papel secundario, las comunidades de política relacionadas con el tema investigado. En evaluación, ocurre justamente lo contrario: en un papel secundario estaría la comunidad científica, y como destinataria principal la comunidad de política. Esta comunidad de política es, en cualquier caso, compleja y amplia. Además del cliente que la encarga, existen muchas otras personas que están implicadas en el programa, de una manera u otra, y, por lo tanto, tienen necesidades de información e intereses en los procesos evaluativos. Es posible que para una evaluación haya un único cliente que la ha encargado, pero siempre va a haber más destinatarios que esperan algo en concreto de esa evaluación. Como muestra, Guba y Lincoln (1989: 40-41) identifican un amplio listado de trece tipos de posibles agentes críticos² (*stakeholders*) en una evaluación, divididos en tres

² El término *stakeholder* suele utilizarse en inglés debido, en parte a su difícil traducción. Aquí lo vamos a traducir como “agente crítico”, alguien que puede verse afectado o afectar al programa y, por lo tanto, a su evaluación.

grandes clases: los “agentes”, los “beneficiarios” y las “víctimas” del programa a evaluar. En este listado se incluyen desde los patrocinadores del programa (agentes) hasta las personas que sufren costes de oportunidad por el uso del evaluando, tales como las personas que hubieran preferido invertir los recursos en otros programas (víctimas). Por lo tanto, el elenco de destinatarios a los que hay que responder y de posibles actores a involucrar es mucho más amplio y complejo en los procesos de evaluación que en los de investigación.

También se pueden establecer diferencias según los **resultados** que deben producir y que se esperan de los dos tipos de indagación; respecto a los resultados se puede hablar, por un lado, de productos esperados (publicación de resultados en un caso y recomendaciones e informe en el otro) y, por otro, de sus efectos (reconocimiento de la comunidad científica en un caso y utilización de los resultados en el otro).

Debido a la naturaleza y finalidad de la evaluación, ésta debe contar con un último nivel de análisis y exposición de los resultados que consiste en la traducción de los resultados de la evaluación en **recomendaciones**. Estas recomendaciones son sugerencias sobre posibles líneas de acción a seguir respecto al programa o política evaluada, cambios propuestos en el programa o cosas a mantener en el mismo o consejos para la mejora dirigidos a los patrocinadores, responsables y/o profesionales del programa. Evidentemente, estas recomendaciones deben seguir la secuencia lógica y sistemática de los niveles de análisis anteriores (hallazgos, interpretaciones y valoraciones). Es decir, las recomendaciones de una evaluación deben estar basadas en las valoraciones realizadas que, a su vez, habrán estado basadas en la recopilación y análisis sistemático de la información. Este “producto” de las evaluaciones no tiene por qué ser exclusivo de las mismas, ya que una investigación podría concluir también con recomendaciones; pero ni es frecuente ni es un elemento crucial. Sin embargo, en las evaluaciones las recomendaciones son un elemento importante ya que son el vínculo formal entre una evaluación y su uso. Más allá, el ofrecimiento de recomendaciones específicas y orientadas a la acción ha sido considerado desde hace tiempo como algo deseable y relacionado con las buenas evaluaciones (Torres, Preskill y Piontek, 196: 105), y esto no ocurre con la investigación. Patton, quien ha desarrollado un enfoque basado en la utilización, da también importancia a las recomendaciones y, aunque insiste en que no sólo éstas sino todo el resto del proceso evaluativo es lo que provoca y estimula el uso de la evaluación, reconoce, sin embargo, que son, a menudo, la parte más visible de un informe de evaluación (Patton, 1997: 324).

Efectivamente, la preocupación por la utilización y el uso de la evaluación ha sido una constante en la literatura especializada en el campo de la evaluación, cuestión que no ocurre respecto a la investigación. Dada la naturaleza específica de la evaluación, su éxito no sólo se debe medir por el grado en que ésta contribuye al aumento o mejora del conocimiento sino, sobre todo, por el grado en el que ha resultado valiosa para cubrir las necesidades específicas de información de un cliente, destinatario o usuario particular.

Una respuesta natural a esta preocupación por la utilización y por responder a necesidades específicas de información, ha sido la atención prestada a los **procesos de comunicación de los resultados** de las evaluaciones, incluyendo la importancia de la elaboración de un buen **informe**, legible y claro (y su difusión y seguimiento posterior),

y de otras formas alternativas de comunicación (como por ejemplo, sesiones de trabajo, presentaciones orales, conferencias de prensa, comunicaciones cortas como faxes, memoranda o correos electrónicos, vídeos, posters, boletines y resúmenes ejecutivos). Un buen exponente de esta preocupación por el elemento comunicativo es la obra de Torres, Preskill y Piontek (1996), la cual se centra en fundamentar la importancia de dicho elemento en los procesos evaluativos y en ofrecer orientaciones prácticas para la comunicación y difusión de los resultados. Es decir, la responsabilidad del evaluador no acaba con la entrega del informe, sino que también tiene responsabilidades de comunicación, difusión y seguimiento de dicho informe.

Otra de las diferencias entre investigación y evaluación, respecto a los resultados esperados de ambos procesos de indagación, es, como señala Weiss (1998: 16) la **publicación** posterior. Mientras el resultado evidente de la investigación es su publicación –ésta es, de hecho, prácticamente la única vía para la diseminación de las investigaciones-, la mayoría de los informes de evaluación no se publican. Por un lado, los responsables de los programas entienden que la información que generan las evaluaciones sirve, sobre todo, para contestar a sus preguntas y cubrir sus necesidades de información, y no tienen especial interés en que dicha información se difunda. Por otro lado, los evaluadores suelen trabajar con presión de tiempo y, normalmente, después de la entrega del informe, suelen ya estar trabajando en otro estudio. Además, los canales para publicar estudios de evaluación son todavía muy escasos, especialmente en países que, como España, tienen todavía una escasa cultura evaluativa y la función de evaluación muy poco desarrollada e institucionalizada. En el ámbito anglosajón existen una serie de publicaciones periódicas especializada en las que pueden publicarse los resultados de estudios de evaluación y el proceso seguido en los mismos, así como debatir sobre métodos enfoques, filosofía y usos de la evaluación. En realidad, en el campo de la evaluación, las cuestiones referidas a los procesos y a la metodología pueden tener más interés incluso que los propios resultados de los estudios de evaluación, ya que al estar estos centrados en un contexto específico y en un programa o política concreta son menos generalizables o extrapolables que los resultados de las investigaciones. Sin embargo, no cabe duda de que los resultados de las evaluaciones de programas y políticas de un determinado sector pueden ser también de interés para la comunidad científica y profesional de dicho sector (por ejemplo, las políticas de cooperación al desarrollo), y que también pueden tener cabida en las publicaciones y foros especializados de cada uno de los sectores.

Respecto a los efectos de los dos tipos de indagación, se puede decir que la investigación pretende, en general, conseguir aportaciones a la comunidad científica y un reconocimiento por parte de la misma; la evaluación, sin embargo, pretende la utilidad concreta y la utilización de los resultados que produzca. Esta utilización puede ser directa e inmediata, para la toma de decisiones (se utilizan los resultados para tomar decisiones concretas sobre el programa o política evaluada) o más indirecta, difusa o mediata, para el aprendizaje organizativo (se aprende de los resultados y de los propios procesos de evaluación y este aprendizaje general supone una mejora para el programa y la organización o colectivo que lo lleva a cabo).

Finalmente, al ser actividades diferenciadas, el papel de los **agentes** (es decir, investigadores y evaluadores) también puede diferenciarse, incluyendo algunos requisitos y necesidades distintas para ambos. Al fin y al cabo los evaluadores se mueven en un contexto determinante y de características muy distintas (con un fuerte carácter político) a los contextos en los que se generan las investigaciones. Esto significa que los evaluadores deben contar con una serie de destrezas, como son, por ejemplo, el conocimiento de los procesos políticos y ciertas habilidades de comunicación, que no son tan importantes en el caso de los investigadores.

De esta manera, para moverse adecuadamente en los determinantes contextos evaluativos, el *evaluador* requiere de unas habilidades diferentes a las del investigador. Los evaluadores deben ser capaces de manejarse con habilidad en los contextos complejos en los que deben intervenir. Necesitan, desde luego, una buena formación metodológica, al igual que cualquier investigador, aunque ésta debe ser de carácter diferente: el evaluador debe ser un metodólogo “todo terreno”, capaz de dar diferentes respuestas metodológicas a distintas situaciones evaluativas. En nuestra opinión, el buen evaluador no es el especialista en un determinado enfoque, una determinada metodología o en varias técnicas concretas, sino el conocedor de un elenco amplio no sólo de los diferentes enfoques, metodologías y técnicas sino de sus ventajas y desventajas, fortalezas y debilidades en cada situación y contexto específico. Además de la formación metodológica, el evaluador requiere también de muchas otras habilidades que debe desarrollar de forma competente: análisis de políticas públicas y conocimiento sobre los procesos políticos en general, conocimiento y capacidad de trato con grupos y personas, habilidades de comunicación, etc. Por último, es evidente que un cierto conocimiento del sector en el que desarrolla el programa o política evaluada es también necesario; por ejemplo, saber sobre política de cooperación al desarrollo y sobre fortalecimiento institucional, si sobre eso versan los programas a evaluar. Sin embargo, para el investigador, además de una formación metodológica, la clave es ser experto especialista en el sector en el que se investiga.

Otra diferencia estriba en la mayor atención que necesariamente deben prestar los evaluadores a las cuestiones de tipo ético. Es claro que también la investigación (especialmente más en unos campos que en otros) debe guiarse por unos principios éticos, pero el carácter político de la evaluación, apegada a un contexto específico al que tiene que responder y en la que siempre existen un número de personas y grupos relacionados de diferente forma con el objeto de estudio, hace que la aparición de dilemas éticos en situaciones de evaluación sea más frecuente que en situaciones de investigación. Esta preocupación por las cuestiones éticas ha sido recogida por las asociaciones profesionales y académicas de evaluadores y, en los últimos años, se han desarrollado códigos deontológicos que suponen, no sólo orientaciones para afrontar posibles dilemas éticos, sino también una forma de consolidar e institucionalizar el campo profesional de la evaluación (Bustelo, 1998).

Para terminar, se expone a continuación una tabla en la que se resumen las principales características de la investigación y la evaluación según cada uno de los criterios discutidos.

TABLA 1.5: DIFERENCIAS ENTRE INVESTIGACIÓN Y EVALUACIÓN

Criterio	INVESTIGACIÓN	EVALUACIÓN
NATURALEZA/ CONCEPTO	<p>Es un tipo de indagación disciplinada que se lleva a cabo para resolver algún problema de forma que se consiga entendimiento o se facilite la acción</p> <p>Modo de indagación académico</p>	<p>Es un tipo de investigación disciplinada que se lleva a cabo para determinar el valor (mérito y/o valía) del objeto evaluado (programa, proyecto, servicio o política) con el fin de mejorarlo, rendir cuentas y/o ilustrar acciones futuras</p> <p>Modo de indagación práctico</p>
FINALIDAD Y CONDICIONAN TES	<p>Puede ser básica o aplicada</p> <p>Construcción de conocimiento</p> <p>Debe ser una contribución relevante para la comunidad científica</p> <p>Trasciende temporalmente</p>	<p>Siempre es aplicada</p> <p>Mejora de los programas Rendimiento de cuentas Ilustración de acciones futuras</p> <p>UTILIDAD</p> <p>Debe tener capacidad de respuesta ante y cierta lealtad hacia los responsables de los programas</p> <p>Debe ser oportuna en el tiempo TEMPORALIDAD AJUSTADA</p>
CONTEXTO	<p>Escenario de investigación (lo prioritario es la investigación)</p> <p>El contexto es un factor más a tener en cuenta (no necesariamente determinante)</p> <p>Carácter más técnico</p>	<p>Escenario de acción (lo prioritario es el programa evaluado y no su evaluación)</p> <p>El contexto es determinante. La evaluación es dependiente del contexto y tiene por objetivo la respuesta a dicho contexto específico</p> <p>Carácter marcadamente político</p>
OBJETO	<p>Referido a problemas (QUÉ se aborda)</p> <p>Se interesa por la evolución de un problema en abstracto</p> <p>Se puede centrar en algunos aspectos concretos</p>	<p>Referido a las intervenciones para abordar dichos problemas (CÓMO se aborda)</p> <p>Se interesa por la evolución de un problema a consecuencia de la intervención en el mismo</p> <p>Tiende a considerar la intervención de forma global</p>
MÉTODO	<p>Referido la selección, recopilación y análisis de la información</p>	<p>Referido al diagnóstico de la situación evaluativa, la elección del tipo de evaluación, a la definición de los criterios de evaluación, a la recopilación y análisis de información, a la valoración y emisión de juicios a la luz de los criterios de evaluación, al establecimiento de recomendaciones y a la comunicación de los resultados para conseguir una mayor utilidad/utilización de la evaluación</p>

<p>DESTINATARIOS</p>	<p>Grupos abstractos Comunidad científica</p> <p>De forma secundaria, comunidad de política o programa</p>	<p>Grupos y personas concretas Cliente, patrocinadores Responsables, profesionales Agentes críticos (<i>stakeholders</i>): agentes, beneficiarios y víctimas del programa evaluado</p> <p>De forma secundaria, comunidad científica</p>
<p>RESULTADOS: PRODUCTOS Y EFECTOS ESPERADOS</p>	<p>Publicación de resultados</p> <p>Aportación a y reconocimiento de la comunidad científica</p>	<p>Recomendaciones para la acción Informes normalmente no publicados Comunicación y difusión activa de los resultados a los interesados</p> <p>Utilidad y utilización de los resultados y los procesos de evaluación</p>
<p>AGENTES (investigadores y evaluadores)</p>	<p>Capacidad investigadora:</p> <p>Conocimientos básicos sobre metodología de la investigación Necesario conocimiento previo y exhaustivo del sector</p>	<p>Conocimientos metodológicos amplios y de sus posibilidades de aplicación</p> <p>Conocimiento del sector (puede desarrollarse a lo largo del proceso)</p> <p>Otras habilidades necesarias: análisis de políticas públicas y conocimiento de los procesos políticos, conocimiento y capacidad de trato con personas y grupos, habilidades de comunicación</p>

Como se ha visto hasta ahora, las diferencias entre investigación y evaluación son de tales dimensiones, que deben reflejarse no sólo en la práctica de la evaluación sino también en la teoría. Los manuales de evaluación, por ejemplo, deberían dejar de hablar de diseños de evaluación como diseños de investigación social, y concentrar sus esfuerzos en las características metodológicas específicas de la evaluación. Otra cuestión pendiente es, en nuestra opinión, la necesidad de una mayor elaboración, tanto desde el punto de vista práctico como del teórico, sobre la valoración y la emisión sistemática de juicios como parte del análisis y la metodología de evaluación. También creemos que se debe seguir insistiendo en la forma de incluir, de forma sistemática y cómo quehacer específico de los procesos evaluativos, el contexto del programa o política evaluada como marco de referencia de la que arranca y al cual responde la evaluación, manteniendo el rigor de la actividad evaluativa.

Por otro lado, la distinción de la evaluación respecto a otros tipos de análisis e investigación ayuda a reforzar su identidad como actividad científica diferenciada. De esta manera, resulta crucial separar y diferenciar los dos tipos de actividad, sobre todo en el terreno de las políticas públicas. Ambas actividades, investigación y evaluación, son necesarias y complementarias y, por lo tanto, no conviene que restemos a una con la otra. Por ejemplo, en el terreno de las políticas de cooperación al desarrollo es necesario que se siga investigando sobre las desigualdades norte-sur y sobre los índices de desarrollo en los diferentes países y territorio. Incluso es interesante que, con objetivos de análisis e investigación, se realicen estudios comparativos sobre las

políticas de cooperación al desarrollo entre diferentes países. Pero también es absolutamente necesario que cada país evalúe su política de cooperación al desarrollo y los diferentes programas e intervenciones que lleva a cabo con fines de mejora y rendimiento de cuentas. En definitiva, no sólo el hecho de que se dé un alto grado de actividad en uno de los modos de indagación no significa que deba haber baja actividad en el otro, sino que tanto la investigación como la evaluación se ven claramente beneficiadas por una existencia mutua.

De esta manera, se debe señalar la importancia de la investigación para ofrecer un marco conceptual que sirva para la realización de evaluaciones. La evaluación de políticas innovadoras en las que existe poca base teórica sobre qué es posible esperar de una determinada intervención (por ejemplo, las políticas de igualdad de género) suele ser más compleja que la evaluación de programas en torno a los cuales se han realizado múltiples investigaciones (por ejemplo, programas de asistencia a drogodependientes). De igual manera, la existencia de programas sistemáticamente evaluados ayuda a conceptualizar y definir problemas y llama la atención sobre campos de estudio en los que existe una necesidad de investigación.

Finalmente, se quiere apuntar otro terreno en el que, en nuestro parecer, ambos tipos de indagación pueden interactuar de forma fructífera; se trata de la investigación cuyo objeto de estudio son los procesos de evaluación (investigación evaluativa o sobre evaluación), como por ejemplo la realización de metaevaluaciones (evaluaciones de procesos de evaluación), investigación sobre la utilización de la evaluación por parte de los poderes públicos, sobre la institucionalización y el desarrollo de la función de evaluación en las diferentes instituciones, análisis de las políticas de evaluación en los diferentes países, etc. Este tipo de investigación, que corresponde a la perspectiva más descriptiva del análisis de políticas públicas, está todavía poco desarrollada. Su práctica generalizada, sin embargo, contribuiría de forma importante al desarrollo, consolidación e institucionalización de la evaluación de programas no sólo como función normalizada en los sistemas político administrativos, sino también como actividad científica diferenciada.

3.2. El carácter político de la evaluación

En los últimos veinticinco años de desarrollo de la disciplina de la evaluación de programas y políticas, se ha ido reconociendo, de forma paulatina y creciente, el carácter político de la evaluación por parte de la comunidad científica. En algunos casos, se entiende que política y evaluación son dos realidades separadas aunque interactuantes. La evaluación debe responder ante el contexto político y, a su vez, ser una herramienta que permita controlarlo. Para otros autores, como Monnier (1995), la evaluación es un acto político en sí mismo.

Un buen indicador del reconocimiento del carácter político de la evaluación es la publicación de la citada obra *"The Politics of Program Evaluation"* en 1987 y cuyo

editor, Dennis J. Palumbo concluía, en su capítulo introductorio, que la **política y la evaluación están intrínsecamente interrelacionadas**; tanto, que la efectividad de la evaluación depende de la incorporación de la realidad política a los procesos de evaluación. Aunque Palumbo reconoce que existe un cierto peligro en lo que él llama la "evaluación política", sin darle un sentido peyorativo, (los poderosos se pueden beneficiar más que otros ya que tienen mayores probabilidades de controlar las condiciones bajo las que se realizan las evaluaciones), la única forma de tener un mayor impacto en las políticas y los programas es incorporando los factores políticos en los diseños de la evaluación y la difusión de sus resultados. Es más, precisamente es la ingenua falta de atención a las dimensiones políticas lo que puede contribuir a que la evaluación se utilice como una herramienta ideológica. Palumbo considera que la política y la evaluación están relacionadas de diversas formas, dos de ellas de carácter positivo: por un lado, si las evaluaciones son utilizadas se convierten en una parte del proceso de decisiones políticas del programa evaluado; por otro, las evaluaciones adquieren una postura sobre el programa y esto es inevitable e inherentemente político. Una tercera es de carácter negativo y es que una evaluación puede realizarse con el propósito de apoyar o crear imagen a un programa. Esta obra representa un hito en el campo de la evaluación no sólo por el reconocimiento explícito de la relación política-evaluación, sino también porque **resta carácter peyorativo al término "política"**: no se trata sólo de partidismos entre partidos políticos sino, sobre todo, de "las interacciones de varios actores dentro y entre burocracias, clientes, grupos de interés, organizaciones privadas y legislaturas y sobre cómo se relacionan unos con otros desde diferentes posiciones de poder, influencia y autoridad" (Palumbo, 1987:18-19). Esta es la primera obra que tiene como propósito explícito aconsejar a los evaluadores la necesidad de comprender la naturaleza política de la evaluación y cómo incorporar la política a su quehacer evaluador.

Una de las primeras autoras que trata el tema de evaluación y política es Carol H. Weiss, quien escribe un artículo en 1973 que también será publicado catorce años más tarde en la obra de Palumbo anteriormente mencionada. En dicho artículo titulado *Donde la política y la investigación evaluativa se encuentran*, la autora sostenía que la evaluación es una tarea racional pero que se lleva a cabo en un contexto político en el que hay que tener en cuenta tres consideraciones: la primera, que las políticas y programas que se evalúan son resultado de decisiones políticas. Segundo, la evaluación al llevarse a cabo para alimentar el proceso de toma de decisiones entra, inevitablemente, en la arena política. Tercero, la evaluación en sí misma adopta una postura política ya que, por su misma naturaleza, expone implícitamente cuestiones tales como la naturaleza problemática o necesidad de algunos programas, la legitimidad de las estrategias y objetivos, etc. Conocer estas resistencias y fuerzas políticas es necesario para realizar una evaluación útil. Años más tarde (Weiss, 1998), la autora sigue insistiendo en que la evaluación no se hace en abstracto, sino en un contexto determinado; más allá, ésta se realiza, principalmente, para facilitar la toma de decisiones sobre el programa o política evaluada. En definitiva, se reconoce el contexto político, incluso como determinante de una situación evaluativa a la que vamos a tener que responder metodológicamente de formas diversas. La evaluación como proceso sistemático reconoce, pero también ayuda a contrarrestar y controlar, ese clima político en el que está inmerso cualquier programa o política, especialmente si es público. De esta manera, para Weiss, el rigor de la evaluación sistemática es particularmente importante si los programas son complejos, las decisiones a tomar son importantes y

caras y si se necesita información para convencer a otros de la validez de las conclusiones.

En la misma línea, Evert Vedung (1997) configura la evaluación como una función que pueden cumplir los académicos y científicos para que gobernantes y poderes públicos reflexionen, comprendan mejor y tomen decisiones más sólidas y fundamentadas. En definitiva, aunque la evaluación se lleve a cabo en y para un "ambiente político", evaluación y política son dos realidades separadas; la primera cumple una función de ayuda y control sobre la segunda. Ambas son necesarias y se mezclan con frecuencia. Así, el evaluador debe reconocer y ser consciente que su trabajo puede verse afectado por determinadas luchas de poder.

Sin embargo, otros autores como Monnier (1995) van un paso más allá, y consideran que no sólo se trata de que la evaluación se desarrolla inevitablemente en un contexto político que hay que reconocer y tener en cuenta, sino que es una actividad política en sí misma y, por lo tanto, forma parte integrante de dicho contexto político. Es decir, evaluación y política no son dos realidades fácilmente separables. De esta manera, la evaluación, al margen del rigor metodológico, debe procurarse una legitimidad política. Y muchas veces, la credibilidad de una evaluación depende más de dicha legitimidad política -que supone el reconocimiento y consideración de los distintos actores sociales- que de una exclusiva legitimidad científico-técnica. Para Monnier, la evaluación sirve, entonces, como "elemento negociador entre Estado y sociedad civil" y, por lo tanto, elemento generador e integrante de la realidad política.

En general, este reconocimiento del contexto político y su interacción determinante con los procesos de evaluación, ha ido aumentando en los últimos años. De esta manera, el carácter político de la evaluación deja de ser un tema relacionado con la práctica evaluativa para ser parte crucial de la teoría de la evaluación. Buen exponente de ello es la conferencia dictada por Eleanor Chelimsky en el encuentro anual de 1997 de la AEA (Asociación Americana de Evaluación) en San Diego, titulada "El papel de la experiencia en la formulación de teorías sobre la práctica de la evaluación" (Chelimsky, 1998) y en la que considera que una de las cuestiones aprendidas con la experiencia, y que deben ser incorporadas a la teoría de la evaluación, es la inequívoca relación entre evaluación y política. La autora cree que existen dos problemas en la forma en la que se ha pensado que la evaluación encaja en el mundo real, especialmente en el mundo político en el que se desenvuelve: en primer lugar, cree que no se ha pensado en ello lo suficiente y, en segundo lugar, piensa que, cuando se ha pensado en ello, se han examinado evaluación y política separadamente, considerando la política como periférica o contextual a la evaluación y no como conectada a aquella de forma visceral. Sin embargo, se debe considerar la política como algo *central* en evaluación.

¿Cuáles serían las implicaciones de esta centralidad política, propuesta por Chelimsky, para la teoría de la evaluación? La primera implicación tiene que ver con la **necesidad de ampliar la visión** sobre la política y el proceso de elaboración de políticas. "Todavía, incluso hoy en día, entendemos la política meramente como el 'contexto' de una evaluación, como algo que se 'entromete' en el buen hacer, más que el motor que la conduce" (Chelimsky, 1998: 39). En general, los evaluadores deben ser

más conscientes del dinamismo político que rodea cualquier situación evaluativa y reconocer su centralidad para la práctica profesional.

La segunda implicación de esta naturaleza política es la **necesidad de credibilidad** para la práctica evaluativa. Si los resultados de la evaluación deben ser útiles y sobrevivir a los avatares políticos, la evaluación misma debe ser percibida como creíble e imparcial. En definitiva, "el evaluador debe ser el proveedor de información honesta en un mar de partidismos" (Chelimsky, 1998: 51). Esta idea choca parcialmente con el reciente planteamiento, muy extendido en parte de la comunidad científica, que consideran que cierto compromiso partidario del evaluador con un ideal (*advocacy evaluation*) es inevitable (Greene, 1997). Este panorama era distinto diez años antes cuando Palumbo decía que "el dilema político con el que se enfrentan los evaluadores es seguir un rumbo entre el reconocimiento de la realidad política de la evaluación y mantener el simbolismo de la neutralidad" (Palumbo, 1987: 20). Es decir, parece que la comunidad científica -al menos parte de ella- reconoce en la actualidad que la neutralidad en evaluación es prácticamente imposible. En cualquier caso, parece adecuado seguir el planteamiento de Chelimsky, quien señala que reconocer que el evaluador tiene sus valores y creencias no quiere decir que se tenga que recomendar cierto partidismo; precisamente, una forma de contrarrestar el sesgo del evaluador es contar con el sesgo opuesto al seleccionar el equipo de evaluación³. En general, se trata de enfatizar la imparcialidad en la realización del trabajo y la credibilidad en su percepción externa.

La tercera implicación es la **temporalidad o necesidad de oportunidad en el tiempo (*timeliness*)**, es decir, que la evaluación se haga en el momento oportuno y no se retrase. Cuando el clima político cambia, presentar los resultados de una evaluación demasiado tarde puede significar lo mismo que no presentarlos. Sin embargo, presentarlos a tiempo puede marcar la diferencia entre ser escuchados o no, que se utilicen los resultados o no, o que se aumente o disminuya la credibilidad del evaluador. En otras palabras, si se tarda mucho en contestar las preguntas de evaluación, puede que el debate que las generó ya no sea el mismo, que esas preguntas sean ya irrelevantes o que los destinatarios de la evaluación hayan desaparecido.

Por último, la cuarta implicación se refiere a la **necesidad de flexibilidad** y que significa una cierta provisionalidad durante el proceso de evaluación. Esto puede significar enfrentar una polémica inesperada cambiando el diseño de la evaluación, añadiendo nuevos consultores o nuevos componentes a la evaluación. En definitiva, cada situación evaluativa puede requerir una respuesta evaluativa única, adaptada a las circunstancias concretas. También implica ser más cautelosos a la hora de hacer valoraciones absolutas, enfatizando las observaciones lo más específicas, precisas y validadas posible.

Se pueden mencionar otras implicaciones de este "reconocimiento político" de la evaluación, ya apuntadas por Palumbo en su obra de 1987. Esto significa reconocer que los imperativos de los evaluadores y los de los gestores públicos son distintos y muchas veces entran en conflicto, que hay que entender el valor simbólico y de cambio de la

³ Esta práctica de incluir en un mismo equipo evaluador sesgos opuestos es lo que hizo la GAO (*General Accounting Office*) con la evaluación de los polémicos programas de educación bilingüe en EE.UU.

información que ofrecen las evaluaciones (la información es poder), que es necesario extender las evaluaciones más allá de los objetivos de un programa y comprender los procesos organizativos que subyacen a toda situación evaluativa. También implica que hay que tener en cuenta a todos los agentes críticos (*stakeholders*) -especialmente a los más indefensos- y reconocer que existen diferentes necesidades de información en diferentes puntos del ciclo de una política.

Finalmente, nos gustaría hacer una reflexión sobre otras posibles implicaciones que tiene este inevitable carácter político de la evaluación. En primer lugar, creemos que la evaluación sólo puede entenderse plenamente desde una **comprensión profunda de los procesos políticos y de las políticas públicas**, mucho más que desde una posición técnica estrictamente metodológica. La evaluación de programas y políticas públicas es algo más que una herramienta metodológica o, si se quiere, la "metodología" de evaluación debe entenderse en un sentido mucho más amplio del que normalmente se le asigna. Son importantes los pasos dados en los últimos años en los que la evaluación va tomando cuerpo como disciplina, diferenciándose de la investigación e incluso, del análisis de políticas públicas (Lincoln y Guba, 1986), y reclamando un cuerpo teórico propio en el que su carácter político es central (Chelmsky, 1998). Esto lógicamente ha de tener implicaciones en cómo se concibe la formación de evaluadores, cuestión que se abordará más adelante en este capítulo. En segundo lugar, reafirma la idea de que, en evaluación existe un **doble componente**: además de un componente **cognitivo**, relacionado con los aspectos sistemáticos, cognitivos y "comparativamente objetivos" de la investigación y la indagación, existe un componente **intuitivo**, relacionado con aspectos intuitivos y cargados de significado por parte del evaluador, sus clientes, los agentes críticos en la evaluación y las organizaciones en las que están inmersos. Este segundo componente se ha ido reconociendo en la última década, considerándose que debe ser integrado con el primero (Torres, Preskill y Piontek, 1996).

3.3. El carácter práctico y aplicado a la acción: el elemento valorativo y el establecimiento de recomendaciones

Ya se ha insistido, a lo largo de este capítulo, sobre la especificidad del **elemento valorativo o de juicio** que tiene la evaluación. Además de recopilar e interpretar sistemáticamente información, la evaluación cuenta con otros dos niveles de análisis (valoración y establecimiento de recomendaciones) que la hacen específica respecto a otros tipos de análisis e investigación.

En los primeros años de desarrollo de la disciplina de la evaluación, en los que dominaba además una concepción epistemológica netamente positivista y se concebía que la ciencia podía y debía estar libre de valores, había algunos autores que creían que esta orientación valorativa conducía a interpretaciones radicales y subjetivas y otorgaba un poder innecesario a los evaluadores. Aunque esta visión está superada en la

comunidad científica y profesional de la evaluación, todavía podemos encontrar algún resquicio hoy en día⁴. Sin embargo, la mayoría de los autores clave en evaluación dan a este elemento valorativo una importancia central en el desarrollo y articulación de las teorías de evaluación (Shadish, Cook y Levinton, 1991).

Ya en los años 80, este elemento valorativo es plenamente reconocido e incluso utilizado como elemento central en la definición de la evaluación. Stufflebeam y Shinkfield en su obra *Evaluación sistemática* (la obra original data de 1985 y la versión castellana de 1987) insisten en que la meta esencial de la evaluación “es determinar el valor de algo que se está enjuiciando. De acuerdo con esta definición, si un estudio no informa de cuán buena o mala es una cosa, no se trata de una evaluación” (Stufflebeam y Shinkfield, 1987: 20). Precisamente, es este elemento el que organiza su clasificación de modelos o enfoques de evaluación, ya que estos autores denominan “cuasievaluaciones” a los enfoques que no incluyen la valoración como parte de la evaluación y “evaluaciones verdaderas” a los que sí lo hacen.

El reconocimiento de este elemento valorativo implica su consideración como elemento sistemático en el proceso de evaluación. Esto significa que no sólo hay que procurar que la información que se recopile sea de calidad sino que se proporcione y clarifique una base lógica que justifique las perspectivas de valoración utilizadas para interpretar y valorar los resultados. Ya que existen distintos públicos y destinatarios de una evaluación, se pueden y deben emplear diferentes perspectivas de valoración. Por ejemplo, de un mismo programa de prevención de drogodependencias en centros educativos, la Agencia Antidroga puede estar interesada en la valoración del programa respecto a la reducción de consumo de drogas en una zona determinada. Sin embargo, el profesorado que participe en los proyectos de prevención de los centros estará más pendiente del interés y motivación de los chicos y de la posibilidad de obtener recursos y desarrollar actividades como el fomento de la autoestima y del deporte que no sólo ayuden a prevenir la drogodependencia y otros problemas sociales, sino que fomenten el desarrollo general del alumnado. Esta necesidad de tener en cuenta las diferentes perspectivas de valoración de los actores críticos en el programa o política evaluada, es una de las claves de la evaluación pluralista propuesta por Monnier. Precisamente, es el compromiso sobre la interpretación de los hechos y sus valoraciones entre observadores y observado, de una parte, y entre los diferentes protagonistas del programa, por otra, lo que hace que una evaluación sea legítima políticamente y, por lo tanto, útil socialmente (Monnier, 1990-91: 143-145).

Siguiendo a Stufflebeam y Shinkfield (1987: 21), además de los diferentes protagonistas, también habrá que tener en cuenta los conceptos de mérito y valor, además de los de viabilidad y equidad. En resumen, las bases de valoración que deben ser tenidas en cuenta son las siguientes:

1. Las distintas expectativas de los clientes. Nosotros añadiríamos, además, las diferentes demandas de los ciudadanos.

⁴ En la parte III de esta tesis doctoral se analiza como en una publicación basada en los resultados de evaluación de uno de los planes autonómicos de igualdad, se menciona expresamente que se exponen los datos recopilados e interpretados, “procurando excluir los juicios de valor... (para que).. la lectora o el lector sean quienes saquen sus propias conclusiones” (Comunidad de Madrid, 1997, *Las mujeres en la Comunidad de Madrid*, p.3) .

2. El mérito o excelencias del servicio, programa o política en cuestión, es decir, ¿hace bien lo que se supone que debe hacer?
3. Hasta qué punto el servicio, política o programa es necesario o adecuado a la comunidad a la que se dirige, es decir si tiene validez.
4. Hasta qué punto es viable, ya que es posible que consuma más recursos de los que necesite o cause infinidad de problemas políticos.
5. La equidad necesaria en toda sociedad democrática para los programas y políticas públicas. Un servicio, programa o política pública debe estar pensado, ser asequible para todos los miembros de la sociedad, todos los grupos deben disfrutar de una participación equitativa en la utilización de los servicios, los niveles de logro o consecución de objetivos no deben ser sustancialmente diferentes para los distintos grupos, y se deberán producir unos resultados similares en la mejora del bienestar de todos los grupos sociales.

Estas bases de valoración se verán reflejadas en los criterios de evaluación que se establezcan en cada proceso evaluativo. Dunn (1981: 343) propone seis grandes criterios de evaluación, que son:

- *Efectividad*: ¿Se ha conseguido los resultados?
- *Eficiencia*: ¿Cuánto esfuerzo se ha requerido para conseguir dichos resultados?
- *Adecuación*: ¿Hasta qué punto el logro de los resultados ha resuelto el problema?
- *Equidad*: ¿Están distribuidos los costes y los beneficios de forma equitativa entre los diferentes grupos?
- *Capacidad de respuesta (responsiveness)*: ¿Satisfacen los resultados de la política las necesidades, preferencias o valores de los grupos implicados?
- *Conveniencia o propiedad*: ¿Merecen realmente la pena los resultados (objetivos) deseados?

Además, las bases de valoración pueden ser relativamente diferentes, y tener sus matices específicos, dependiendo del ámbito al que pertenezca la política o programa evaluado. Así, por ejemplo, en la evaluación de las políticas de cooperación al desarrollo existe un consenso generalizado sobre la importancia de valorar a través de los criterios generales de *eficiencia* (hasta qué punto los resultados alcanzados derivan de una utilización eficiente de los recursos financieros, técnicos y humanos), *eficacia* (hasta qué punto se han alcanzado los objetivos propuestos), *impacto* (consecuencias previstas y no previstas para la sociedad, tanto negativas como positivas), *pertinencia* (hasta qué punto el programa evaluado se ajusta, en este caso, a los objetivos globales de desarrollo, a la política del donante y del receptor y a las necesidades y prioridades locales) y *viabilidad* (definida en este ámbito como hasta qué punto los cambios positivos logrados como consecuencia del programa se mantienen después de que éste haya finalizado) (NORAD, 1997, Ministerio de Asuntos Exteriores, 1998).

Por la propia naturaleza de la evaluación, ésta debe contar con un último nivel de análisis y exposición de los resultados que consiste en la traducción de los resultados de la evaluación en **recomendaciones**. Estas recomendaciones son sugerencias sobre posibles líneas de acción a seguir respecto al programa o política evaluada, cambios propuestos en el programa o cosas a mantener en el mismo o consejos para la mejora dirigidos a los patrocinadores, responsables y/o profesionales del programa. Evidentemente, estas recomendaciones deben seguir la secuencia lógica y sistemática de los niveles de análisis anteriores. Es decir, las recomendaciones de una evaluación deben estar basadas en las valoraciones realizadas que, a su vez, habrán estado basadas en la recopilación y análisis sistemático de la información.

Las recomendaciones son, además “el vínculo formal entre una evaluación y su uso”. Además, el ofrecimiento de recomendaciones específicas y orientadas a la acción ha sido considerado desde hace tiempo como algo deseable y relacionado con las buenas evaluaciones (Torres, Preskill y Piontek, 1996: 105). Patton, quien ha desarrollado un enfoque basado en la utilización, da también importancia a las recomendaciones y, aunque insiste en que no sólo las éstas sino todo el resto del proceso evaluativo es lo que provoca y estimula el uso de la evaluación, reconoce, sin embargo, que son, a menudo, la parte más visible de un informe de evaluación (Patton, 1997: 324).

Sin embargo, esta idea sobre la necesidad y conveniencia de las recomendaciones también tiene algunos detractores. Michael Scriven, en su obra *Hard-Won Lessons in Program Evaluation* expone 31 tesis, aprendidas, según su autor, de los errores cometidos en la práctica de la evaluación (Scriven, 1993). Una de sus tesis, con el irónico título de “Una evaluación sin una recomendación es como un pez sin una bicicleta” intenta persuadir sobre la idea de que las evaluaciones no tienen por que concluir con recomendaciones; es más, para él ésta es una tarea que no debe realizar el evaluador (siempre situándose desde la perspectiva del evaluador externo). Según él, es una falacia lógica pensar que juzgar el mérito y la valía conduce directamente a las recomendaciones; la principal función del evaluador es valorar y hacer juicios, no la de hacer recomendaciones. A continuación, analizamos la tesis de Scriven y su fundamentación y cuáles son los argumentos que, en nuestro parecer, se pueden utilizar para contrarrestarla.

El primer argumento que utiliza Scriven es que, en muchos casos, el evaluador no está lo suficientemente cualificado para dar ese paso de pasar de las conclusiones a las recomendaciones. Puede que necesite un mayor conocimiento del contexto local o conocimientos específicos o determinados (por ejemplo, de tipo legal). Es cierto que los evaluadores deben contar con un alto conocimiento tanto del contexto como de otras cuestiones relacionadas con los programas o políticas que evalúan. Sin embargo, creemos que los evaluadores conocen necesariamente a fondo dicho contexto, ya que han tenido que diagnosticar previamente la situación evaluativa para responder con un proceso evaluativo adecuado. Por otro lado, existe una alta probabilidad de que los evaluadores competentes aprendan mucho sobre el programa, su contexto, el ámbito o tema sobre el que versa dicho programa y las diferentes perspectivas en torno al mismo. Además, una práctica muy utilizada en los equipos de evaluación es contar con el asesoramiento de especialistas en los temas tratados, de manera que se pueda contrarrestar esa falta de conocimientos específicos que ayudan, definitivamente, a

contextualizar las recomendaciones. Teniendo en cuenta estas circunstancias, creemos que el evaluador está, precisamente, en una situación privilegiada para poder ofrecer recomendaciones, entre otras cosas y sobre todo, porque éstas emanan del estudio riguroso y sistemático del programa y su contexto.

El segundo argumento que utiliza Scriven es que, desde el punto de vista del consumidor que quiere información para elegir entre diferentes “productos”, no tiene ningún interés ni las recomendaciones para la acción (“compre este determinado producto”) ni para la mejora (esto le interesará al responsable del producto, no al consumidor). Este enfoque orientado al consumidor, que parece especialmente adecuado para orientar a tomar decisiones en una situación de mercado -al estilo de la famosa publicación *Consumer Reports*- no parece tan adecuado para la evaluación de programas y políticas públicas. No sólo en la mayoría de los casos, la ciudadanía no tiene la oportunidad de elegir entre varios servicios, programas o políticas concreta, sino que, incluso cuando esto es así, esta información es válida única y exclusivamente en un momento muy determinado (en el de la elección de entrar o participar en un determinado programa). Pongamos un ejemplo en el que podríamos emular una cierta situación de mercado. Unos padres tienen que elegir el colegio público al que asistirán sus hijos, ya que los criterios de zonificación en su municipio son amplios y pueden elegir, al menos entre 3 o 4 centros⁵. En el momento de la selección, esta perspectiva de “consumidores” va ser especialmente interesante. Pero una vez seleccionado el colegio, la perspectiva cambia radicalmente, porque rara vez se cambia de colegio a un niño porque hemos comprobado que otro funciona un poco mejor o para probar en cual entre varios le va mejor. Definitivamente, a los padres les interesará que el colegio de sus hijos sea evaluado y que se ofrezcan recomendaciones para la mejora. Por último, esta visión liberal de la evaluación de Scriven presupone una alta madurez a la ciudadanía “consumidora” de los recursos públicos. Si a esto le añadimos la falta de cultura evaluativa en general y de los poderes públicos en particular, como es el caso de España, creemos que “desperdiciar” la capacidad del evaluador para establecer recomendaciones y dejar a poderes públicos poco acostumbrados y con poco tiempo la tarea de “traducir” los análisis en posibles cursos de acción, parece, cuanto menos, insensato. Por otro lado, los ejemplos que utiliza Scriven para ilustrar sus argumentaciones son de unos determinados tipos de evaluación (de personal y de productos) que, en cierto modo, están alejados de una concepción más general de evaluación de programas o políticas.

Patton (1997) también critica la postura de Scriven. Para Patton, aunque el consejo de Scriven sobre evitar hacer recomendaciones cuando uno carece de las habilidades necesarias es sabio, falla a la hora de considerar que la búsqueda de dichas habilidades y la captación de expertos para ello forman parte de la responsabilidad del evaluador (Patton, 1997: 327).

En cualquier caso, la literatura sobre evaluación sólo ha prestado especial atención al asunto de las recomendaciones de forma reciente. Incluso Patton llega a decir que las recomendaciones son la parte más débil de la evaluación. “Hemos

⁵ También podemos suponer que estos criterios de elección existen en cualquier caso. En algunos países, como en EE.UU. es práctica habitual que las familias sitúen su lugar exacto de residencia en función del colegio elegido para los hijos.

progresado enormemente en las formas de estudiar programas, en la diversidad metodológica y en una variedad de técnicas y diseños para la recopilación de datos. Las consecuencias de dichos avances a menudo culminan en recomendaciones, pero hemos progresado comparativamente menos en cómo construir recomendaciones útiles. Encuentro que enseñar a los estudiantes sobre cómo pasar de los datos a las recomendaciones es a menudo el mayor reto en la enseñanza de la evaluación. Es un proceso que no es simple ni linear.” (Patton, 1997: 328).

Esta atención a las recomendaciones ha hecho que se hayan desarrollado una serie de orientaciones y sugerencias prácticas para su elaboración (Torres, Preskill y Piontek, 1996; Patton, 1997; Chelimsky, 1998), las cuales nos parece importante señalar porque ellas son reflejo de las teorías de la evaluación que se han ido construyendo con una importante base práctica (Shadish, Cook y Levinton, 1991). Todas estas sugerencias denotan la importancia concedida a la elaboración de recomendaciones como colofón de los procesos de evaluación. Al fin y al cabo, cuando se incluyen recomendaciones en un informe de evaluación, éstas “atraen la atención del lector como el néctar de las flores a las abejas” (Patton, 1997: 324).

En primer lugar, una de las cuestiones a tener en cuenta al desarrollar las recomendaciones es quiénes participan en su elaboración y de qué forma. Esta cuestión va a depender, sobre todo, del enfoque con el que se haya diseñado y emprendido la evaluación, especialmente si es un enfoque de tipo participativo. La extensión con la que participen en las recomendaciones los agentes críticos en el programa puede ser diversa (Patton, 1997: 333) considerarse en un continuo (Torres, Preskill y Piontek, 1996: 106). En un extremo estarían los evaluadores que desarrollan ellos mismos las recomendaciones y se las entregan a los agentes críticos (*stakeholders*) y en el opuesto estarían las situaciones en las que son los propios implicados o *stakeholders* los que desarrollan las recomendaciones. En los puntos intermedios se encontrarían los procesos en los que los evaluadores presentan recomendaciones para que sean criticadas y modificadas por los implicados y agentes críticos y también las situaciones en las que evaluador e implicados colaboran conjuntamente para su elaboración. En cualquier caso, estos autores que prestan especial atención a las recomendaciones suelen dar una gran importancia a la implicación de los destinatarios de la evaluación en el desarrollo de las mismas. Así, sugieren que, en caso de que las recomendaciones estén hechas por el evaluador, éste debe entregarlas antes para permitir la revisión, discusión y/o modificación. También desde su punto de vista es importante que los agentes críticos (*stakeholders*) sientan como suyas las recomendaciones; para aumentar este sentido de pertenencia lo mejor es desarrollarlas conjuntamente con ellos. Incluso, para aumentar el máximo sentido de pertenencia del proceso de evaluación en general y su uso posterior, recomiendan implicar en el análisis de los datos a los participantes en el programa evaluado (Torres, Preskill y Piontek, 1996: 106-7).

En segundo lugar, hay que tener en cuenta el tipo de recomendaciones que se quiere realizar. Las recomendaciones pueden ser:

- simples (se desarrolla la mejor opción)
- se desarrollan varias opciones, analizando las debilidades, amenazas, costes y beneficios de cada una

- opciones basadas en diferentes escenarios futuros con sugerencias para el seguimiento
- diferentes recomendaciones para diferentes destinatarios

La idea de desarrollar varias opciones puede resultar especialmente interesante. El propio Patton insiste en que algunos decisores prefieren recibir opciones múltiples que verse avocados a una única posibilidad de acción, y en cualquier caso señala la conveniencia de que, siempre que sea posible hay que incluir en las recomendaciones algunas consideraciones sobre los costes y beneficios de seguir dichas recomendaciones. También Chelimsky señala que, en su experiencia como responsable de los estudios de evaluación de la GAO (*General Accounting Office*), han llegado a la conclusión de que es muy importante medir los costes de las recomendaciones para aumentar la probabilidad de su uso. Para ella, la inclusión de la cuestión de los costes en el proceso de evaluación mejora la percepción de credibilidad del estudio, ayuda a las evaluaciones a resistir los ataques políticos y aumenta la probabilidad de que los resultados sean utilizados con el tiempo” (Chelimsky, 1998: 47). Estas sugerencias hacen que se dé una mayor importancia a esta fase final de las recomendaciones: no sólo el evaluador puede tener capacidad para hacerlo sino que tiene que dedicarle especial tiempo, esfuerzos y cuidado (aumentando, por lo tanto, los costes de los estudios de evaluación). Esto resalta el papel del evaluador no sólo como metodólogo e investigador, sino como asesor y consultor.

Otras cuestiones relacionadas con las recomendaciones y que se encuentran en la literatura (Patton, 1998; Torres, Preskill y Piontek) son las siguientes:

1. Las recomendaciones deben seguir claramente y deben ser apoyadas por los hallazgos de la evaluación.
2. El tipo de recomendaciones que van a aparecer en un informe final es materia de negociación con los clientes de la evaluación (patrocinadores de la evaluación y agentes críticos –*stakeholders*- implicados en el programa). En realidad se debe negociar la naturaleza y la totalidad del contenido de los informes, una vez que el propósito de la evaluación ha sido clarificado y antes de que se recopile la información. Las recomendaciones no son una excepción.
3. Hay que distinguir entre diferentes tipos de recomendaciones. La cuestión aquí estriba en que las listas largas e indiscriminadas difuminan el foco de atención y disminuyen el poder de las recomendaciones cruciales. Patton (1997) sugiere 6 formas de diferenciarlas:
 - centrales y secundarias
 - sumativas y formativas
 - de aplicación inmediata/a medio plazo/a largo plazo
 - según a quién van dirigidas: patrocinadores, gestores del programa, profesionales, clientes o participantes

- si están sólidamente apoyadas por los datos (y todos los miembros del equipo de evaluación están de acuerdo) o menos sólidamente apoyadas y existe algún disenso respecto a las mismas

- si implican una creencia sólida sobre que se deba actuar o simplemente pretenden estimular la discusión o hacer sugerencias que pueden formar parte de la agenda en un futuro.

4. Las recomendaciones deben ser realistas, o mejor dicho, se deben enfocar hacia acciones bajo el control de los destinatarios de la evaluación, es decir, sobre lo que puedan hacer algo. En esta misma línea, se sugiere ofrecer datos y descripciones que tengan que ver con los elementos clave del contexto del programa evaluado.

5. Al escribir las recomendaciones, hay que ejercitar la sensibilidad política. Se recomienda ponerse en la piel de los destinatarios de la evaluación y trabajar con ellos, de forma anticipada las implicaciones políticas de cada recomendación.

6. Poner especial cuidado en la redacción de las recomendaciones, ya que se pueden diluir en un lenguaje vago y obtuso. Se debe intentar evitar palabras que confunda o distraigan del mensaje central.

7. Dedicarle suficiente tiempo a la redacción de las recomendaciones, para trabajarlas con los destinatarios y para comprobar su claridad, comprensión, carácter práctico, utilidad y significado. Se insiste en que, a menudo, el problema es que los evaluadores acaban haciendo las recomendaciones de prisa y corriendo, sin dedicarle el tiempo suficiente para hacer un buen trabajo.

8. Desarrollar estrategias para que las recomendaciones se tomen en serio. Se sugiere que el evaluador ayude a los decisores a tomar decisiones sobre las recomendaciones con, por ejemplo, sesiones de trabajo que asignen responsabilidades de seguimiento y fechas tope para la puesta en práctica de dichas recomendaciones.

3.4. El carácter práctico y aplicado a la acción: La preocupación por la utilización y el uso de la evaluación

La preocupación por la utilización y el uso de la evaluación ha sido una constante en la literatura especializada a lo largo de la evolución de la disciplina, aunque para algunos autores, como Monnier, “hicieron falta muchas desilusiones por parte de los investigadores ante la poca atención otorgada por los responsables políticos o por los gestores a sus resultados (...), para que la cuestión de la utilización de los resultados se planteara de forma clara” (Monnier, 1990-91: 142). Dada la naturaleza específica de la evaluación, su éxito no sólo se debe medir por el grado en que ésta contribuye al

aumento o mejora del conocimiento sino, sobre todo, por el grado en el que ha resultado valiosa para cubrir las necesidades específicas de información de un cliente, destinatario o usuario particular. Esto hace que, necesariamente, se distingan diferentes necesidades de información.

Efectivamente, una de las primeras consecuencias de esta preocupación por la utilización de la evaluación, ha sido la de esclarecer y clasificar los diferentes usos que de una evaluación pueden hacerse y las diferentes necesidades de información que pueden cubrirse.

Los diferentes usos vienen determinados, en principio, por la función dominante de cada evaluación. Si la función principal es la mejora del programa o política a evaluar, el posterior uso vendrá dado por dicha función y las necesidades de información estarán determinadas por una perspectiva de desarrollo o mejora. Si la función dominante es la rendimient de cuentas, las necesidades de información vendrán determinadas por una perspectiva de *accountability*, en la que se necesitará información que permita, sobre todo, hacer juicios generales sobre las responsabilidades del programa. Finalmente, si domina la función de conocimiento o *enlightenment*, las necesidades de información serán más generales y la perspectiva de uso estará más determinada por los valores académicos.

Chelimsky, en su artículo *Linking Program Evaluation to User Needs* (Chemimsky, 1987: 73) explora las relaciones entre 1) las funciones y los usuarios de la evaluación, 2) los diferentes tipos de necesidades de información, 3) los diferentes momentos del ciclo de vida de un programa o política y 4) los diferentes tipos de estudios de evaluación. Resultante de este cruce de los cuatro aspectos mencionados, la autora realiza una interesante tabla en la que expone las necesidades de información para seis tipos de destinatarios gubernamentales de las evaluaciones (tomando como modelo EE.UU., distingue tres en la rama ejecutiva y tres en la legislativa) y según sea una propuesta para un programa nuevo o un programa ya existente o recién ejecutado (Chelimsky, 1987: 83-85).

Así, para la **función** referida a la formulación de políticas (*policy formulation*) el tipo de información requerida es referente a tres áreas: información sobre el problema, sobre los resultados de otros programas similares e información que permita la selección de una alternativa sobre las otras. Para la ejecución de políticas (*policy execution*) se necesitará información referida a la implementación del programa, a la gestión y dirección del programa y a la información actualizada sobre el estado actual del problema que se intenta solucionar Finalmente, para el propósito del rendimient de cuentas (*accountability*) se necesita información sobre los resultados de un programa, sobre si ha conseguido los objetivos propuestos y sobre los efectos (esperados e inesperados). También los diferentes **destinatarios o usuarios** de la evaluación van a tener diferentes necesidades de información, dependiendo de sus funciones específicas, propósitos y nivel de responsabilidad: algunas serán comunes a todos y otras diferentes. La capacidad para servir a diferentes usuarios, aunque deba intentarse, siempre se va a ver constreñida por la necesidad del evaluador de contestar a las preguntas precisas que haya hecho el usuario principal (normalmente el cliente de la evaluación). Las necesidades de información van ser también distintas según el **tipo de programa o política y el momento del ciclo** en el que se encuentre. Por ejemplo, va a influir la

antigüedad del programa (que sea nuevo o ya establecido) o si es aceptado o, por el contrario, controvertido, o si es un programa caro o modesto. Finalmente, Chelimsky distingue cinco **actividades básicas** respecto al programa o política a evaluar: 1) planificar y racionalizar un programa y su evaluación, 2) ejecutar y gestionar un programa, 3) justificar la efectividad de la ejecución y gestión, 4) demostrar la efectividad de un programa y 5) medir el problema y el progreso del programa. Cada una de estas actividades requiere de diferente información evaluativa.

Aunque las conclusiones concretas de Chelimsky son difícilmente aplicables de forma directa a la situación de España, creemos que su idea de clasificar y relacionar los cuatro factores anteriormente mencionados (funciones, destinatarios y necesidades de información, momento de la política y tipo de evaluación) puede ser especialmente clarificadora y fructífera para analizar los procesos de utilización de las evaluaciones.

Además de a la diferenciación y clasificación de los posibles usos de la evaluación, la preocupación por la utilización también ha llevado a la **identificación de elementos prácticos facilitadores de dicha utilización en los procesos de evaluación**. Así, la literatura sobre **uso** y **utilización** de la evaluación (Alkin, 1985; *Joint Committee*, 1994; Patton, 1985, 1997; Shadish, Cook y Levinton, 1991) coincide en que los evaluadores deben:

- tener en cuenta el contexto específico de la evaluación
- identificar los destinatarios (audiencias) de la evaluación e implicarles (especialmente a clientes y agentes críticos *-stakeholders-*) en el diseño de la evaluación
- mantener un fuerte y frecuente contacto y aportar resultados intermedios a lo largo del proceso de evaluación
- ajustar los informes a las necesidades de los destinatarios, usando una variedad de formatos, incluyendo informes cortos y resúmenes, presentaciones orales y oportunidades para la interacción
- presentar ilustraciones reales y concretas de los hallazgos
- presentar los resultados con una temporalidad adecuada a una variedad de destinatarios
- usar un lenguaje simple y claro

Efectivamente, una respuesta típica a esta preocupación ha sido la atención prestada a los procesos de comunicación de los resultados de evaluación, incluyendo la importancia de la elaboración de un buen informe, legible y claro, y su seguimiento y de otras formas alternativas de comunicación. Un buen exponente de esta preocupación por el elemento comunicativo es la obra de Torres, Preskill y Piontek (1996). Una consecuencia de esto es el aumento de responsabilidades en el papel del evaluador: la responsabilidad del evaluador no acaba con la entrega del informe, sino que también tiene responsabilidades de comunicación, difusión y seguimiento de dicho informe. De hecho, tiene todo el sentido esta preocupación con los procesos de comunicación ya

que, según demuestra una encuesta realizada a evaluadores profesionales de la Asociación Americana de Evaluación (Torres, Preskill y Piontek, 1997), dichos profesionales estaban sólo moderadamente satisfechos con sus esfuerzos y resultados en la información y comunicación de resultados. Los dos impedimentos más mencionados para una buena información y comunicación tenían que ver con la falta de tiempo invertida en esta tarea, en primer lugar, y la complejidad política y organizativa en los procesos de evaluación, por otro. Las razones mencionadas son: falta de claridad entre los agentes críticos sobre sus necesidades de información, falta de respuesta ante los esfuerzos por comunicar e informar, cambio de cliente o destinatario de la evaluación en mitad del proceso, situaciones muy cargadas políticamente, resistencia ante los hallazgos negativos, características de individuos particulares e interpretación errónea de los hallazgos.

Otros autores señalan la necesidad de enfocar la totalidad del proceso de evaluación (y no sólo las fases finales) a la maximización de su uso o, al menos, considerar la cuestión del uso desde el principio (Weiss, 1998: 324). El máximo exponente es Patton, quien elabora un modelo teórico-práctico que se llama, precisamente “Evaluación orientada a la utilización” (*Utilization-Focused Evaluation*). Este modelo, además, se ha visto reforzado con los años. La primera edición de su libro con el mismo título es del año 1978. En 1997 publica una tercera edición introduciendo nuevos elementos y reforzando su enfoque. La evaluación orientada a la utilización se hace “con y para usuarios específicos e intencionados para usos específicos e intencionados” (Patton, 1997: 23). Es decir, todo el proceso de evaluación, desde su diseño hasta su comunicación está enfocado a responder las preguntas necesarias a determinados destinatarios identificados previamente. No sólo la utilización de la evaluación es su sentido último, sino que la concepción general de los procesos evaluativos debe estar subordinada a dicho fin.

Sin embargo, de forma paralela a esta preocupación por la utilización, la comunidad científica dedicada a la evaluación, se ha ido haciendo más consciente y realista sobre cuáles pueden ser, realmente, los usos de las evaluaciones. De esta manera:

- En la toma de decisiones sobre un programa o política, la evaluación es sólo uno de los factores a tener en cuenta. Muchas otras variables, además de unos buenos y útiles resultados de evaluación, intervienen en la toma de decisiones. En primer lugar, los resultados de la evaluación deben ser aceptados como correctos y útiles. Después, hay que reconocer que van a interactuar con lo que la gente ya cree sobre el programa, sus valores y creencias sobre lo bueno y lo malo, sus propios intereses y los intereses de la organización o estructura que lance el programa, etc. (Weiss, 1998: 324).
- La utilización no siempre es directa e inmediata. Esta utilización directa e inmediata sí puede darse en ocasiones, sobre todo cuando hay buenas condiciones (el tema no es polémico, implica pocos recursos o propone cambios para los que existe un cierto consenso). Pero lo normal es que no sea así, y que los cambios, aunque se den, sean más sutiles y graduales. Lo importante aquí es que, aunque no se dé un uso directo, la evaluación afecta la manera en la que la “comunidad de programa o política” (conjunto de personas afectadas o implicadas en un

programa o política) ve el mundo, es decir, la hace más consciente (Weiss, 1998: 325).

- El uso de la evaluación puede ser simbólico (Palumbo, 1987: 25). Cómo y por qué las organizaciones usan las evaluaciones es parte de cómo se utiliza la información en general. La información se utiliza, sobre todo, con un valor simbólico; los gestores de programas y políticas tienden a recopilar más información de la que necesitan y si la información es favorable tiene mayores probabilidades de ser utilizada que si no lo es.
- Chelimsky (1998: 50) hace una reflexión sobre la necesidad de reconsiderar el concepto de uso. Para ella, la justificación de la evaluación debe ser menor en cuanto al uso que se haga de sus resultados (aunque éste sea importante para la mejora de la formulación de políticas) que en cuanto a su éxito como proveedora de información de interés público. Es decir, que además de los intereses particulares y específicos de los destinatarios específicos de una evaluación, se puede y debe pensar en el interés público y general.

Hacerse conscientes de esta realidad ha derivado en una **visión más flexible de la utilización de los resultados de la evaluación**. Se ha pasado de pensar que el uso de la evaluación consistía en el uso de los resultados para tomar decisiones a ser consciente de unas mayores dimensiones en el concepto de uso. La evaluación no sólo puede y debe servir para facilitar la toma de decisiones concreta respecto a un programa. También puede contribuir al aprendizaje organizativo y de los grupos humanos, y su utilidad puede ir más allá del programa o política concreta que evalúa, no sólo respecto al ámbito o campo que toque, sino también respecto a los propios procesos de evaluación. Cada proceso de evaluación puede y debe contribuir y ofrecer una información útil sobre cómo evaluar determinadas políticas en determinados contextos, qué enfoques, perspectivas, metodologías y técnicas han sido adecuadas, etc.

En definitiva, se puede pensar en el concepto de **uso de la evaluación** en un sentido amplio. A este respecto, nos parece clara la definición que hacen Torres, Preskill y Piontek (1996: 2). Para ellas, el **uso** en evaluación es cualquier consecuencia o resultado de una evaluación e incluye:

- a) Los efectos de los mismos procesos de evaluación, como una mayor comprensión y/o apreciación de la evaluación.
- b) Una mayor comprensión o nuevas perspectivas en una situación, programa u organización.
- c) Las acciones que se llevan a cabo como resultado de una evaluación.

En la misma línea, para Carol H. Weiss (1998b), bajo esta visión ampliada sobre los usos de la evaluación, se puede hablar de cuatro tipos de uso:

- 1) Uso instrumental: se usan los resultados para tomar decisiones.
- 2) Uso conceptual: se utiliza la evaluación para cambiar la comprensión de lo que el programa o política es, hace o pretende.

3) Uso persuasivo: la evaluación sirve para movilizar apoyos para conseguir cambios necesitados en el programa.

4) Influencia en otras instituciones y eventos, más allá del programa o política evaluada. Las conclusiones de una evaluación pueden influir en las redes de política (*policy networks*) y en los profesionales, alterar paradigmas de política, y producir cambios en la agenda política.

De esta manera, para Weiss (1998b: 27), además de utilizarse los resultados, hallazgos y recomendaciones de un estudio de evaluación, se usan las ideas y generalizaciones extraídas de la evaluación, el mismo hecho de evaluar (ya que la gente en torno de un programa se hace más consciente del mismo), el foco del estudio y las medidas utilizadas, así como el propio diseño de investigación utilizado en el estudio de evaluación (ya que puede influir en las operaciones de programa).

Nos resulta especialmente interesante insistir en conceptualizar como uso de la evaluación los efectos que se pueden producir por **el propio hecho de evaluar**. Ya hace años se describió el efecto *Hawthorne*, por el que se demostró que las personas estudiadas que sabían estaban siendo observadas realizaban la tarea significativamente mejor que las que no eran conscientes de estar siendo observadas⁶. Haciendo referencia a las consecuencias positivas este efecto, se sabe que las personas que participan en un proceso de evaluación tienen mayores probabilidades de hacerse más conscientes del programa o política y sus objetivos. Como Greene (1988) ha señalado, los profesionales que participan en definir y diseñar un proceso de evaluación, comienzan a pensar con mayor profundidad sobre lo que se está tratando de conseguir. Puede que esto incluso lleve a una mayor cohesión del equipo de profesionales a cargo del programa. En definitiva, la evaluación es reflexión y revisión; el mero hecho de dar un espacio formalizado para ello tiene efectos positivos en sí mismos. Otros efectos positivos del hecho de llevar a cabo una evaluación pueden ser introducir una cultura de rendimiento de cuentas o legitimar el programa en cuanto a que se manda un mensaje sobre que se están tomando las cosas en serio y hay algo que aprender del mismo. Dependiendo del enfoque adoptado para la realización de la evaluación, se pueden obtener otros efectos colaterales. Por ejemplo, con los enfoques participativos se ayuda a las personas en torno al programa evaluado a reflexionar sobre la práctica, a pensar de forma crítica y a preguntarse por qué el programa funciona como lo hace. Con una evaluación basada en la teoría, los profesionales deben analizar y hacerse conscientes de las premisas subyacentes a su trabajo.

⁶ Este efecto se describió por primera vez en 1939 por dos investigadores (Roethlisberger y Dickson, 39) que pretendían hallar la intensidad de luz óptima para la productividad de unas obreras que trabajaban con piezas eléctricas pequeñas, en la planta de la *Western Electric Company* en *Hawthorne*. Constataron que cualquier cambio en la intensidad de la luz producía un aumento de la productividad. Concluyeron que el hecho de pertenecer al grupo experimental y ser observadas significaba una señal de atención personalizada de la empresa hacia ellas o hacia su bienestar personal.

CAPÍTULO 2

LA EVALUACIÓN DE POLÍTICAS PÚBLICAS DESDE UNA PERSPECTIVA HISTÓRICA

INTRODUCCIÓN

Aunque los orígenes de la evaluación de programas y políticas se sitúan en la década de los años cuarenta con la aparición de los trabajos de Ralph Tyler, el desarrollo de la misma como campo científico y profesional diferenciado es relativamente reciente, ya que se ha producido realmente en los últimos treinta años. A pesar de esta juventud, la evaluación ha sufrido una evolución acorde con el contexto histórico en el que se ha desarrollado. Por lo tanto, se pueden reconstruir históricamente las propuestas realizadas por parte de cada autor e identificar algunos factores que han influido en el desarrollo histórico de la evaluación de programas, como, por ejemplo, el grado de institucionalización y consolidación profesional de la misma, que es distinto según los países, o la propia evolución en el estudio y la concepción de los problemas sociales. Por lo tanto, en este segundo capítulo, se realiza una **breve historia** de la evaluación y se identifican los factores que han influido en su desarrollo y evolución. Además, se exponen los **enfoques** más destacables que se han ido proponiendo a lo largo del desarrollo de la evaluación como campo científico diferenciado. De esta manera, se analizan las cuestiones más relevantes de cada enfoque o autor, destacando los elementos que han sido más utilizados en la teoría y la práctica de la evaluación.

Una cuestión estrechamente relacionada con los enfoques y el desarrollo histórico de la evaluación es la referida a las diferentes **tipologías de evaluación** propuestas por cada autor o perspectiva. A lo largo de estos años de evolución histórica, se han establecido múltiples clasificaciones, identificándose diferentes tipos de evaluación posibles. Después de un análisis exhaustivo de las clasificaciones y tipologías disponibles, resaltamos, en este segundo capítulo, las que, desde nuestro punto de vista, aportan una mayor claridad conceptual y pueden traducirse de forma más operativa a la práctica de la evaluación de las políticas públicas. Además de las cuatro tipologías que se proponen aquí como “básicas”, también se mencionan otras clasificaciones utilizadas en la literatura especializada y se discuten otros tipos de evaluación no incluidos en las tipologías básicas anteriormente mencionadas, como son la evaluación de la evaluabilidad, y la metaevaluación.

También dependiendo del enfoque de evaluación propuesto, la definición de quiénes son los **agentes críticos** en la evaluación de un programa o política, así como cuál debe ser el **papel del evaluador**, pueden ser muy diferentes. Una de las características de los contextos evaluativos es la presencia de diferentes personas o colectivos que tienen algo que ver con el programa o la política a evaluar y que, por lo tanto, pueden estar implicados en o afectados por la evaluación. La definición de

quiénes son estos agentes críticos o *stakeholders* y, sobre todo, la inclusión o no de los mismos en el diseño de evaluación va a depender del enfoque que se adopte, siendo los últimos enfoques participativos y pluralistas los que hacen una definición más amplia de agentes críticos y los que consideran clave la participación de los mismos en la evaluación. En cualquier caso, cada enfoque suele dar prioridad a un tipo de destinatarios de la evaluación (patrocinadores, directores, gestores y profesionales, usuarios, comunidad política o científica, etc.). Asimismo se discute en este capítulo el diferente papel que puede cumplir el evaluador dependiendo también de los diferentes enfoques y estructuras institucionales en donde se produzca la evaluación. De esta manera, se describe el continuo cuyos extremos estarían representados, por un lado, por el evaluador distanciado o despegado y, por el otro, por el evaluador comprometido o apegado, y se identifican los diferentes tipos de evaluadores según se encuentren en el organigrama o *staff* de las organizaciones públicas, trabajen en una organización externa especializada en evaluación o formen parte de un proyecto de evaluación.

Un tema de especial auge hoy en día desde todos y cada uno de los enfoques de evaluación propuestos son las posibles **implicaciones éticas** relacionadas con la práctica de la evaluación. Efectivamente, en los últimos años se le ha prestado mucha atención a las cuestiones éticas relacionadas con la evaluación. Esta preocupación, en cierta lógica, ha sido paralela al reconocimiento del carácter político de la evaluación - y, por lo tanto, al poder que ésta pueda tener- y a la necesidad de que el evaluador interactúe en el contexto en el que debe intervenir. En este capítulo se analizan las cuestiones éticas que se tratan de forma creciente en los manuales de evaluación, así como la aparición y desarrollo de códigos deontológico para evaluadores.

Finalmente, exponemos en el último punto de este capítulo, los aspectos que, desde nuestro punto de vista, y una vez revisados los diferentes elementos que componen el concepto de evaluación, son más importantes en la construcción de un posible modelo general para la evaluación de las políticas públicas. El objetivo de este apartado es ofrecer una **visión general e integradora** de las cuestiones y elementos clave en la evaluación de los programas y las políticas públicas.

1. LA HISTORIA DE LA EVALUACIÓN Y LOS DIFERENTES MODELOS O ENFOQUES PROPUESTOS A LO LARGO DE LA MISMA

El estudio de los diferentes enfoques o modelos de evaluación propuestos a lo largo de la historia de la evaluación, ofrece una idea no sólo de las distintas propuestas de los autores clave, sino también de la propia evolución de la evaluación como actividad científica diferenciada. Esta evolución, además, no es ajena al desarrollo del propio análisis de políticas públicas en general. Por ejemplo, existe un claro paralelismo entre la superación de los modelos racionales en el *policy analysis* (con la aparición de los modelos incrementalistas) a partir de los años setenta y los enfoques de evaluación que reconocen y asumen la parte intuitiva de la misma.

En este apartado se pone en contexto la evaluación como actividad científica diferenciada, realizando una breve historia de la evaluación y estudiando los factores que inciden en su evolución. Esto permite entender mejor las respuestas que los especialistas han ido dando, proponiendo diferentes enfoques o modelos para entender y practicar la evaluación de programas. Posteriormente, se exponen las clasificaciones de modelos propuestas por diferentes autores y el sentido de su estudio y, finalmente, se discuten, de una forma sucinta y a grandes líneas, los modelos o enfoques más importantes en evaluación de programas y políticas.

1.1. Breve historia de la evaluación

La historia de la evaluación de programas es relativamente reciente, ya que se puede hablar de un campo que se ha desarrollado, en realidad, en los últimos treinta años. Cuando se estudian los orígenes de la evaluación de programas, hay que tener en cuenta dos factores. En primer lugar, los intentos de utilizar los datos que proporcionaban las investigaciones para comprender mejor los problemas públicos y la conducta social y para formular políticas públicas más inteligentes; en general tiene que ver, con la idea de utilizar la ciencia para un mejor gobierno y dirección de los asuntos públicos. En segundo lugar, también influyeron los intentos de medir el aprendizaje, la conducta y otros factores, como la inteligencia y la personalidad, de los individuos, a efectos de clasificación educativa (para la selección de alumnos para determinadas fases o secciones de estudio) y psicológica (por ejemplo, para la selección de futuros trabajadores). De hecho, muchos de los que luego serían especialistas en evaluación de programas, venían del mundo educativo y de la psicología, de las ramas del diagnóstico (se pasa del interés de medir determinadas características o conductas de los individuos a medir programas).

En una breve historia de la evaluación, Weiss (Weiss, 1998: 10-15) menciona algunos estudios en el siglo XIX que están en el origen de la evaluación de programas, tal y como la conocemos hoy en día, y que podrían ser etiquetados como evaluativos. En 1833, un francés, Guerry, publica un estudio estadístico que intentaba demostrar que la educación no reducía el crimen. Poco después, en 1844, un colega suyo, Deputit trata de establecer la utilidad de las infraestructuras públicas (carreteras, canales, etc.), haciendo cálculos sobre hasta cuánto estaría dispuesta a pagar la gente por dichas infraestructuras. En general, a finales del siglo XIX y principios del XX, se utilizaba la ciencia social para documentar la dimensión de los problemas y para identificar a la gente afectada por dichos problemas. Entonces, estudiar el cómo se abordaban los problemas y el éxito de dicho abordaje no tenía sentido ya que se suponía que los remedios que se establecían iban a resolver los problemas, es decir, no se ponía en duda que no fueran a hacerlo.

Los estudiosos en los ámbitos de la educación y la salud son los primeros que empiezan a realizar estudios sistemáticos sobre los resultados de su trabajo. En EE.UU., Cabot en 1912 examina 3000 informes de autopsia y los compara con los diagnósticos realizados previamente para evaluar la calidad de la diagnosis médica. En 1914, Codman intenta evaluar el desempeño de los cirujanos a través de la medición del estado de los pacientes después de dejar el hospital. Uno de los estudios más famosos en educación en aquellos años (1933) fue el "*Eight-Year Study*", dirigido por Ralph Tyler y en el que se comparaban los resultados en 15 institutos de secundaria progresistas y 15 tradicionales. También en 1934, Palmer, de la Asociación Americana de Salud Infantil realizó un estudio sobre la efectividad de programas de salud en las escuelas.

Seguimos haciendo una breve historia de la evaluación, tomando sobre todo como referencia a EE.UU., ya que éste ha sido el país donde no sólo se comienza a desarrollar la evaluación como actividad científica diferenciada, sino también uno de los países en los que existe un mayor grado de institucionalización de la evaluación, como veremos más adelante.

En los años 40, algunas fundaciones privadas comienzan a financiar evaluaciones de programas sociales innovadores que ellas mismas habían puesto en marcha. Un estudio famoso fue el del programa Cambridge-Somerville de jóvenes trabajadores para prevenir la delincuencia. En los años 50, el Gobierno Federal reacciona, ante el lanzamiento del *Sputnik* soviético y su consiguiente miedo a un posible analfabetismo científico de los norteamericanos, financiando nuevos currículos educativos (como el *Harvard Project Physics*). Como consecuencia se realizan evaluaciones para comprobar el éxito de dichos currículos.

La *Guerra Contra la Pobreza* de mediados de los años 60, supone el comienzo de las grandes evaluaciones financiadas federalmente en Estados Unidos. También se dan evaluaciones a gran escala en el nivel internacional, evaluándose, sobre todo, programas de planificación familiar en Asia, de nutrición y cuidado de la salud en América Latina y de desarrollo agrícola en Africa. También en EE.UU se desarrolla el análisis coste-beneficio (sobre todo, a través de la RAND Corporation y el Departamento de Defensa) y otros métodos de análisis y evaluación económica. Como consecuencia de esta actividad, se comienzan a desarrollar todo tipo de métodos e

instrumentos de evaluación. Estos años son, además, años de gran optimismo y fe en la ciencia.

El desarrollo y la institucionalización de la evaluación en EE.UU. es una consecuencia de las grandes sumas de dinero federal para esta actividad. Ya en los años 70, la evaluación era común en todas las agencias federales norteamericanas. Se establecen empresas y consultoras para llevar a cabo las evaluaciones financiadas federalmente, los centros e institutos de investigación universitarios asumen también tareas de evaluación y los institutos de investigación y grandes empresas ya establecidos se suman a la tarea evaluativa. Como comenta Weiss, en aquellos años, “muchos investigadores reinstrumentalizan sus habilidades para aprovecharse de las nuevas corrientes de dinero” (Weiss, 1998: 13). Un punto álgido para la evaluación fue la realización, en estos años, de una serie de experimentos sociales para examinar ideas y políticas innovadoras previamente a su aprobación (uno de los más famosos es el *Negative Income Tax Experiment*). Se ponían en marcha experiencias piloto lo suficientemente grandes como para que se dieran las condiciones reales; así se suponía que los resultados experimentales ayudarían a los formuladores de políticas a decidir si se debían generalizar dichos programas. En aquel entonces, resultó, sin embargo, que en el momento en el que estuvieron listos los resultados (eran grandes experimentos, que requerían tiempo para su realización), el clima político había cambiado, y las evaluaciones no resultaron tan útiles como se había pensado en un principio.

En EE.UU. la evaluación sigue creciendo hasta los años 80, en los que la administración de Reagan corta drásticamente la financiación para las nuevas iniciativas sociales y los programas innovadores en el nivel federal (que era el tipo de programas que generaba mayor movimiento de evaluación). Hacia mediados de los 80, y dados estos recortes en la financiación disponible para los estudios de la evaluación, el panorama era pesimista: en la introducción al *Evaluation Studies Review Annual* del año 1986, sus editores señalaban que había varios signos de que la evaluación, como actividad científica diferenciada, se estaba “desinstitucionalizando” (Cordray y Lipsey, 1987: 32). Aunque en los 90 se recupera parte de esa financiación (U.S. General Accounting Office, 1987), existe una tendencia general al “adelgazamiento” del Estado Federal, y las políticas sociales (y por lo tanto, su evaluación) se han situado más al nivel central y local.

Finalmente, resulta interesante hacer un resumen de cuáles son las tendencias en la evolución de la evaluación de programas. Weiss (1998: 14) identifica cuatro tendencias recientes en la evaluación de programas y políticas:

- Existe una tendencia hacia estudios de evaluación con dimensiones menores. Es decir, se pasa de los grandes estudios a pequeños estudios de evaluación.
- Otra tendencia es hacia estudios que utilizan en mayor medida metodologías de tipo cualitativo. Es decir se pasa de lo cuantitativo y experimental a lo cualitativo, los estudios de caso y las metodologías participativas. En cualquier caso cada vez hay una mayor aceptación de formas de investigación que no estaban tan legitimadas anteriormente, y una mayor flexibilidad en la utilización de diferentes enfoques

epistemológicos, métodos y técnicas (Yin, 1997). En los últimos años, se ha insistido, precisamente, en los beneficios y la conveniencia de utilizar métodos mixtos (Greene y Caracelli, 1997).

- En EE. UU. se ha pasado de una predominancia de los estudios de evaluación en el nivel federal a una mayor localización de la actividad de evaluación en los niveles central y local.
- En los últimos años se han creado y desarrollado asociaciones profesionales de evaluación. Esto ha contribuido a un desarrollo paralelo de la profesión de evaluador.

La consideración de estas tendencias generales es una buena base para la comprensión de la propia evolución de las diferentes propuestas y enfoques que se han ido proponiendo. En definitiva, dadas estas tendencias es explicable, por ejemplo, la proliferación de enfoques de tipo pluralista y participativo y la atenuación de los modelos experimentales.

1.2. Factores que influyen en la evolución de la evaluación y en la propuesta de modelos o enfoques: grado de institucionalización y consolidación profesional y evolución en la concepción y estudio de los problemas sociales

Uno de los factores que ha influido claramente en el desarrollo de la evaluación ha sido el **grado de institucionalización y consolidación profesional** de la misma, que ha sido diferente en cada país (Rist, 1990; Ballart, 1993).

La evaluación se produce principalmente en los países en los que existe un desarrollo importante de políticas de bienestar social (EE.UU., Canadá, Alemania, Suecia y Gran Bretaña) y en los organismos internacionales (en general en la ONU y especialmente en el Banco Mundial, OMS, FAO y UNESCO) especialmente en aquellos dedicados a programas de ayuda internacional (Ballart, 1993: 209). El país con un mayor nivel de institucionalización es, sin duda, EE.UU., en donde existe una política de evaluación sistemática de las políticas y los programas públicos y en donde la función de evaluación está normalizada en el sistema político-administrativo a través de unidades especializadas, tanto en el nivel ejecutivo (existen estructuras de evaluación en los departamentos y agencias públicas que gestionan programas sociales), como en el nivel legislativo (*General Accounting Office, Congressional Research Service, Congressional Budget Office y Office of Technological Assessment*). En los países europeos se produce una evaluación *ad-hoc*, en relación con demandas específicas y no alcanza todavía el carácter de una función asumida como propia por las instituciones (Ballart, 1993: 209-211). En el caso concreto español en el que el grado de institucionalización es muy bajo, Ballart (Ballart, 1993: 221-2) señala cuáles son los posibles factores del retraso:

- 1) la incorporación de España a las nuevas políticas de intervención social llega en momentos de crisis económica,
- 2) el sistema parlamentario español – en los niveles central y autonómico– tiende a limitar la función de control del Parlamento ya que se basa en una fuerte disciplina de partido,
- 3) el Tribunal de Cuentas y sus homólogos en el nivel autonómico limitan su papel al de supervisión de las cuentas públicas de acuerdo con las leyes de presupuestos y otras aplicables (aunque en teoría deberían examinar la eficacia y eficiencia de la actuación de los organismos públicos),
- 4) el número de profesionales con habilidades evaluativas es escaso (tampoco existen programas de formación específicos para ello)¹, y
- 5) a pesar del avance en estos últimos años, los sistemas de formación internos son débiles, la disponibilidad de datos en general todavía es escasa y las bases de datos no se han construido desde una perspectiva de evaluación.

Uno de los efectos de un alto grado de institucionalización es, lógicamente, un mayor grado de consolidación profesional. La consolidación profesional de una disciplina se puede medir a través de una serie de factores como el desarrollo de programas académicos superiores dedicados a la formación de profesionales, la edición de publicaciones periódicas especializadas, la existencia de asociaciones profesionales y la presencia de un mercado de trabajo con un cierto “volumen de negocio” con empresas, agencias consultoras e institutos de investigación, además de profesionales, dedicados y pagados por ello (Ballart, 1993: 211; Bustelo, 1998: 152). Paralelamente al grado de institucionalización, la consolidación profesional es también especialmente elevada en EE.UU., aunque empieza a haber algunos signos de consolidación en Europa en los últimos años (creación de asociaciones de evaluación nacionales -como la británica, la alemana y la italiana- y europeas -Sociedad Europea de Evaluación- y la publicación de algunas revistas especializadas -como por ejemplo, *Evaluation*).

El segundo factor que influye en la evolución de la evaluación como actividad científica diferenciada y en la propuesta de modelos y enfoques es la propia **evolución en el estudio y la concepción de los problemas sociales**.

A principios de la década de los años sesenta, existía, en EE.UU. una gran confianza en la capacidad del gobierno para afrontar los problemas públicos. Los liberales pedían reformas de acuerdo con el principio de igualdad y creían que el Gobierno Federal debía ser el principal artífice de las mismas. En 1964, con el ascenso de Johnson al poder y la mayoría demócrata en las dos cámaras, se lanzan los grandes

¹ Comienza a haber algunas tímidas excepciones. En la Universidad Complutense, por ejemplo, se impartió durante los cursos 1998/99 y 1999/2000 un Experto en Evaluación de Proyectos, Programas y Políticas de Cooperación al Desarrollo organizado por el Instituto Complutense de Desarrollo y Cooperación (IUDC-UCM). Este Experto, sin embargo, se ha reconvertido en el curso 2000/01 en un Experto en Gestión del Ciclo de Proyecto que, aunque incluye una parte importante de evaluación, ya no tiene el carácter específico de las dos primeras ediciones.

programas *Great Society* y *War on Poverty* que significan una amplia y audaz reforma social que fomenta mucha investigación y evaluación. Las medidas y programas se basaban, supuestamente, en decisiones racionales respaldadas científicamente. Como comenta Aguilar Villanueva “Este aire intelectual, técnico, de las políticas reformistas contribuyo a aumentar la confianza en la capacidad del gobierno. ‘Gran sociedad’ y ‘Guerra a la pobreza’ eran políticas no sólo legítimas, justas, sino correctas. (...) Esa actitud explicará la magnitud del escepticismo posterior en el gobierno” (Aguilar, 1996 (4ª): 25).

Efectivamente, en la década de los años setenta los resultados fueron desalentadores para los politólogos ya que fueron conscientes de que los problemas públicos rebasaban por mucho las políticas públicas. La “potencia” de las políticas públicas no era tanta como se esperaba para resolver problemas. Las investigaciones y evaluaciones realizadas en aquella época, incluyendo los pesimistas primeros estudios de implementación, fueron socavando la esperanza en la capacidad del gobierno como el agente de la gran reforma social, además de la credibilidad en las posibilidades de la ciencia.

(Los estudios de implementación) demostraban que la promulgación de leyes buenas o el buen diseño (científico-tecnológico) de un programa gubernamental no eran condiciones suficientes para el éxito efectivo de una política (...). Más aún. el proceso de implementación fue también la ocasión para descubrir la inconsistencia de muchas relaciones causales espurias que la legislación y el programa de gobierno habían incorporado, confiando en los teoremas de la economía, la sociología, la pedagogía, la ciencia política, el urbanismo. Las conexiones causales entre fenómenos, contempladas por teorías, modelos y tecnologías, o simplemente no sucedieron o tuvieron lugar en formas imprevistas, con un cortejo de consecuencias secundarias indeseadas” (Aguilar, 1996 (4ª): 26-27).

Este proceso es paralelo a ser consciente de que las grandes evaluaciones de los grandes programas sociales, bajo un paradigma estrictamente positivista-experimentalista no resultaban útiles para la mejora de las políticas reales. ¿Qué había que evaluar pues? De esta manera, se comienzan a desarrollar enfoques que enfatizan la evaluación de los procesos, los contextos, los pequeños resultados, algo que ayudara a ir tanteando y comprendiendo mejor el abordaje posible y real de los problemas sociales y no tanto algo que pudiera dictaminar la resolución masiva, directa y final de los problemas que, además, estaba claro que no se conseguía. Esta pérdida general de la ingenuidad y la fe ciega en la aplicabilidad directa de la ciencia, así como el reconocimiento de la complejidad de los procesos políticos que intervienen e una política pública son también factores importantes para entender la evolución y el desarrollo de las propuestas teórico-prácticas en el campo de la evaluación.

1.3. Las diferentes clasificaciones de modelos o enfoques y el sentido del estudio de los mismos

Los modelos que aquí se reseñan, más que modelos teóricos son grandes enfoques o aproximaciones sobre cómo concebir y enfrentar la evaluación de programas o políticas; en muchos casos, se trata de propuestas teórico-prácticas para realizar estudios de evaluación. El sentido de su estudio y exposición en este trabajo estriba en ofrecer un marco general de los enfoques más importantes en evaluación, entender cuál ha sido la evolución de las diferentes propuestas y, finalmente, estar en condiciones de extraer conclusiones sobre cuáles enfoques pueden resultar los más adecuados para la evaluación de las políticas de igualdad de género.

Se han ofrecido diferentes clasificaciones de los modelos para su exposición. Así, Ballart (1992) señala dos grandes bloques: La evaluación clásica por objetivos y la evaluación pluralista. Monnier (1990-91 y 95) distingue entre cinco grandes familias de enfoques: el enfoque por los OBJETIVOS, por los EFECTOS, por la FORMA EN QUE EL PROGRAMA SE PUSO EN PRÁCTICA, por los PROCESOS y por los ACTORES SOCIALES. Vedung (1997) distingue tres grandes bloques: modelos de efectividad, económicos y profesional (de revisión de pares o colegas). Los modelos económicos los divide, a su vez, en modelos de productividad y de eficiencia (coste-efectividad y coste-eficiencia) y los de efectividad los divide según el elemento en el que hacen más énfasis: metas, resultados, cliente y participación (ver tablas).

Una de los trabajos más interesantes respecto a la exposición de diferentes modelos o enfoques en evaluación, es la que realizan Shadish, Cook y Levinton (1991) en su obra *Foundations of Program Evaluation*, ya que basan dicha clasificación en los fundamentos teóricos de cada uno de los autores propuestos (Scriven, Campbell, Weiss, Wholey, Stake, Cronbach y Rossi). Estudian a fondo estos autores y extraen sus “teorías” respecto a cinco aspectos o componentes de la evaluación de programas:

- la *programación social* (cómo se desarrollan, mejoran y cambian los programas y políticas sociales, especialmente lo relacionado con los problemas sociales),
- la *construcción de conocimiento* (la forma en que los investigadores aprenden sobre la acción social,
- la *valoración* (la forma en que los valores se adhieren a las descripciones de los programas),
- el *uso del conocimiento* (la forma en la que la información de las ciencias sociales, incluida la evaluación, se usa para modificar programas y políticas), y
- la *práctica de la evaluación* (las tácticas y estrategias que los evaluadores siguen en su trabajo profesional, especialmente dadas las limitaciones que encuentran.

Así, clasifican las teorías en: 1) Teorías del estadio uno, que son las que intentan traer la verdad a la resolución de los problemas sociales, y en las que incluyen a Scriven (“La ciencia de valorar”) y a Campbell (“El metodólogo de la sociedad

experimentadora”), 2) Teorías del estadio dos, que son las que generan alternativas que enfatizan el uso y el pragmatismo, y en las que incluyen a Weiss (“La relación entre evaluación e investigación de políticas”), a Wholey (“La evaluación para la mejora de los programas”) y a Stake (“La evaluación sensible y los métodos cualitativos”). 3) Por último, las teorías del estadio tres son las que intentan integrar el pasado, y en ellas incluyen a Cronbach (“El diseño de la evaluación funcional para un mundo de acomodación política”) y a Rossi (“Las evaluaciones extensas, a medida y dirigidas por la teoría. Un abanico de opciones”).

Otra de las clasificaciones clásicas es la realizada por Stufflebeam y Shinkfield (1987), en la que distinguen entre pseudoevaluaciones, cuasievaluaciones y evaluaciones verdaderas. Las pseudoevaluaciones representarían usos perversos de la evaluación (investigaciones encubiertas y de relaciones públicas), las cuasievaluaciones estarían determinadas por su visión parcial y por la ausencia del elemento valorativo, es decir, en ellas se contestan a una serie de preguntas pero no se evalúa el valor o mérito de algo. Los dos enfoques que los autores consideran cuasievaluaciones son las evaluaciones de objetivos y los estudios experimentales. Finalmente las evaluaciones verdaderas examinan el valor o mérito de algo y en su trabajo exponen los estudios orientados a la toma de decisiones, estudios centrados en el cliente, los estudios de políticas y los estudios orientados hacia el consumidor.

En nuestra exposición de modelos o enfoques se toma como base esta última clasificación, aunque se añaden otros modelos o enfoques propuestos en los últimos años. Para cada enfoque hablamos de los autores pioneros y los más representativos, es decir, los que han tenido una mayor influencia y son más citados en el ámbito de la evaluación de programas y políticas. De esta manera, proponemos exponer de forma sintética o sucinta las principales características de los siguientes enfoques:

- Enfoque de evaluación por objetivos
- Enfoques experimentales de evaluación
- Enfoques orientados a la toma de decisiones
- Enfoque centrado en el cliente
- Enfoque centrado en el consumidor
- Enfoque orientado a la utilización
- Enfoques participativos

Como ya se ha mencionado anteriormente, se exponen los modelos o enfoques que han tenido eco -y siguen teniendo en muchos casos- en la teoría y la práctica de la evaluación de programas y políticas. La exposición de enfoques sigue un cierto hilo histórico, explicándose en último lugar las propuestas más recientes. Para cada enfoque se presentan, de forma escueta, sus presupuestos básicos, sus contribuciones más representativas al campo de la evaluación, así como algunas de sus fortalezas y debilidades para la práctica evaluativa.

A. Enfoque de evaluación por objetivos

Este es el primer modelo que se propone en el ámbito de la evaluación de programas, y se debe a Ralph TYLER que desarrolla su modelo en 1942 (Smith y Tyler, 42), a partir del *Eight-Year Study*. A Tyler se le considera uno de los precursores en materia de evaluación (Monnier, 1990-1991) y el padre de la evaluación sistemática de programas (Stufflebeam y Shinkfield, 1987). Su influencia ha sido muy grande durante mucho tiempo y su modelo de evaluación, basado en los objetivos, domina el campo de la evaluación hasta bien avanzados los años 60 y extiende algunos coletazos hasta prácticamente nuestros días, en los que no es extraño encontrar algunos estudios de evaluación con este enfoque.

Tyler define la evaluación como el proceso en el cual se comprueba si existe una congruencia entre objetivos y resultados, es decir, su modelo pretende ofrecer un marco de análisis para saber hasta qué punto se cumplen los objetivos. Tyler desarrolla su modelo en el ámbito educativo y tiene mucha importancia entonces, ya que trasciende la concepción de la evaluación de los logros de los alumnos para proponer una evaluación del currículo educativo, concibiéndola como parte del desarrollo de la planificación educativa. Como ya se ha comentado, el modelo de Tyler va más allá del ámbito educativo y se convierte en el modelo dominante de la evaluación de programas en general. Durante mucho tiempo se sostuvo la idea de que evaluar un programa o política consistía en saber si se había cumplido y hasta qué punto, los objetivos propuestos en su formulación.

Este modelo presupone que existen objetivos claros, precisos y medibles con los que se pueden relacionar los efectos, considerados asimismo medibles. Una de las mayores críticas recibidas es que, la mayoría de las veces los objetivos de los programas y las políticas son generales o ambiguos y, en muchas ocasiones, incoherentes y hasta contradictorios. Este enfoque, además, pretende y necesita traducir en términos operativos, racionales y jerarquizados los objetivos como paso previo a su evaluación. Esto ha condicionado, durante mucho tiempo, el concepto de “evaluabilidad” de un programa o política y se han desarrollado mecanismos y técnicas para medir y aumentar dicha “evaluabilidad”². Incluso suponiendo que una cierta “evaluabilidad” existen una serie de limitaciones técnicas importantes, ya que no todos los objetivos son igualmente precisos, claros y medibles. Esto produce, y de hecho ha producido, un proceso de filtración en el se escogen los objetivos más fácilmente reducibles a términos observables y operativos, de manera que puedan ser medibles.

También presupone este enfoque que la fase de ejecución de un programa o política reproduce fielmente lo planificado previamente, cuando, sin embargo sabemos que la puesta en práctica de un programa difiere en menor o mayor medida y de forma inevitable del dispositivo inicialmente previsto (Monnier, 1990-1991: 133).

² Se trata en mayor profundidad esta cuestión de la evaluabilidad en el punto referente a tipos de evaluación.

Otra crítica posible a este modelo es que concibe la evaluación de una manera terminal, tendiendo a medir el éxito o fracaso final, y que da sólo una idea parcial del programa o política evaluada. Cuestiones sobre la adecuación y pertinencia de los objetivos propuestos y sobre la operatividad y funcionamiento de los programas, incluyendo la eficiencia con la que se gestionan los recursos quedan fuera de este modelo.

B. Enfoques experimentales de evaluación

El enfoque experimentalista de la evaluación se basa en la aplicación de las reglas de la investigación experimental que pretende contrastar hipótesis previamente formuladas y establecer lazos causales entre variables previamente definidas. El investigador-evaluador manipula y controla las variables, de forma que se puedan establecer dichas relaciones de causa-efecto.

Bajo este enfoque se propone la aplicación de diseños experimentales o, en su defecto, cuasi-experimentales. Una de las obras más importantes para este enfoque es la de Campbell y Stanley (63), no específica del campo de la evaluación sino de la metodología de las ciencias sociales en general. En esta obra, los autores proponen una serie de diseños cuasi-experimentales que permiten el control de las variables estudiadas (a base sobre todo de mediciones múltiples y utilización de grupos no equivalentes), aún cuando no se cuente con alguna de las condiciones necesarias para los estudios estrictamente experimentales (especialmente la asignación aleatoria de sujetos a los grupos experimentales y de control y la asignación aleatoria de tratamiento o no -experimental y control- a los grupos). La obra de Campbell y Stanley representa un hito importante, ya que “suaviza” las condiciones en las que se puede llevar a cabo una evaluación.

El autor más representativo de este enfoque es Edward Suchman y su enfoque científico de la evaluación. Suchman (67) distingue entre “evaluación” (proceso de juzgar el mérito o valor de alguna actividad) e “investigación evaluativa” (uso específico del método científico-experimental para realizar dicho proceso). Este autor se define claramente por la investigación evaluativa que, según él, no utiliza una metodología específica sino que debe aproximarse al máximo al método científico experimental. Aunque Suchman reconoce que en los estudios de evaluación hay que tener en cuenta una serie de consideraciones y problemas prácticos que no son tan importantes en investigación básica, cree que los evaluadores son básicamente investigadores que deben hacer un balance entre el método riguroso que deben aplicar y la situación práctica en la que deben actuar.

Para Suchman, la evaluación no mide directamente el fenómeno a estudiar sino índices del mismo, de manera que se pueden dar toda una serie de fuentes de variación sistemática. Por eso, los problemas de validez y fiabilidad son sumamente importantes y hay que intentar controlarlos. El autor distingue siete tipos de validez en el proceso

evaluativo (de proposición, de instrumento, de muestreo, de observador, de sujeto, de administración y de análisis) y cinco de fiabilidad (de sujeto, de observador, de situación, de instrumento y de procesamiento). También hace una clasificación y sistematiza las diferentes categorías de efectos: *esfuerzo* (cantidad y calidad de la actividad que se da en un programa), *realización* (resultados del esfuerzo), *adecuación de la realización* (si la realización ha sido adecuada a las necesidades), *eficiencia* (si hay una proporción en cuanto al esfuerzo) y *proceso* (explicaciones de las razones de los hallazgos).

Entre las ventajas de este enfoque se encuentra haber ampliado la concepción tyleriana de la eficacia, ya que también se propone determinar las razones de los éxitos y los fracasos y de descubrir los principios subyacentes a un programa exitoso. También establece relaciones causales y controla variables, de forma que es más difícil la manipulación externa. De hecho, para los defensores de este enfoque la evaluación experimental puede controlar y dejar de lado el contexto político de los programas y políticas evaluadas.

Entre las muchas críticas que ha recibido este enfoque, está la idea de que el método científico-experimental (o aproximaciones al mismo) no es aplicable en muchas situaciones evaluativas. Muchas veces por acercarse a este método se sacrifican aspectos importantes en la evaluación, como son, tener una visión parcial del programa evaluado y emplear mucho tiempo en la investigación, de manera que la información llega demasiado tarde para ser realmente útil. Efectivamente, este enfoque pretende el control de variables. En una situación tan compleja como es un programa o política pública, o bien sólo se estudia una pequeña parte o componente del mismo, haciendo caso omiso de los restantes, o bien se hacen grandes estudios que permitan el control de un número alto de variables. Este tipo de grandes estudios tiene una alta probabilidad de no resultar operativos ni útiles, ya que son muy caros y llevan mucho tiempo. Por otro lado, los defensores de este enfoque están más preocupados por la metodología de investigación, es decir, por conseguir un estudio fiable, válido y representativo que en que esa investigación evaluativa arroje una información útil para el programa evaluado. Ya se ha comentado anteriormente la conveniencia de distinguir entre evaluación e investigación, procesos que deben ser complementarios pero, desde nuestro punto de vista, claramente diferenciados.

C. Enfoques orientados a la toma de decisiones

Como reacción a los enfoques anteriores, una serie de autores desarrollan sus perspectivas de evaluación desde un punto de vista más pragmático, enfatizando el uso y la utilidad de la evaluación para la toma de decisiones en el programa o política

evaluada. Como representantes de este tipo de enfoques se pueden mencionar, entre otros, a Cronbach³, Stufflebeam y Weiss.

Una fortaleza de este enfoque es que anima a los profesionales y a las instituciones y organizaciones a usar la evaluación de forma continua y sistemática. La evaluación sirve realmente para la toma de decisiones porque se enfoca hacia dicho proceso. También representa una forma en la que dichos profesionales se hacen responsables y pueden explicar las decisiones tomadas en el desempeño de su trabajo. Sin embargo, una posible limitación es que la colaboración que se requiere entre evaluador y tomador de decisiones puede hacer que haya resultados sesgados.

Lee J. *Cronbach* ha sido uno de los autores más prolijos en el campo de la evaluación. Se puede considerar que marca un hito en el momento en el que publica en 1963 un artículo (*Course improvement through evaluation*) y en el que expone que el enfoque tyleriano no es suficiente. Al igual que Tyler, Cronbach se inicia a través de la evaluación educativa aunque a lo largo de los años expande su filosofía a la evaluación de programas en general. Buenos exponentes de la misma son sus obras de 1980 (*Toward the Reform of Program Evaluation*) y de 1982 (*Designing Evaluations of Educational and Social Programs*).

Ya en 1963, Cronbach señala que la evaluación sirve, sobre todo, para mejorar los programas y las asignaturas. De esta manera la evaluación se convierte en algo fundamental para el desarrollo de los currículos educativos, siendo la herramienta que permite recoger y utilizar información que ayuda a tomar decisiones sobre los programas educativos. En su versión posterior, la evaluación es un elemento necesario para la mejora de los programas llegando a ser una instrumento que, en última instancia, ayuda a desarrollar las políticas sociales y educativas. También introduce en dicho artículo la importancia de una evaluación *formativa* que describa (y no mida necesariamente) los cambios y los resultados del programa, se fije en los efectos colaterales e identifique los aspectos que necesitan revisión.

Para Cronbach, la evaluación debe estar bien planificada para que resulte provechosa y el mejor diseño es el que promete aumentar el beneficio social a partir de la evaluación. También debe tener una capacidad de respuesta flexible ya que las preocupaciones de la comunidad a la que sirve son cambiantes. Por esta razón, aunque la evaluación ha de ser realista, científica y razonada, debe también penetrar en el pensamiento de la comunidad y de los interesados en la evaluación. El evaluador es, por lo tanto, una especie de profesor que ayuda a hacer y contestar preguntas buenas y apropiadas a cada situación; por lo tanto, trabaja en equipo con los responsables del programa.

Cronbach trata de integrar los enfoques “científico” y naturalista. Para él, la evaluación da lugar a todo tipo de investigación y se pueden utilizar ambos enfoques, dependiendo de la naturaleza del estudio. Más que una elección de enfoque, propone

³ Shadish, Cook y Levinton (91) clasifican a Cronbach en las teorías del estadio tres que son las integradoras. Sin embargo, el primer Cronbach (63) insiste en la idea de la mejora de los programas a través de la evaluación.

dar el balance adecuado entre los dos estilos a cada situación evaluativa o a cada pregunta de evaluación. Para seleccionar las preguntas de evaluación adecuadas propone una primera fase divergente, de apertura y una posterior fase convergente en la que se establecen prioridades. Para el establecimiento de prioridades propone guiarse por dos factores: el grado de incertidumbre de una pregunta (a mayor incertidumbre respecto a la respuesta, más prioridad se debe otorgar a la pregunta) y el grado de influencia (que define como el peso relativo que tiene dicha pregunta en el proceso de toma de decisiones que afectan al programa evaluado).

La evaluación debe indicar lo que se espera si un determinado plan de intervención se adopta en un cierto tipo de situación. Así, las conclusiones pueden llegar a ser predicciones sobre la reacción de los clientes, el programa, el cambio institucional, etc. Atendiendo a este carácter predictivo de la evaluación, Cronbach define los siguientes elementos: *unidades* (individuos o clases sobre los que se realiza la investigación), *tratamientos* (cualquier tipo de acción), operaciones de *observación* (impresiones o datos que se recogen antes, durante y/o después del tratamiento, a través de las cuales se pone de relieve sus efectos en las unidades) y el contexto o *setting* en el que están inmersas las unidades. Utilizando estos cuatro elementos habla de “utos”, “UTOS” y “*UTOS”.

- El símbolo “utos” se refiere a la muestra de unidades, tratamientos, observaciones y contexto que participan en un estudio.
- “UTOS” se refiere a la población de unidades, tratamientos, observaciones y contextos realmente representada en la muestra observada. El diseño ideal es el que define UTO y selecciona un uto que lo representa. Así, U es la población y u la muestra; T es el tratamiento a seguir y t el que realmente se experimenta; O es el universo de todas las posibles observaciones y o son las que hemos realizado.
- el “*UTOS” se define como el campo de aplicación. La evaluación debe permitir la extrapolación de lo observado a las circunstancias de los *UTOS, que representan lo que interesa, lo aplicable. La generalización a los *UTOS es la función principal de la evaluación, permitiendo la transferencia de conocimiento a diferentes contextos, momentos y programas.

Finalmente, para Cronbach es importante la transmisión de los resultados de la evaluación. La utilidad de la misma depende del grado en el que las personas que toman decisiones asumen sus resultados y consideran posible la aplicación de las recomendaciones. Esto debe tenerse en cuenta desde los primeros pasos de la evaluación. El informe de los resultados no tiene por qué ser un informe escrito. Según Cronbach, el informe más potente puede ser una conversación informal. En cualquier caso es importante hacer informes para una comunidad amplia, resaltando los aspectos que interesan a cada uno de los sectores implicados.

Daniel L. *Stufflebeam* forma parte del Comité *Phi Delta Kappa* que se reúne en 1969 para evaluar el estado de la cuestión en evaluación educativa y ofrecer algunos consejos a la hora de realizar evaluaciones. Este Comité dictamina que la evaluación está “enferma” (escepticismo, falta de directrices, “síndrome de las diferencias no

significativas”, etc.) y decide que hay cinco cuestiones que hay que tener en cuenta: 1) hay que redefinir la evaluación, 2) la toma de decisiones está íntegramente relacionada con la evaluación, 3) la evaluación no puede estar exenta de valores, 4) existen diferentes niveles de evaluación porque existen diferentes clases de público que necesitan información distinta y 5) hay que distinguir entre evaluación e investigación.

La evaluación queda redefinida, entonces, como el proceso de delinear, obtener y ofrecer información descriptiva y valorativa sobre el valor (capacidad de respuesta a las necesidades) o mérito (calidad) de los objetivos de un programa, su diseño, ejecución e impacto para guiar la toma de decisiones, controlar responsabilidades y promover entendimiento sobre el fenómeno evaluado. Stufflebeam asume la redefinición de la evaluación como un proceso para ofrecer información que pueda resultar útil para la toma de decisiones, y desarrolla lo que él mismo bautiza como el *modelo CIPP*. Este modelo identifica 4 tipos de evaluación:

Evaluación del Contexto: pretende identificar el contexto institucional, los destinatarios de un programa o política y evaluar sus necesidades y problemas. Ayuda a desarrollar objetivos que sean coherentes con las necesidades y, en general, ayuda a tomar decisiones sobre metas y objetivos y a ofrecer una base para juzgar los resultados.

Evaluación del Input: su objetivo es valorar la capacidad del sistema para llevar a cabo las estrategias planteadas con los recursos dados. Identifica y valora el mérito de proyectos alternativos y ayuda a diseñar los programas y a dar forma a las propuestas y proyectos. Sirve para tomar decisiones sobre la estructuración del programa: identificación de fuentes de apoyo y diseño de estrategias y procedimientos. Ofrece una base para juzgar la ejecución o implementación del programa.

Evaluación del Proceso: pretende identificar y pronosticar, durante el proceso, los fallos de la planificación y describir y juzgar los aspectos del procedimiento con el que se está llevando a cabo el programa. Ayuda a guiar la realización del programa y sirve para tomar decisiones sobre la ejecución del mismo. Ofrece una base lógica sobre el proceso real y que servirá posteriormente para interpretar los resultados.

Evaluación del Producto: su principal objetivo es evaluar los resultados y relacionarlos con los objetivos y las informaciones sobre el contexto, el input y el proceso, para interpretar su valor y su mérito. Ayuda a tomar decisiones sobre continuación, terminación o modificación de los programas y a presentar información sobre los efectos de los mismos (intencionados o no, positivos y negativos).

Bajo una perspectiva sistémica, este modelo es una estrategia para la mejora de sistemas y evaluar significa obtener y utilizar *feedback* sistemáticamente, de manera que se convierte en una herramienta para ayudar a que los programas funcionen mejor para las personas a las que intentan servir. La evaluación cumple un papel crucial en la

estimulación y planificación de cambios en una institución (debe formar parte integral del funcionamiento cotidiano de la misma).

En la planificación del diseño de una evaluación hay que ser flexibles, ya que la dinámica de la misma hace difícil prever con exactitud cuáles van a ser las necesidades específicas de información. Esta contradicción (necesidad de planificación pero también de flexibilidad) es fácilmente resoluble asumiendo el diseño como un proceso en el que hacer revisiones periódicas para una readaptación y adecuación a la realidad del mismo.

Finalmente, Stufflebeam preside el *Joint Committee on Standards for Educational Evaluation* en sus primeros años y su influencia es crucial para la elaboración de las Normas para la Evaluación de Programas en su primera edición (*Joint Committee*, 1981), que marcan un hito para el desarrollo y consolidación de la evaluación como actividad científica diferenciada (Bustelo, 1999).

Las aportaciones de Carol H. *Weiss* al campo de la evaluación han sido enormes, sobre todo en lo que se refiere a entender la evaluación como una parte de la investigación y análisis de políticas. Ella es una de las primeras autoras en identificar y describir las dificultades políticas y organizativas que rodean la evaluación y en denunciar el inevitable carácter político de la misma.

En sus primeros años, Weiss adopta los métodos y estrategias experimentalistas de Campbell y Stanley (63). Pero con el tiempo, se aleja de dichas teorías idealistas y focalizadas en los efectos, porque, según ella, fallaban en la producción de resultados que pudieran ser realmente utilizados y en la apreciación de la realidad política y social de las políticas sociales (Shadish, Cook y Levinton, 1991: 181).

Entre sus aportaciones a la teoría de la evaluación de programas y políticas a lo largo de más de 30 años, se encuentran las siguientes (Shadish, Cook y Levinton, 1991):

Incompatibilidad entre las ciencias sociales y los programas sociales. El investigador de laboratorio puede controlar tratamientos experimentales y variables dependientes. Pero ni la investigación es el objetivo principal de los programas sociales ni los cambios realizados en los programas se realizan para hacer mejor o más fácil la investigación. Cuando entran en conflicto las necesidades de evaluación y las necesidades del programa, el programa tiene prioridad. Los evaluadores deben adaptar sus estándares metodológicos a la realidad del programa y obviar cuestiones que los profesionales del mismo no vean como legítimas. Para Weiss, la emergencia de métodos cualitativos ha conseguido remediar algunos problemas de los métodos experimentales, aunque aquellos también tengan sus límites, sobre todo la falta de credibilidad. Sin embargo, reconoce su flexibilidad y su capacidad dinámica y de demostración del mundo desde la perspectiva de otros. Ella recomienda a los evaluadores que utilicen tanto métodos cualitativos como cuantitativos.

El contexto político de los programas sociales. Weiss es una de las primeras autoras que señala y teoriza sobre el carácter político de la evaluación⁴. Para ella la evaluación es una tarea racional que se lleva a cabo en un contexto político. Las evaluaciones son resultado de decisiones políticas, se realizan para obtener información para el proceso de toma de decisiones y por lo tanto, entran de lleno en la arena política y, por último, las evaluaciones toman posiciones -que so políticas- sobre el programa o política evaluada.

Los límites de la solución de problemas sociales a corto plazo. En este asunto, Weiss rompe la ingenuidad de los primeros evaluadores que pensaban que los programas sociales cambiarían sustancialmente los problemas sociales. Los cambios en los programas, y más en los problemas, son lentos e incrementales.

Un contexto de múltiples agentes críticos (stakeholders) en los programas y políticas. Todos los programas tienen múltiples agentes críticos que tienen diferentes intereses y diferentes poderes para el uso de los resultados; enfrentan diferentes decisiones y, por lo tanto, tienen diferentes necesidades de información.

Una visión ampliada del uso de la evaluación. En su primera época, Weiss abogaba por un uso instrumental de la evaluación. A lo largo de su carrera va aumentando el concepto de uso, dándole un mayor énfasis a los usos indirectos, graduales y de iluminación general para acciones futuras y para una mayor comprensión de los problemas sociales⁵.

Finalmente, una de las últimas preocupaciones y desarrollos teóricos de Carol Weiss (1997, 1998) es la utilización, en evaluación, de la teoría de programas⁶. Por teoría de programas, Weiss entiende “el conjunto de creencias que subyace a la acción.(...) Es el conjunto de hipótesis sobre las que la gente construye y planifica los programas. Es una explicación de las relaciones causales que unen los inputs de un programa a sus resultados esperados o como dice Bickman (1987:5) ‘un modelo plausible y sensible sobre cómo un programa se supone que debe funcionar’ (Weiss, 1998a: 55). Basándose en estas ideas, Weiss desarrolla lo que ella llama las “teorías de cambio” de un programa: la “teoría del programa” y la “teoría de la implementación”. La teoría del programa “se refiere a los mecanismos que median entre la distribución (y entrega) de un programa y la aparición de los resultados. El mecanismo operativo de cambio no son las actividades del programa *per se* sino las respuestas que dichas actividades generan” (Weiss, 1998a:57). La teoría de la implementación tiene que ver con lo que se requiere para traducir los objetivos en la entrega de un servicio o en las operaciones de programa. Mientras que la teoría del programa pone énfasis en las respuestas de la gente a las actividades, la teoría de la implementación lo hace en los servicios que se ponen en marcha y cómo se ofrecen. Weiss explica la conveniencia y las ventajas de analizar y exponer estas “teorías del cambio” y de utilizarlas como una guía para la evaluación.

⁴ Su pensamiento sobre esta cuestión ha sido expuesto en el punto 3.2. del capítulo primero.

⁵ Para un mayor desarrollo de su teoría, véase el punto 3.4. del capítulo primero.

⁶ Weiss ha utilizado el enfoque dirido por la teoría (*Theory-Driven Approach*), que se podría considerar otro enfoque más (Chen, 1990, 1994; Chen y Rossi, 1983; Bickman, 1990). En este trabajo obviamos su referencia específica por abordarlo en las aportaciones de Weiss.

D. Enfoque centrado en el cliente

Los estudios orientados a la toma de decisiones suelen ir dirigidos en muchas ocasiones a directivos y responsables de los programas y políticas evaluadas, ya que en los niveles altos suelen tomarse más decisiones. En el caso del enfoque centrado en el cliente, el énfasis es en ayudar a la gente implicada en gestión cotidiana del programa. Las preguntas de evaluación se organizan en torno a las preocupaciones y cuestiones de los profesionales de base. El propósito de este enfoque es ayudar a dichos profesionales a entender el funcionamiento de las actividades de su servicio o programa y el grado en el que son valoradas por expertos y destinatarios del programa.

Como fortaleza de este enfoque se suele mencionar su enfoque de investigación-acción en la que la gente que ejecuta los programas es ayudada a dirigir su propia evaluación. Su máxima debilidad es, pues, su posible falta de credibilidad externa y el posible sesgo hacia los profesionales por el control que estos ejercen sobre el proceso evaluativo.

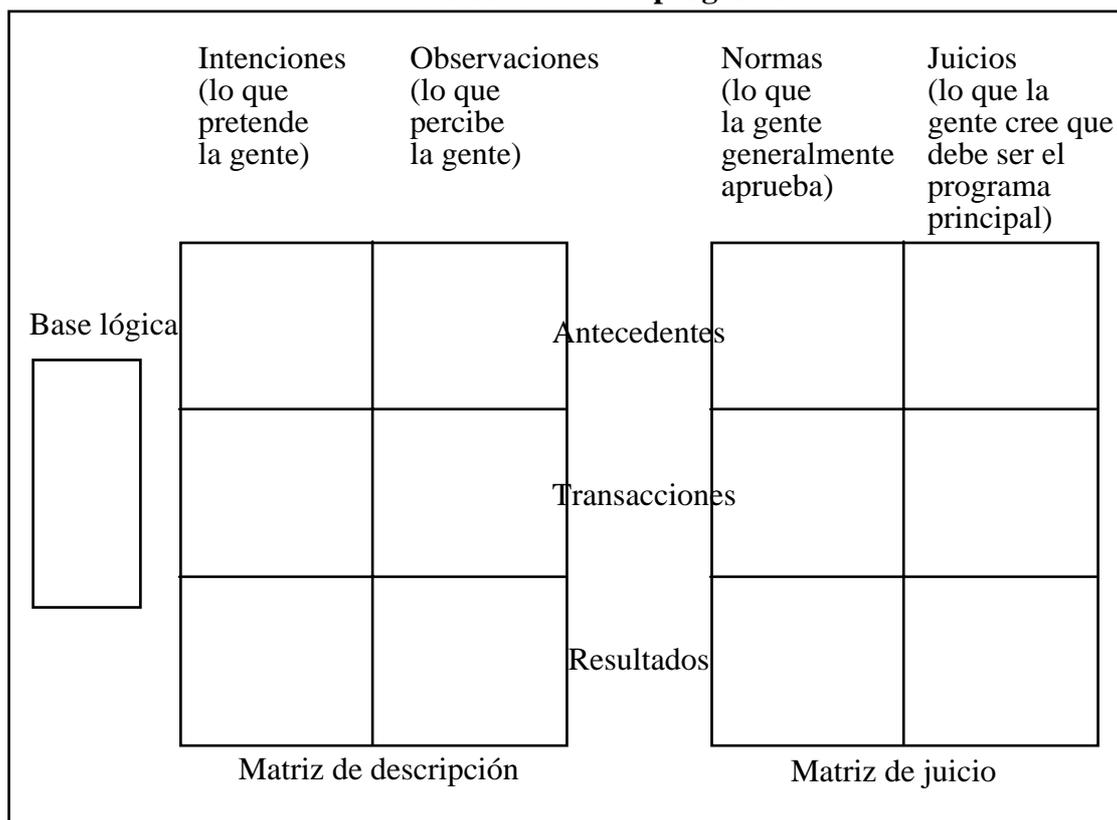
Robert *Stake* es el pionero de este enfoque. También pueden mencionarse a Barry MacDonald (75) en el Reino Unido. En realidad, el modelo propuesto por Stake es el precursor de los enfoques participativos que veremos más adelante, en los que se incluirán otros autores. Stake propone en 1967 el “Modelo de la Figura” (*The Countenance of Evaluation*) que posteriormente evoluciona a un modelo más flexible que denomina “Evaluación sensible” (*Responsive Evaluation*) (Stake, 76).

Con el “Modelo de la Figura” se pretende describir y juzgar el programa evaluado por completo:

- La *descripción* es el acto básico de la evaluación y hay que describir antecedentes (toda condición existente anterior al programa y que puede relacionarse con los resultados), operaciones o transacciones y resultados.
- El *juicio* es el segundo elemento de la evaluación y aunque los evaluadores no deben ser los que juzguen directamente, sí son los únicos cualificados para recopilar y procesar objetivamente las opiniones y los juicios de otras personas.
- El tercer elemento es el *esquema para la recopilación de datos*. (tabla 2.1). Hay que recopilar información sobre los antecedentes, las transacciones y los resultados en cuanto a dos ejes: lo supuesto, lo que se planifica o pretende (las *intenciones*) y lo real, lo que se da u observa (las *observaciones*). Además hay que investigar sobre la *base lógica* que representa los antecedentes e intenciones filosóficas del programa y que, una vez formulada, se puede utilizar para juzgar las intenciones del programa. Con estos datos se hacen dos tipos de análisis: de la congruencia (si lo que se pretendía realmente ocurrió) y de la contingencia (si las relaciones entre los elementos son lógicas y coherentes).

- Por último, estaría la *elaboración de normas* que son los criterios explícitos para valorar la calidad de un programa. Hay dos tipos de normas: absolutas (convicciones personales de lo que es bueno o deseable para un programa) y relativas (Características de programas similares que se consideran satisfactorios).

TABLA 2.1: Esquema de Stake para la recopilación de datos en la evaluación de un programa



Fuente: Stufflebeam y Shinkfield, 1987. *Evaluación sistemática*. Madrid: Paidós/MEC. p.244

El propio Stake reconoce que su modelo es limitado y difícil de llevar a cabo y propone la “Evaluación sensible”. En este segundo modelo es más importante la observación, la recopilación de juicios, la identificación de las necesidades del cliente y la preparación de informes informales que la preparación de instrumentos, la administración de tests y el procesamiento de datos. Pretende ofrecer un servicio a personas o grupos específicos, teniendo en cuenta la situación y el contexto.

Por último, Stake también ha contribuido al desarrollo de los estudios de caso como método especialmente utilizado en evaluación (Stake, 1995).

E. Enfoque centrado en el consumidor

Este enfoque enfatiza el punto de vista del “consumidor” de programas. Su propósito es el de juzgar servicios y bienes alternativos, ayudando a contribuyentes y profesionales a elegir más inteligentemente los bienes y servicios a consumir.

Este enfoque goza de mucha credibilidad entre grupos de consumidores. Sin embargo, puede ser tan independiente de los profesionales que puede no ayudarles a servir mejor a los consumidores. Además, se requieren expertos competentes y con una alta credibilidad y suficientes recursos para permitir al experto hacer su estudio (suele ser caro).

Michael *Scriven* es el máximo representante de este enfoque. Scriven es uno de los autores más activos en el campo de la evaluación y sus aportaciones a la campo de la evaluación son múltiples. Además de este enfoque desde el punto de vista del consumidor, Scriven ha desarrollado teóricamente el elemento valorativo de la evaluación. De hecho, Shadish, Cook y Levinton (1991: 73) caracterizan su aportación teórica como la “ciencia de la valoración”. Efectivamente, para Scriven evaluar es justamente la determinación del mérito o valor de lo que se evalúa (evaluando). Critica la idea de que la evaluación deba servir para ofrecer una información utilizable; para él, para lo que se utilice la evaluación es un asunto que no tiene que ver con la función de valoración de la evaluación.

Para Scriven, los informes para consumidores en los que se ofrecen estudios comparativos de productos en el mercado (*Consumer Reports*) son un ejemplo de evaluación. El fin último de la evaluación es satisfacer las necesidades de los consumidores; para esto, un evaluador externo e independiente debe emitir un juicio basándose en pruebas procedentes de la comparación de distintos programas. En esta línea, es muy importante que la evaluación esté libre de costes, es decir, debe arrojar un balance positivo coste-beneficio o, en otras palabras, debe merecer la pena claramente.

Una de las aportaciones más importantes de Scriven es su distinción entre evaluación *formativa* y evaluación *sumativa*. La primera ayuda a desarrollar el programa que pretende obtener información para la modificación del mismo mientras la intervención o actividad se está llevando a cabo. Se hace para ofrecer *feedback* a las personas que están intentando mejorar el programa evaluado. La evaluación sumativa, sin embargo, calcula el valor del programa una vez desarrollado y lanzado al mercado. Para Scriven, en este tipo de evaluación, el evaluador debe ser externo y los resultados de la evaluación deben hacerse públicos. Ambos tipos de evaluación son importantes para Scriven, aunque tiene una clara preferencia por la evaluación sumativa.

En el intento de evitar los sesgos y servir los intereses de los consumidores, Scriven propone la *Evaluación sin objetivos*. La filosofía subyacente a esta propuesta es la idea de que fijarse de antemano en los objetivos del programa puede condicionar su evaluación; a propósito se ignoran entonces las metas fijadas y se valoran los efectos (esperados o no) producidos por el programa. Esta evaluación es, según el autor, menos

intrusiva, más adaptable a los cambios repentinos de objetivos, menos propensa a la tendenciosidad social y más estimulante profesionalmente. Además puede ser reversible (se puede empezar por no considerar los objetivos para luego retomarlos) y complementaria (dos evaluadores pueden trabajar simultáneamente, uno con y otro sin objetivos). Al no comparar los resultados con los objetivos propuestos, Scriven propone hacer una evaluación de las necesidades de los consumidores.

Otras dos contribuciones de Scriven son su concepto e insistencia de *metaevaluación* (la evaluación debe ser evaluada) y la elaboración de su Lista de control de elementos clave en evaluación (*Key Evaluation Checklist*)⁷, que nace de la consideración *multidimensional* de la evaluación; la evaluación comporta múltiples dimensiones y por lo tanto, debe emplear diferentes perspectivas, introducir distintos niveles de valoración y utilizar una variedad de métodos. Propone que los evaluadores utilicen la lista de control para plantearse todos los aspectos, aunque no tienen por qué hacerlo necesariamente en la misma secuencia en la que se exponen.

F. Enfoque orientado a la utilización

El énfasis de este enfoque se pone en ofrecer una información útil y utilizable por parte de los agentes críticos y los responsables del programa. La utilización de los resultados de la evaluación es el elemento organizador de este enfoque. Como principal ventaja se encuentra el aumento de la probabilidad de uso de las evaluaciones, aunque se corre el riesgo de sesgar información por no considerarse relevante.

El representante de este enfoque es Michael Q. *Patton*, quien ha escrito múltiples libros en evaluación y se ha dedicado a la práctica de la evaluación durante más de 25 años. Aunque ha estado relacionado con el mundo académico (durante años fue profesor de la Universidad de Minnesota y director del Centro de Investigación Social de dicha universidad), se le puede considerar más como evaluador profesional que ha elaborado teóricamente sobre su experiencia práctica. Una de sus obras *-Utilization-Focused Evaluation-* es la que expone el enfoque que aquí tratamos. Su primera edición data de 1978 y la tercera, con el nada modesto subtítulo de “El texto del nuevo siglo”, y en la que ha añadido partes sustanciales, es tan reciente como 1997. Su influencia es enorme en el campo de la evaluación: tiene una alta presencia en los foros profesionales y en las publicaciones especializadas y es uno de los autores multicitados en los textos sobre evaluación.

Patton (1978, 1997) cree que el mayor problema de los estudios de evaluación consiste en que los resultados de los mismos no se utiliza; es decir la mayoría no son útiles para los potenciales usuarios. La falta de utilización de las evaluaciones la enmarca Patton en la crisis general de la información: se ha avanzado muchísimo en la

⁷ Véase tabla 4.3. *Key Evaluation Checklist* de Scriven. Capítulo 4.

capacidad de generar, guardar, recuperar y comunicar instantáneamente la información, pero no tanto en estar al día, clasificarla, absorberla, procesarla y *usarla*. Su propuesta, entonces, la resume en el lema “uso pretendido por quienes lo pretenden” (*intended use by intended users*) y se enfoca hacia la potenciación de la utilización. Para esto, se tienen que dar una serie de premisas (Patton, 1997: 382-383):

- 1) La preocupación por la utilización debe ser la guía conductora de la evaluación.
- 2) Esta preocupación debe ser continua y se debe dar desde un principio.
- 3) El factor personal (intereses y compromisos personales) contribuye significativamente al uso.
- 4) Se debe hacer un análisis cuidadoso y profundo de los agentes críticos (*stakeholders*) y esto debe servir para identificar a los usuarios principales.
- 5) Las evaluaciones deben centrarse en algún aspecto, ya que debido a las limitaciones de recursos y tiempo es imposible que una evaluación preste atención a todas las cuestiones posibles. El camino más útil es centrarse en “el uso pretendido por quienes los pretenden”.
- 6) Centrarse en el uso pretendido requiere hacer elecciones deliberadas y profundas. Se puede juzgar el mérito o valor (evaluación sumativa), mejorar el programa (uso instrumental) o generar conocimiento (uso conceptual). Los usos del proceso de la evaluación también pueden ser distintos: fomento de una visión compartida, reforzar intervenciones, fomentar el compromiso de participantes o desarrollo de programas y organizaciones. Todos los usos son importantes y legítimos y el valor relativo de cada uno depende de la situación concreta. Los usos pueden cambiar y evolucionar a medida que el programa o política madura.
- 7) Las evaluaciones útiles se deben diseñar y adaptar según la situación. Las recetas estandarizadas no funcionan. El valor relativo de una utilización particular sólo puede ser juzgado en el contexto de un programa específico y de los intereses de los usuarios particulares.
- 8) Los usuarios principales deben estar implicados activa y personalmente en la toma de decisiones sobre la evaluación. Esa es la forma de fomentar su compromiso de uso.
- 9) La participación de calidad y no de cantidad es el objetivo. La cantidad de tiempo de interacción de grupo puede estar inversamente relacionada con la calidad. Por eso, los evaluadores bajo este enfoque deben ser buenos facilitadores de grupo.
- 10) La implicación de calidad de los usuarios de la evaluación resultará en evaluaciones de calidad.
- 11) Los evaluadores deben aumentar su credibilidad siendo activos (en la presentación del enfoque y los métodos que consideran adecuados), reactivos (al

escuchar atenta y respetuosamente a las preocupaciones de otros) y adaptativos (al buscar formas de diseñar evaluaciones que incorporen los intereses distintos, incluidos los suyos, y, a la vez, respeten estándares de calidad de la práctica profesional).

12) Los evaluadores tienen la responsabilidad de entrenar o formar a los usuarios y agentes críticos en la evaluación en los procesos de evaluación y en los usos de la información.

13) El uso es diferente del informe y la difusión del mismo. Estos son medios para facilitar el uso. El uso puede ser tomar decisiones, mejorar programas, cambiar formas de pensar, facultar o “empoderar” a los participantes o generar conocimiento.

14) Se deben tratar seriamente las cuestiones de los costes económicos y de calendario. Los beneficios relacionados con estos costes se manifiestan en un mayor uso.

Patton apela a las Normas del *Joint Committee* (1994) para resaltar que, ante todo, una evaluación debe ser útil. Para que se dé este tipo de evaluación, tienen que darse cinco requisitos:

a) Capacidad de respuesta situacional: no hay una manera única de hacer evaluaciones y cada situación evaluativa es única.

b) Flexibilidad metodológica: es necesario conocer una gran variedad de métodos para aplicar el más adecuado en cada momento. De hecho, Patton habla del “Paradigma de la Elección” (*Paradigm of Choices*) que va más allá de la polémica de lo cuantitativo/experimentalista y lo cualitativo/naturalista.

c) Múltiples roles del evaluador: éste debe ser un investigador, pero también un profesor, un buen comunicante y relaciones públicas, etc. El factor personal es muy importante.

d) Sensibilidad política (*political sophistication*): se deben conocer los distintos intereses que por la evaluación pueden tener los diferentes grupos implicados en la misma. Estos intereses son diferentes en naturaleza, grado y se deben identificar y establecer prioridades en caso de que no todos puedan ser atendidos. En casi todas las evaluaciones existen implicaciones políticas que hay que asumir.

e) Creatividad: a la hora de enfrentar una evaluación. Existen dos tendencias que hacen que esto sea muy importante: 1) tendencia a alejarse del paradigma dominante hipotético-deductivo e ir hacia el “paradigma de elección” que enfatiza el uso de múltiples métodos y enfoques alternativos, y 2) los métodos de decisión ya no dominan el campo de la evaluación, el foco ha cambiado hacia el uso de la información, las cuestiones prácticas de la ejecución de programas y la ética de la práctica evaluativa.

La adecuación de diferentes enfoques y métodos a las diferentes situaciones evaluativas es una idea clave en Patton. Para demostrar que las situaciones evaluativas pueden ser muy diferentes entre sí, hace un listado de “variables situacionales” que son los aspectos o condiciones que puede variar y que configuran una situación evaluativa, y que aparecen resumidas en la tabla 2.2.

Tabla 2.2. Variables situacionales de Patton

<p>Algunas de las variables situacionales que hay que tener en cuenta son:</p> <ol style="list-style-type: none"> 1.- Número de responsables a tratar. 2.- Historia del programa (corta-larga) 3.- Actitud de los profesionales que trabajan en el programa (<i>staff</i>) ante la evaluación (entusiasmo-resistencia). 4.- Conocimientos de evaluación de dichos profesionales. 5.- Patrones de interacción (dirección-trabajadores-beneficiarios o clientes) en el programa (cooperación-conflicto). 6.- Historia evaluativa del programa. 7.- Nivel educativo y cultural de los profesionales y beneficiarios. 8.- Homogeneidad-heterogeneidad de profesionales y beneficiarios. 9.- Localización del programa (única o múltiple). 10.- Recursos para la evaluación. 11.- Número de fuentes de financiación. 12.- Naturaleza del programa: si es simple (ej: programa para dejar de fumar) o multidimensional (ej: programa para la prevención de la droga). 13.- Tipo de actuaciones o tratamiento (estables para el mismo grupo durante un tiempo, individuales o grupales, etc.) 14.- Estructura en la toma de decisiones en el programa (vertical-horizontal). 15.- Claridad-ambigüedad sobre el fin y la función de la evaluación. 16.- Datos existentes sobre el programa. 17.- Énfasis en la evaluación (voluntaria, impuesta). 18.- Tiempo disponible para la evaluación (mucho o poco, tiempo límite flexible o cerrado).

Fuente: Patton, M.Q., 1980. *Practical Evaluation*. Newbury Park. Sage.

Este listado de variables situacionales resulta ser una buena guía para realizar un diagnóstico de la situación evaluativa, operación necesaria que debe realizarse antes de delimitar una evaluación. El hecho de ser consciente de estas dieciocho variables va a facilitar la respuesta a los interrogantes sobre por qué, para qué y para quiénes

realizamos una evaluación. Esto a su vez ayudará a definir la estrategia de evaluación, decidiendo el tipo de evaluación a realizar así como el enfoque más adecuado.

En la última edición de su libro *Utilization-Focused Evaluation* introduce una serie de temas nuevos que reflejan, en cierto modo, la evolución de la evaluación en los últimos años. Algunos de estos temas son las cuestiones éticas en la práctica de la evaluación (tema candente en el ámbito anglosajón) y la utilización de procesos participativos de evaluación para cambiar la cultura del programa y contribuir al aprendizaje organizativo. De esta forma el concepto de utilización de agranda, ya no es sólo la utilización de los resultados de la evaluación sino también de los propios procesos evaluativos.

G. Enfoques participativos

Los enfoques participativos están teniendo especial auge en los últimos años en general en la evaluación de programas y, de forma especial, en algunos sectores como el de las políticas de cooperación al desarrollo y ayuda internacional. Por participativos, en sentido amplio, entendemos los enfoques que hacen especial énfasis en la participación de los agentes críticos en el programa en los procesos de evaluación. Es decir, la participación es el elemento central de los mismos. Aquí vamos a hablar de tres de estos enfoques:

- ***La evaluación de cuarta generación.*** Propuesta por Guba y Lincoln (1986, 1989), supone un enfoque alternativo desde el punto epistemológico y metodológico. Desde el paradigma alternativo naturalista, proponen un enfoque de evaluación que fundamenta teóricamente la participación de los diferentes agentes críticos. Su importancia e influencia es indudable, sobre todo, desde el punto de vista de la justificación ontológica y epistemológica de estos enfoques.
- ***La evaluación para el empoderamiento***⁸ (*Empowerment Evaluation*). Impulsada por David Fetterman (Fetterman et al.,1996) ha concitado y dado origen a toda una serie de experiencias participativas en evaluación. Su éxito e influencia es importante, entre otras cosas, por la importancia que está ganando el concepto de *empowerment* en otros ámbitos (por ejemplo, en el desarrollo de las políticas de igualdad de género).
- ***La evaluación pluralista*** centrada en los actores sociales. Este enfoque (Monnier, 1996) tiene especial influencia en el ámbito de las políticas y la

⁸ No hemos encontrado un término en español que pueda expresar el sentido del concepto de *empowerment*. Aún siendo conscientes de que no existe el término en español, hemos decidido utilizar una traducción literal que exprese el sentido del mismo. Pedimos disculpas por ello y agradeceríamos cualquier sugerencia para un término más adecuado.

gestión pública. También propone la necesaria participación de los actores sociales en torno a una política pública.

Por supuesto, estos no son los únicos enfoques participativos, aunque sí nos parecen los más interesantes desde el punto de vista de la evaluación. Se han propuesto diferentes métodos desde esta perspectiva participativa, en algunos casos en el ámbito de la planificación y gestión de proyectos, aunque han tenido también su reflejo en evaluación. Un ejemplo en el ámbito de la cooperación al desarrollo y los organismos internacionales, es el “Diagnóstico Rural Participativo” (*Participatory Rural Appraisal* o PRA). El PRA consiste en “una serie de aproximaciones y métodos que permiten a los habitantes de las comunidades locales (rurales o urbanas) expresar, ampliar, compartir, analizar sus conocimientos de las condiciones en las que viven, planificar acciones futuras y llevarlas a la práctica (...) La utilización del PRA implica ciertos cambios de orientación con respecto a la evaluación tradicional en cuanto a los marcos de análisis, las modalidades de trabajo, las relaciones y el poder” (Ministerio de Asuntos Exteriores, 1998: 92).

Antes de pasar a explicar cada uno de los enfoques, se exponen aquí algunas de las fortalezas y debilidades de este tipo de enfoques. Para Weiss (1998: 105 y ss.), la evaluación participativa da voz a los agentes críticos en el programa y fomenta la *cooperación* con el evaluador. También, a menudo ayuda a que dichos agentes críticos comprendan mejor el programa o política, es decir, que se produzca *aprendizaje sobre el programa*, además de que también fomenta el *aprendizaje sobre evaluación*, sobre todo de los profesionales del programa, pero también de otros agentes críticos. En el sentido de la justicia y la igualdad, las evaluaciones participativas tienden a compensar la balanza y a un mayor equilibrio, aunque la misma Weiss señala que tampoco hay que ser ingenuos al respecto: “sentarse juntos en una mesa de trabajo, en una atmósfera igualitaria, no elimina las distinciones que hay puertas afuera. Todavía son importantes las distinciones de clase, edad, raza, nacionalidad, ocupación, género y otras (...) La superación de las diferencias de poder y de conocimiento entre los diferentes agentes críticos no es fácil” (Weiss, 1998a: 106).

Weiss también es crítica y menciona una serie de limitaciones de este tipo de enfoques. En primer lugar, no todos los agentes críticos suelen estar interesados en la evaluación, ya que la mayoría de ellos van a estar más interesados en el programa que en su evaluación. Esto puede hacer que resulte difícil conseguir la participación de todos y que haya ausencias en las reuniones de trabajo. Además, Weiss cree que aunque los procedimientos participativos parece que aumentan la relevancia y la factibilidad del estudio, no está claro que lo haga respecto a la calidad técnica, que puede verse como secundaria. Por otro lado, aunque estos enfoques aseguran que se consigue aumentar las probabilidades de que la gente que participe en el estudio utilice los resultados, la verdad es que algunos agentes críticos en los programas no toman decisiones sobre, y ni siquiera influyen en, el futuro del programa, como por ejemplo, los beneficiarios. Cuando estén los resultados es probable que no tengan mucha autoridad para hacer algo con ellos y lo tendrán que hacer a través de la presión a otros. Por último, si se hace evaluación en un programa, servicio o política que tenga un cierto grado de movilidad entre sus profesionales y responsables, es posible que la gente con la que se ha trabajado en el comienzo del proceso no sea la misma que cuando estén listos los resultados.

Efectivamente, los enfoques participativos se encuentran con algunas contradicciones en su puesta en práctica en contextos con poca cultura participativa. En el ámbito de las políticas de cooperación al desarrollo, el análisis de las experiencias de varios países y agencias donantes, el Grupo de Evaluación del CAD (Comité de Ayuda al Desarrollo de la OCDE) (CAD, 1997, citado en Ministerio de Asuntos Exteriores 1998: 93-94) señala una serie de conclusiones referidas a cuestiones metodológicas de la evaluación participativa. Por un lado, dicha evaluación no es fácilmente integrable en los mecanismos rutinarios de la agencias donantes, ya que es percibida como un instrumento cuyas conclusiones son poco fiables y altamente subjetivas (en contraste con las que se obtienen de herramientas y métodos cuantitativos formales). Sin embargo, algunas técnicas participativas como las entrevistas con grupos focales o de discusión, comienzan a ser adoptadas como estrategias comunes para la obtención de información en las evaluaciones convencionales. El CAD realiza una serie de recomendaciones para las evaluaciones participativas:

- Se deben realizar en contextos favorables, que reconozcan la importancia de la participación y la incorporen en la planificación de las intervenciones.
- Debe ser entendida como un proceso en el que los involucrados participen desde sus primeras etapas, distinguiendo los proyectos con origen participativo de los que no lo tienen, y en los que la evaluación sea la etapa final de un seguimiento.
- Los conceptos y marcos de análisis deben ser definidos clara y precisamente, identificando de forma clara los grupos involucrados y los participantes.
- Los indicadores que se utilicen deben reflejar dichos conceptos y ser resultado de un análisis sistemático basado en la colaboración.

Según el CAD, lo que queda pendiente es introducir el concepto de participación en las líneas prioritarias de las agencias donantes, tarea a la que las propias evaluaciones pueden contribuir pero que debe tener un apoyo institucional más amplio y desde todos los niveles. También es necesario seguir desarrollando métodos, técnicas y marcos analíticos, procedimientos e indicadores que sean adecuados a los complejos niveles participativos de las agencias de cooperación.

G.a. La Evaluación de Cuarta Generación

Egon Guba y Yvonna Lincoln (1987, 1989) desarrollan un enfoque en evaluación al que denominan “Evaluación de cuarta generación”. Su propuesta ha tenido una importante repercusión ya que ha supuesto la base para muchos otros enfoques de tipo participativo y porque propone una perspectiva alternativa desde el punto de vista epistemológico.

La denominación de cuarta generación se debe a que ellos plantean cuatro generaciones de estudios de evaluación, que son como siguen:

Primera generación: La generación *técnica*. La evaluación se caracterizaba por la *medición*. La evaluación se veía como un medio para determinar si los alumnos alcanzaban una serie de especificaciones a través de una serie de exámenes o tests. El evaluador es un técnico.

Segunda generación: La generación de la *descripción*. Nace con Tyler y pretende describir las fortalezas y debilidades respecto a una serie de objetivos determinados. El evaluador es un descriptor.

Tercera generación: La generación del *juicio y la valoración*. Se pretende valorar el mérito o valor del programa evaluado. Y el evaluador es un juez. En esta generación se desarrollan toda una serie de modelos (Stake “Modelo de la Figura”, Scriven “Evaluación sin objetivos”, Stufflebeam, “CIPP”, etc.).

Cuarta generación: La generación *sensible (responsive)*, con capacidad de respuesta y basada en la *negociación*. En esta generación se toma como punto de referencia no los objetivos, las decisiones o los efectos, sino las preocupaciones, quejas y cuestiones de los diferentes agentes críticos en el programa evaluado y, por lo tanto, pueden verse afectados por dicha evaluación. En esta generación se pueden mencionar, además de a los propios Guba y Lincoln, al segundo Stake (Modelo sensible), la evaluación iluminativa de Parlett y Hamilton, y a Patton.

Una de las mayores aportaciones de estos autores es que su enfoque supone una propuesta ontológica, epistemológica y metodológica distinta. La evaluación que proponen se basa en un paradigma naturalista y constructivista (Lincoln y Guba, 1985; Guba, 1990). Esto quiere decir que, desde el punto de vista *ontológico* (lo que se considera que es la naturaleza de la realidad), piensan que las realidades existen como constructos o construcciones mentales y que son relativas a las personas que los sostienen. Desde el punto de vista *epistemológico* (cuál es la relación entre el conocimiento y el conocedor), creen que el conocimiento y el investigador forman parte de una misma entidad subjetiva. Los hallazgos e información que se consiguen son el resultado de una interacción. También significa que hay que estudiar los fenómenos tal y como se dan, en su contexto, sin manipulación de variables por parte del investigador. Desde el punto de vista *metodológico* (cómo se puede obtener el conocimiento), proponen identificar, comparar y describir las diferentes construcciones que existen. Esto significa una propuesta radicalmente distinta al paradigma positivista e hipotético-deductivo hasta entonces dominante, en el que se consideraba que la realidad existe y se dirige por leyes de causa-efecto que pueden ser conocidas, que la investigación puede y debe hacerse libre de valores y que las hipótesis pueden ser comprobadas de forma empírica. E incluso distinto del postpositivismo que “suaviza” el paradigma anterior ya que asume que la realidad existe pero no puede ser totalmente

comprendida y que existe una multiplicidad de causas y efectos, que la objetividad es un ideal, y aunque sea difícil hay que pretender conseguirla de forma crítica y que es crítico con el experimentalismo, enfatiza los enfoques cualitativos, la teoría y el descubrimiento. Esta comparación de paradigmas se refleja en la tabla 2.3.

TABLA 2.3: TABLA COMPARATIVA DE PARADIGMAS

Guba, en su libro *The Paradigm Dialog*, de 1990, distingue cuatro grandes paradigmas para la indagación social y contrasta:

- a) Su ontología (lo que se considera que es la naturaleza de la realidad),
- b) Su epistemología (cuál es la relación entre el conocimiento y el conocedor)
- c) Su metodología (cómo se puede obtener el conocimiento)

Positivismo

- a) La realidad existe y se dirige por leyes de causa-efecto que pueden ser conocidas.
- b) La indagación puede estar libre de valores.
- c) La hipótesis pueden ser comprobadas empíricamente.

Post-positivismo

- a) La realidad existe, pero no puede ser enteramente comprendida o explicada. Hay una multiplicidad de causas y efectos.
- b) La objetividad es un ideal, pero requiere de una comunidad crítica.
- c) Es crítico con el experimentalismo, y enfatiza las aproximaciones cualitativas, la teoría y el descubrimiento.

Teoría crítica

- a) La realidad existe, pero no puede ser enteramente comprendida o explicada. Hay una multiplicidad de causas y efectos.
- b) Los valores median en la indagación.
- c) Propone la eliminación de una falsa conciencia y facilita y participa en la transformación.

Constructivismo

- a) La realidad existe como constructos mentales que son relativos a aquellos que los sostienen.
- b) El conocimiento y el conocedor son parte de una misma entidad subjetiva. Los hallazgos son el resultado de la interacción.
- c) Identifica, compara y describe las varias construcciones que existen (hermeneútico y dialéctico).

Fuente: Parsons, W. 1995. *Public Policy. An Introduction to the Theory and Practice of Policy Analysis*. Chetelham: Edward Elgar, p. 71.

Los principios fundamentales de esta nueva propuesta son organizados por sus autores a través de cinco conceptos axiomáticos: pluralismo de valores, construcciones de los agentes críticos, justicia, mérito y valor y negociación.

A. Se ha de reconocer un *pluralismo de valores*, ya que las generaciones anteriores tendían a asumir el consenso de valores. Aquí se reconocen y se valoran las diferencias. De esta premisa arrancan tres principios:

1. En una evaluación que tenga en cuenta las diferencias de valores es más esperable el conflicto que el consenso.
2. Los procesos técnicos, descriptivos y valorativos de las anteriores generaciones de evaluación no son suficientes para tratar el desacuerdo partidario; se deben disponer algunos medios que permitan la acomodación entre diferentes posturas de valores.
3. Las partes que representan diferentes valores deben implicarse si se quiere tener alguna esperanza de que estos tengan cabida.

B. El concepto de las *construcciones de los agentes críticos*: Como se ha visto antes, este enfoque está enraizado en una concepción ontológica relativista, en la que se concibe que las realidades son construcciones sociales relativas a quienes las sostienen. Esto supone dos principios:

4. El evaluador a veces aprende sobre cuáles son las construcciones de los agentes críticos, y en el mismo acto de aprender, ayuda a crear esas construcciones.
5. El evaluador a veces enseña y comunica las construcciones de otros agentes críticos y, en este proceso, clarifica las posiciones de valores que tienen que ver con ellos.

C. El concepto de *justicia*: Nadie debe tener especial preferencia en los procesos de evaluación. Esto añade otros cuatro principios:

6. El evaluador está éticamente ligado a solicitar cuáles son las preocupaciones, quejas y cuestiones de todos los agentes críticos identificables, y a clarificar los valores que subyacen a las mismas.
7. La equidad demanda que el evaluador dé igual validez a las cuestiones de cada uno de los grupos agentes críticos.
8. Se debe recoger información de modo que resulte útil a los protagonistas y a los antagonistas del programa.
9. El evaluador debe demostrar y enseñar los datos recogidos de forma continua, para dar oportunidad de valoración por parte de los agentes críticos.

D. Los conceptos de *mérito y valor*: Estos dos conceptos son distintos entre sí, dependiendo el mérito de las características internas del evaluando y el valor, de características externas, siendo un concepto relativo al entorno en el que se sitúa dicho evaluando. De esto, se desprenden otros tres principios:

10. Mientras el mérito puede ser valorado en solitario, el valor sólo puede valorarse en relación a un contexto y unas necesidades determinadas. Si se cambia el contexto, se cambiarán las necesidades y el valor.

11. El evaluador no puede llevar a cabo una valoración de las necesidades o el valor, excepto con métodos de campo que tomen en cuenta de forma plena el contexto local.

12. La transferencia de las valoraciones de necesidades y/o valor de un contexto a otro depende de la similitud empírica de los contextos.

E. El concepto de *proceso y resultados negociados*: Si se asumen los axiomas anteriores, está claro que una evaluación justa sólo puede hacerse a través de la negociación. Finalmente, esto implica otros cinco principios:

13. A los agentes críticos se les debe dar la oportunidad de reaccionar ante los resultados por razones éticas y de validez y para satisfacer el requisito de colaboración.

14. El evaluador debe implicar a todos los agentes críticos en todos los momentos del proceso de evaluación, solicitando y teniendo en cuenta sus puntos de vista en las decisiones a tomar.

15. Los agentes críticos deben enrolarse en un proceso de negociación que determine las acciones a llevar a cabo en la evaluación. Es este grupo de agentes críticos el que actúa como juez y que realiza las recomendaciones. El evaluador actúa como un mediador.

16. El evaluador debe permitir la posibilidad de desacuerdos entre los agentes críticos. El objetivo del proceso de negociación no es tanto el consenso como la comprensión.

17. El evaluador debe actuar como mediador y agente de cambio en estas negociaciones, animando al seguimiento de las decisiones acordadas y actuando como monitor que valora dicho seguimiento.

De esta manera, la evaluación tiene un significado muy distinto del que había tenido en las tres generaciones anteriores. Guba y Lincoln distinguen ocho rasgos distintivos de este nuevo enfoque, que se resumen de la siguiente manera:

1. La evaluación es un proceso socio-político.
2. La evaluación es un proceso de enseñanza-aprendizaje.
3. La evaluación es un proceso continuo, recurrente y divergente.
4. La evaluación es un proceso que crea "realidad".
5. La evaluación es un proceso emergente.
6. La evaluación es un proceso con resultados impredecibles.

7. La evaluación es un proceso de colaboración.

8. La agenda para la negociación se manifiesta mejor en un formato de estudio de caso, siendo descritas las cuestiones que requieren negociación respecto a las particularidades de ese caso.

Los autores, sin embargo, reconocen que la evaluación de cuarta generación tiene algunas limitaciones. Estas son:

1. Falta de legitimación del nuevo enfoque. Este es el análisis que hacen los autores a finales de los años 80. Creemos que, hoy en día, una década más tarde, aunque el problema de la legitimidad de los enfoques participativos sigue existiendo, su credibilidad ha ido y sigue aumentando de forma paulatina y su aceptación es cada vez mayor, si no de los enfoques metodológicos globales, sí de algunas de las técnicas que se proponen (CAD, 1997, citado en Ministerio de Asuntos Exteriores, 1998: 93-94).
2. Disfuncionalidad política. Si se toma en serio el aspecto de colaboración, la gente con poder tiene que estar dispuesta a delegarlo.
3. Disfuncionalidad humana: Para los autores de este enfoque, todas las evaluaciones son disfuncionales para el desempeño humano, creando ansiedades y miedos que son difíciles de combatir. Cuanto más intensas y personales sean las experiencias, más aumentará esta tendencia. Por eso, la creación de un clima de confianza es importante, pero éste lleva tiempo.
4. Los límites del tiempo y el contexto: Esta evaluación está ligada al momento y al contexto determinado en el que se lleva a cabo. Son evaluaciones que no son generalizables.
5. Incapacidad de un diseño específico previo: En este enfoque, el diseño de evaluación es emergente, no puede tener una planificación previa.
6. La naturaleza abierta del enfoque: Es muy difícil alcanzar un punto final; las evaluaciones se acaban cuando se acaba el tiempo o el dinero. La contestación de una pregunta levantará unas cuantas preguntas más.

G.b. La Evaluación para el Empoderamiento

Como ya se ha comentado anteriormente, el catalizador y propulsor de esta propuesta es David Fetterman, cuya conferencia presidencial y propuesta de tema general para el encuentro anual de la Asociación Americana de Evaluación del año 1993 trató de ese tema. A partir de ese año, el enfoque se ha ido desarrollando y tomando cuerpo y sus principios, premisas y experiencias se reflejan en la obra del año 1996 (Fetterman, Kaftarian y Wandersman, 1996).

La evaluación para el empoderamiento “consiste en el uso de los conceptos, técnicas y hallazgos de la evaluación para promover mejoras y autodeterminación” (Fetterman, Kaftarian y Wandersman, 1996: 4). Este enfoque tiene un claro propósito: ayudar a la gente a ayudarse a sí mismos y a mejorar sus programas utilizando un formato de autoevaluación y reflexión. Los participantes en los programas dirigen sus propios programas: un evaluador externo a menudo sirve como árbitro y facilitador dependiendo de las capacidades internas del programa. Su lema podría ser el de “mejor que ofrecer el pescado, ofrece la caña y enseña a pescar”. Desde luego, este es un enfoque que da absoluta prioridad a los profesionales de base de los programas y políticas, a lo que podríamos llamar, en términos de Lipsky, los “burócratas a nivel de calle” cuya importante influencia en el proceso de elaboración de las políticas públicas ha quedado demostrado.

En la evaluación para el empoderamiento se propone emplear tanto metodología cuantitativa como cualitativa y aunque se pueda aplicar a individuos, organizaciones o comunidades, su foco está en los programas. Requiere necesariamente una actividad de grupo y de forma cooperativa, es decir, no es una actividad individual del evaluador. Esto no quita que tenga que mantenerse el rigor del estudio ni que la objetividad sea una preocupación. Aunque se reconoce que la evaluación no es nunca neutral y que es un acto político en sí mismo, se aduce que se contrarrestan posibles sesgos al estar la evaluación en manos de un grupo, que puede actuar como controlador de los posibles sesgos individuales. Es, por lo tanto, un proceso democrático y no tiene por qué ser excluyente de una evaluación externa. Aquí es importante el concepto de empoderamiento y su significado. El empoderamiento no es algo que alguien otorga sino algo que surge de las mismas personas. Aplicado a la evaluación, el evaluador no empodera a nadie sino que la gente se empodera a ella misma con ayuda y arbitraje.

Las facetas o fases de desarrollo de la evaluación para el empoderamiento y que forman parte integral del proceso de evaluación son (Fetterman, Kaftarian y Wandersman, 1996: 9 y ss.):

- **Formación:** se da formación a los participantes y esto se utiliza para hacer mapas del terreno, estableciendo las preocupaciones y categorías que se quieren estudiar.
- **Facilitación:** como árbitros y facilitadores para que los otros realicen una autoevaluación. Esto consiste en una orientación general y hacia el proceso de la evaluación. El evaluador puede asistir a las reuniones de trabajo, remover ciertos obstáculos, clarificar patrones de mala comunicación, ayudar a crear equipos de facilitación, a trabajar con unidades reacias y dar ideas nuevas para trabajar con unidades cansadas. También puede ayudar a crear el diseño de evaluación.

Cualquiera que sea su contribución, el evaluador debe asegurar que la evaluación permanece en manos del personal del programa.

- **Apoyo- Defensa (*Advocacy*):** se trata de ayudar a los grupos a empoderarse a través de la evaluación y de defender el programa evaluado, trabajando incluso con representantes políticos, escribiendo en la arena pública para cambiar la opinión pública y ofreciendo información relevante en los momentos adecuados del proceso de toma de decisiones. Muchas veces los evaluadores deben ser defensores de grupos que no tienen control sobre sus vidas como por ejemplo, la población sin techo o los fracasados escolares. En un contexto de empoderamiento, los evaluadores deben permitir a los participantes cambiar la dirección de la evaluación, sugerir soluciones ideales a sus problemas y, entonces, tratar de conseguir el cambio social.
- **Iluminación:** se trata de ayudar a abrir los ojos e ilustrar con ejemplos, datos, hechos e ideas sobre los roles, las estructuras y las dinámicas de los programas. Se puede ayudar a conseguir datos que ilustren efectos o a identificar cuáles son las fortalezas y planteárselas como objetivos a trabajar. “La experiencia de iluminación ofrece el mismo tipo de intoxicación intelectual que cada uno de nosotros ha experimentado cuando por primera vez nos topamos con una pregunta de investigación. El proceso crea una comunidad dinámica de aprendices a medida en que la gente se engancha en el arte y la ciencia de evaluarse a sí mismos” (Fetterman, Kaftarian y Wandersman, 1996: 16).
- **Liberación:** es el acto de verse liberado o liberarse a sí mismo de roles y limitaciones preexistentes. A menudo implica nuevas conceptualizaciones sobre uno mismo y los otros. Este tipo de evaluación permite a los participantes encontrar nuevas oportunidades, ver los recursos ya existentes de otra manera y redefinir sus identidades y sus papeles futuros.

Los pasos a dar en una evaluación para el empoderamiento son los siguientes (Fetterman, Kaftarian y Wandersman, 1996: 20)

- 1) Determinación de donde se encuentra el programa, incluyendo fortalezas y debilidades (normalmente se les pide a los participantes que clasifiquen el programa en las dimensiones que se deciden importantes),
- 2) Establecimiento de objetivos y metas, determinando hacia dónde se quiere ir en el futuro con un énfasis explícito en la mejora del programa,
- 3) Desarrollo de estrategias y ayuda a los participantes para que determinen sus propias estrategias para cumplir los objetivos, y
- 4) Documentación del programa: ayuda a los participantes del programa para que determinen el tipo de información necesaria para documentar un progreso creíble hacia los objetivos.

G.c. La Evaluación Pluralista

Para entender el modelo de la evaluación pluralista propuesto por Monnier (1990-1991, 1995), es conveniente tener en cuenta su visión sobre la acción de los poderes públicos, las políticas públicas innovadoras y las relaciones entre el Estado y la sociedad civil.

Para Monnier, las políticas públicas innovadoras requieren siempre de la negociación, ya que existen fuertes intereses políticos en torno a las mismas. A esto hay que añadir, además, la complejidad de las relaciones entre los poderes públicos y la sociedad civil. Para analizar esta realidad hay que superar los modelos jurídicos y técnico-económico de acción de los poderes públicos (que son precisamente, según Monnier, los dos modelos que dominan en la administración francesa). Bajo un modelo jurídico, y siguiendo el esquema weberiano, las organizaciones de naturaleza burocrática tienen su fuente de legitimidad en el marco jurídico instituido: así, la administración pública ejecuta las leyes votadas en el Parlamento, elegido a su vez democráticamente por la sociedad civil. Por lo tanto, bajo este modelo, como la actuación estatal se ajusta exclusivamente a derecho, se proscriben toda interacción entre la sociedad civil y el Estado. En el modelo técnico-económico la fuente de legitimidad se basa en referencias científico-técnicas. Este modelo presupone la existencia de objetivos explícitos y compartidos y excluye cualquier dimensión que no sea técnico-económica, es decir, por ejemplo, la evaluación se declara apolítica.

Como dice Monnier, “Los modelos precedentes pueden ser relativamente adecuados a la hora de analizar políticas rutinarias, con un alto grado de formalización (...) pero distan de resultar apropiados en el marco de las iniciativas para hallar soluciones a un problema social nuevo, o para mejorar el modo de funcionamiento o de organización del aparato estatal” (Monnier, 1995: 105). Un tercer modelo, que llama “Modus Vivendi” o negociado, explicaría mejor la realidad. Según este modelo, el contexto sociopolítico es cambiante y la formulación de las políticas es compleja y ambigua. Los programas y las políticas innovadoras siempre van a concitar conflictos, ya que hay diferentes intereses. Se tiene que dar, pues, un proceso de negociación entre los diferentes actores sociales que tienen intereses distintos y llegar a un compromiso. En definitiva, las decisiones no se adoptan, sino que “afloran” de la interacción entre diferentes perspectivas, puntos de vista, argumentos, es decir, son productos de un “modus vivendi”. “A todas luces, gobernar equivale a gestionar la acción colectiva” (Monnier, 1995: 261).

Mientras que para el modelo jurídico la evaluación correspondería al control de la regularidad y en el técnico-económico se tiende a la medida estadística, al modelo negociado y “modus vivendi” le corresponde la evaluación pluralista. Esta evaluación debe ser útil socialmente y esta utilidad social de la evaluación reside más en la credibilidad y la legitimidad del proceso de evaluación por parte de los actores críticos que en la metodología y el rigor técnico. Se trata más de una legitimidad política que de una científico-técnica. El enfoque pluralista asocia, pues, los actores sociales con el desarrollo de la evaluación, para aumentar su credibilidad y legitimidad. Por lo tanto, la evaluación nunca es ajena a las relaciones de poder ni neutral, convirtiéndose en un acto político en sí mismo que permite la relación entre el Estado y la sociedad civil.

Su propuesta pluralista consiste en que “la evaluación genere verdaderamente un *espacio para negociar*” (Monnier, 1995: 149). Esta construcción de la legitimidad política de la evaluación implica que, además de un evaluador profesional, se debe contar con una instancia de evaluación, depositaria del mandato político. Esta instancia que Monnier llama “comité de pilotaje” agrupa a los diferentes protagonistas del programa o política a evaluar, selecciona el equipo evaluador y hace un seguimiento de su trabajo. Así, los comanditarios (los que encargan) la evaluación no tienen mayor poder que los demás miembros del comité, y además pueden beneficiarse de la reflexión colectiva.

El dispositivo de evaluación debe cumplir, al menos, tres papeles:

- a) Ser un instrumento de observación que permita percibir mejor la realidad
- b) Ser un instrumento de medida de los fenómenos sociales, y
- c) Ser “un espejo reflectante que ayude a los actores a tomar conciencia de sus expectativas y a ajustarlas a la realidad” (Monnier, 1995: 269).

Según la situación podrá darse mayor o menor importancia a uno u otro componente, pero ninguno de ellos puede soslayarse totalmente.

Finalmente, para Monnier, el marco de referencia de la evaluación debe ser el análisis de políticas públicas. El análisis de políticas públicas reconoce la complejidad de los procesos y las dificultades inherentes a la acción pública; desde esta perspectiva también debe entenderse la evaluación. De hecho, para el autor “el tema de la evaluación está estrechamente ligado con el reconocimiento de la política como modalidad de acción colectiva; es decir, con la identificación de la acción pública en términos de proceso” (Monnier, 1995: 266). El tradicional enfoque por objetivos representaba una visión centralizada y unitaria del poder; la visión pluralista declara una dimensión necesariamente política de toda acción pública, que implica una multiplicidad de centros de poder, lo que obliga a la evaluación a estar muy apegada al terreno. El análisis de políticas públicas y, más en concreto, la evaluación, puede permitir una mayor viabilidad y legitimación de las ciencias sociales para estudiar las políticas públicas sobre el terreno y contribuir a su mejora.

2. TIPOS DE EVALUACIÓN

En la literatura especializada en teoría y metodología de la evaluación abundan las clasificaciones sobre diferentes tipos de evaluación (Alvira, 1991; Chelimsky, 1989; Stufflebeam y Shinkfield, 1987). De hecho, parece que la práctica de clasificar la evaluación en diferentes tipos es un ejercicio de definición y aprehensión del propio concepto de la evaluación. En esta línea, Scriven (Scriven, 1967), uno de los autores más prominentes en evaluación y que viene originariamente del campo de la filosofía de la ciencia, articula parte de su discurso inicial sobre la evaluación alrededor de polos duales (evaluación formativa vs. sumativa, intrínseca vs. final, “amateur” vs. profesional, etc.).

De todas las posibles clasificaciones, se exponen aquí las cuatro tipologías básicas que a nosotros nos parecen clave para distinguir la naturaleza de la evaluación. Estas cuatro tipologías se construyen a partir de cuatro dimensiones: el papel que cumple la evaluación, el contenido o aspecto de la política que se evalúa, el agente evaluador y el momento en el que se evalúa. Las denominamos básicas ya que son aplicables simultáneamente, es decir, cada evaluación podría ser clasificada por todas y cada una de estas cuatro clasificaciones. Así, por ejemplo, la evaluación del II Plan de Igualdad de Oportunidades para las Mujeres (1993-1995) llevada a cabo por el Instituto de la Mujer (IM a partir de ahora), es una evaluación sumativa, de proceso y resultados, mixta (tiene partes de evaluación interna y partes de evaluación externa) y *ex post facto*. Pasamos a explicar cada una de estas cuatro clasificaciones.

2.1. Según el papel que cumple la evaluación

Cronbach (63) es el primero que aplica el término de evaluación formativa a la evaluación de programas y a Scriven (67) se le debe la distinción entre evaluación **formativa** y evaluación **sumativa**. Esta distinción se realiza según el papel que cumple la evaluación, es decir, y como sus nombres indican, según se quiera utilizar la evaluación para ir (con)formando el evaluando (programa, servicio o política evaluada) -evaluación formativa- o se quiera utilizar la evaluación para hacer recapitulación del evaluando y realizar una valoración global del mismo -evaluación sumativa-. A Stake se debe un símil que creemos ilustra muy bien la distinción entre evaluación formativa y sumativa. En un restaurante, la evaluación formativa es la que realiza el cocinero probando la sopa mientras está acabando su faena culinaria, con el fin de comprobar si necesita rectificar añadiendo algún ingrediente. La evaluación sumativa es la que realizan los clientes del restaurante cuando les han servido la sopa en la mesa.

Esta tipología de evaluación formativa-sumativa surge inicialmente de la evaluación educativa aplicada a la evaluación individual del alumno. La evaluación sumativa es la que hace una valoración global recapitulando el desempeño del alumno, a efectos de calificación y clasificación del mismo; la evaluación formativa es la que

pretende que la propia evaluación sea un instrumento docente que permite obtener *feedback* en una doble dirección (profesor-alumno) y que guía el propio desempeño del alumno⁹. La transferencia de esta idea a la evaluación de programas y políticas provoca esta tipología básica según el papel que queramos cumpla la evaluación.

La traducción práctica de esta distinción se puede ilustrar con un ejemplo de una evaluación de un programa que lleve, en su ejecución, el ritmo del curso académico. En su versión formativa, los agentes evaluadores definirían las preguntas o criterios de evaluación en el mes de septiembre. Nada más empezar la aplicación del programa, se comenzaría a recopilar y a analizar información para contestar a dichas preguntas, se valoraría y, sobre la marcha se utilizaría la información evaluativa para la modificación y ajuste del programa evaluado; dichas modificaciones generarían más preguntas de evaluación que se intentarían contestar y así, sucesivamente. El resultado sería que el programa evaluado se habría ido modificando y reforzando sobre la marcha como consecuencia de la evaluación formativa. En su versión sumativa, también el equipo evaluador tendría definidas las preguntas o criterios de evaluación y antes del comienzo de la aplicación del programa en septiembre habría decidido cuáles serían los puntos y momentos en los que, a lo largo de la aplicación del programa en el período septiembre-junio, habría que recoger información. La información se recogería en dichos momentos, pero se iría acumulando para hacer un análisis y una valoración conjunta en el mes de junio. El resultado sería una evaluación global del programa durante el curso académico que permitiría mejoras y modificaciones para ediciones posteriores.

El papel que se le confiere a la evaluación en cada tipo es, por lo tanto, distinto. La evaluación formativa se asocia más a la función de mejora de la evaluación, y la sumativa, a la de rendimiento de cuentas. Aunque este paralelismo no es absolutamente lineal. También la evaluación formativa puede cumplir una función de responsabilidad y, por supuesto, una sumativa la de mejora (aunque no sea de aplicación inmediata). Algunos autores como Cronbach (Cronbach, 1963; Cronbach et al., 1980) han defendido la evaluación formativa por encima de la sumativa, haciendo hincapié, precisamente, en la función de mejora de la evaluación. Para Scriven, sin embargo, esta postura no es realista; para él, toda decisión que no sea arbitraria descansa sobre una evaluación sumativa (y relacionada con cómo resolver problemas como la droga, el control de crímenes, la cirugía, etc.). En definitiva, aunque se haga buena evaluación formativa (que suele resultar, además, cara), nunca se debe dejar de hacer sumativa. De hecho la tesis número diez de su obra *Hard-Won Lessons in Program Evaluation* señala que “La evaluación formativa es atractiva, pero la evaluación sumativa es un imperativo (Scriven, 1993: 59).

⁹ Desgraciadamente, la práctica totalidad de la evaluación que se aplica al alumnado en el actual sistema universitario español, es de tipo sumativo. Además de la propia cultura e inercia del sistema universitario, existen al menos tres factores que, a nuestro juicio, contribuyen a la dificultad de aplicar una evaluación formativa del alumnado: el poco reconocimiento de la labor docente del profesorado universitario fuera del aula (tanto de forma oficial como elemento cultural aprehendido por la comunidad universitaria), la existencia de grupos grandes de alumnos y el sistema cuatrimestral de asignaturas.

2.2. Según el contenido o aspecto de la política que se evalúa

La segunda clasificación básica se establece según el aspecto o contenido del programa o política que se evalúa. Así, podemos distinguir entre la evaluación del **diseño** o conceptualización del programa o política, la evaluación del **proceso** de puesta en práctica de dicho programa o política y, finalmente, la evaluación de **resultados**¹⁰.

La evaluación del **diseño** o conceptualización del programa estudia y valora cómo está diseñado o formulado un programa o política. Los aspectos que trata este tipo de evaluación tienen que ver con si el programa está bien fundamentado, si está basado en un buen análisis de la realidad, si las metas y objetivos están bien formulados, definidos, jerarquizados, especificados, si las acciones que se proponen responden de forma lógica a los objetivos, si se prevén los recursos necesarios (incluidos los humanos y su organización), si está explicitada una metodología de intervención, si se han diseñado y adecuados los mecanismos de información, seguimiento y evaluación, si es coherente el programa con la filosofía y objetivos de la institución o instituciones que lo formulan, ejecutan y evalúan, etc. En realidad, este tipo de evaluación tiene que ver con lo que algunos autores (Weiss, 1998; Chen, 1990) denominan la “teoría de programa”, que consiste en todo lo relacionado con las creencias subyacentes a una intervención. Como se ha señalado en el punto 1.4. de este capítulo, referido a los diferentes enfoques, Weiss (1998) habla de la teoría de programa y de la teoría de la implementación¹¹: ambas teorías equivaldrían al contenido de la evaluación del diseño y la conceptualización de un programa o política. Este tipo de evaluación resulta clave en las políticas públicas porque ofrece una necesaria visión de la intervención pública y porque es, además, el tipo de evaluación menos practicado.

La evaluación del **proceso** de puesta en práctica (o si se quiere, tomando como referencia el lenguaje más utilizado en análisis de políticas públicas, de la **implementación**) se refiere a todos los aspectos relacionados con el funcionamiento y ejecución de una determinada política o programa. En términos de Pressman y Wildavsky, la implementación (ejecución, puesta en práctica) de una política es “el conjunto de acciones que transforma las intenciones en resultados observables” o “el espacio intermedio del actuar entre el extremo de un concepto de política como ‘declaración de intenciones’ –anterior a toda acción- y el extremo del otro concepto de

¹⁰ Se podría incluir en esta clasificación un primer tipo de evaluación: la evaluación de necesidades que coincidiría con la evaluación del contexto del modelo CIPP de Stufflebeam (Stufflebeam y Shinkfield, 87). Aquí ha decidido descartarse por dos razones. Por un lado, para hacer una evaluación de necesidades o contexto no tiene por qué existir necesariamente una intervención por lo que no se estaría hablando estrictamente de evaluación de programas; por otro, el análisis de la realidad o diagnóstico de necesidades puede llevar una lógica metodológica muy distinta a las evaluaciones.

¹¹ Como se explicaba en el punto 1.4., la teoría del programa se refiere a los mecanismos que median entre un programa y sus resultados y la teoría de la implementación tiene que ver con lo que se requiere para traducir los objetivos en las operaciones de programa.

política como ‘declaración de resultados’ –posterior a toda acción– (Aguilar, 1996d:44). Este tipo de evaluación se centra en dicho espacio, es decir, en cómo una política o programa se pone en práctica, cómo funciona en el terreno, cuáles son las dificultades que encuentra en dicha ejecución y si se está llevando a cabo tal y como se tenía previsto (incluido el grado de cumplimiento de las acciones previstas) y cuál es el grado de desviación. Esto incluye lógicamente la valoración de los recursos que se utilizan para esta puesta en práctica, incluyendo los recursos humanos. Aquí debe incluirse, pues, todo lo relacionado con el funcionamiento del equipo que lleva a cabo el programa, las relaciones intergubernamentales e interinstitucionales que se establecen para ello y los actores críticos y participantes que intervienen para la puesta en práctica (facilitando o dificultando la misma). También se tiene en cuenta si se está llegando al grupo de destinatarios previsto y, en general los flujos de entrada, estancia y salida de dichos destinatarios en el programa, así como si se cumple el calendario previsto y cuál es la “anatomía del retraso” (de nuevo en términos de Pressman y Wildavsky) y, en general, todo lo que tiene que ver con la dimensión temporal de la ejecución de un programa o política. Es evidente que este tipo de evaluación es necesario ya que siempre existe diferencia entre lo que está diseñado y previsto y lo que luego se pone en práctica. La explicación y valoración de esa puesta en práctica ofrece claves ineludibles para la posterior interpretación y evaluación de los resultados.

Finalmente, además del diseño y del proceso de puesta en práctica, se pueden evaluar los **resultados** de un programa o política pública. Aquí resulta interesante distinguir entre **productos** (lo que produce o general un programa o política, lo que correspondería al término de *outputs*) y los **efectos** (los efectos que tienen dichos productos, tanto esperados como no esperados, lo que equivaldría al término *outcomes*). Por ejemplo en un curso de formación de postgrado, los productos serían las horas de formación (incluyendo su estructuración, metodología, etc., es decir, su calidad docente), la documentación y material entregado además de la producida por los alumnos, las relaciones que se establecen entre profesorado-alumnado y entre el propio alumnado, etc. y los efectos serían el aprendizaje y capacitación de los alumnos, así como la aplicación de los conocimientos aprendidos en los puestos de trabajo o la entrada al mercado de trabajo, pasando por la utilización de las relaciones establecidas para la realización de actividades profesionales conjuntas, etc. Entre los efectos, se podría distinguir entre **efectos directos** (los efectos sobre el grupo de destinatarios previstos) y **efectos indirectos** (los efectos que se producen en grupos más amplios o en la sociedad en general) . La evaluación de estos efectos indirectos es a veces denominada evaluación del **impacto** de un programa o política. La distinción entre productos y efectos es una distinción interesante para la evaluación de políticas. Es lógicamente más fácil identificar y evaluar los productos de un programa que sus efectos, que suele ser más costoso. Sin embargo es imprescindible valorar la calidad de los productos para poder tener una referencia sobre los efectos que han o no han producido. Por otro lado, el hecho de evaluar los productos no exime necesariamente del intento de medición y valoración de los efectos (ésta es la tendencia que ha habido en educación ocupacional no formal y de postgrado en España, en la que se ha desarrollado una buena metodología de evaluación del producto formativo, pero en mucha menor medida, la evaluación de los efectos). Además, la dificultad de la evaluación de los efectos de un programa o política estriba en la identificación de sus **efectos netos** (los efectos achacables al programa o política y no a otros factores). Sin embargo, la evaluación exclusiva de los efectos de un programa (sin que ofrezca

información sobre los productos que los han generado, los procesos de puesta en práctica y el diseño y teoría subyacente a dicho programa) suele ofrecer una visión parcial que es, además poco útil para la mejora de los programas y el rendimiento de cuentas y para la explicación de las razones de que se hayan dado o no una serie de efectos.

2.3. Según el agente que evalúa

Según el agente que evalúa, se puede distinguir entre evaluación **interna** y evaluación **externa**. La evaluación **externa** es la que se realiza por parte de alguien que no tiene que ver ni con la financiación, ni con la formulación ni con la ejecución del programa o política que se evalúa. Por lo tanto, la evaluación **interna** es la que se hace por parte de una institución, unidad de evaluación, equipo o agente que ha participado o participa en el programa o política en alguno o más de los aspectos antes mencionados (financiación, formulación y ejecución). Se pueden dar también procesos **mixtos** que consisten en dividir el proceso de evaluación y realizar alguna parte de forma interna y otra externa (por ejemplo, diseñar la evaluación internamente y ejecutarla externamente o, por el contrario, contar con asesores externos para el diseño del estudio de evaluación y posteriormente ejecutarlo internamente).

Entre las ventajas de la evaluación interna se encuentran un mejor conocimiento del programa o política a evaluar y su contexto, un menor coste (al menos en costes directos), una mayor probabilidad de implicación de la comunidad de programa en los procesos de evaluación y una mayor probabilidad de utilización de la evaluación por parte de profesionales y agentes críticos por haber sido protagonistas de la misma. Por el contrario, los procesos de evaluación externa gozan normalmente de una mayor credibilidad externa y de una mayor probabilidad de ser objetivos (en el sentido de ausencia de sesgos, por el hecho de que el agente evaluador se sitúa externamente y puede situarse desde todas las diferentes perspectivas de los distintos grupos y personas implicadas en el programa evaluado), una mayor probabilidad de cumplimiento del calendario (las evaluaciones externas siempre se terminan, las internas tienen mayor probabilidad de retraso y de no conclusión) y una mayor posibilidad de realizar procesos de evaluación técnicamente sofisticados y que requieran de una especialización concreta.

Sin embargo, estas ventajas y desventajas de un tipo y otro de evaluación son siempre relativas: un buen evaluador externo debe conocer profundamente el programa a evaluar y su contexto, los procesos de evaluación interna pueden ser tremendamente costosos si se miden los costes indirectos de los mismos, los evaluadores internos deben también procurar la ausencia de sesgos y los enfoques participativos, en los que profesionales e agentes críticos son protagonistas, no sólo pueden sino que, en muchos casos, conviene que estén dirigidos por un equipo externo. Por último, el cumplimiento del calendario previsto es una cuestión esencial en cualquier proceso de evaluación, ya sea interno o externo. La elección de una evaluación interna o externa depende, pues, de la situación evaluativa concreta, de los recursos disponibles y de la finalidad de la evaluación y la utilización posterior que vaya a dársele.

Finalmente, se quiere señalar dos cuestiones respecto a esta tipología. En primer lugar, en el proceso de las políticas públicas debe existir siempre un nivel de evaluación interna que no debe obviarse por la eventual existencia de evaluaciones externas. Es decir, el hecho de que exista la posibilidad de realizar una evaluación externa no elude las responsabilidades de evaluación interna de la gestión pública. Aunque lógicamente hay que procurar que no existan solapamientos entre ambos procesos, normalmente suelen darse en niveles distintos (es posible que la evaluación externa se centre en aspectos más globales y la interna más en los procesos internos de gestión, por ejemplo). En segundo lugar, la realización de buenos procesos de evaluación externa es una excelente manera de ir introduciendo una cultura evaluativa en las organizaciones que, posteriormente, pueden evaluar internamente de forma normalizada.

2.4. Según el momento en el que se evalúa

La última clasificación se refiere al momento en el que se realiza la evaluación y, en su versión más simple, consta de dos grandes categorías: la evaluación *ex ante* (la que se realiza antes de que el programa o política evaluada se ponga en marcha) y la evaluación *ex post facto* (la que se realiza durante o después de que el programa o política evaluada se haya puesto en marcha). Se considera evaluación *ex ante*, además de a la evaluación que se pueda realizar antes de que un programa se ponga en marcha en sentido estricto (en este caso, se trataría lógicamente de una evaluación del diseño y conceptualización del programa, “sobre el papel”), a la evaluación de las experiencias piloto. Este es un tipo de evaluación que se ha practicado bastante en algunos organismos internacionales: antes de generalizar un programa nuevo de grandes dimensiones, se hace una experiencia piloto en uno o varios lugares determinados. La evaluación de dicha experiencia piloto (a la que, además, se le dedica una gran cantidad de esfuerzo y recursos) se la considera evaluación *ex ante*. La evaluación *ex post facto* es cualquier evaluación que se realiza durante o una vez el programa se ha ejecutado (a veces, se distingue entre estas dos posibilidades: *in via* o evaluación “durante” y *ex post facto* o evaluación “después”). En los países en los que existe poca cultura y poca práctica de evaluación, como el nuestro, la práctica totalidad de evaluaciones que se realizan suelen ser *ex post facto*. De hecho, muchas evaluaciones no se prevén con la suficiente antelación (teniéndolas previstas al inicio del programa o plan a evaluar, aún cuando sea una evaluación de tipo sumativo), planteándose su diseño y ejecución en las fases finales del programa a evaluar, limitándose así las posibilidades de recopilación sistemática de información y, por lo tanto, de la evaluación en general.

TABLA 2.4. TIPOS DE EVALUACION
Clasificación según las cuatro tipologías básicas

SEGUN EL PAPEL DE LA EVALUACION

Formativa: Se hace con la intención de mejorar sobre la marcha el diseño y la ejecución del programa a evaluar.

Sumativa: Se hace una recapitulación del programa durante un periodo de tiempo determinado para ofrecer una valoración global y final del mismo.

SEGUN EL CONTENIDO O FASE DEL PROGRAMA

Evaluación de Diseño: Analiza y valora el diseño y conceptualización del programa.

Evaluación de Proceso: Analiza y valora el proceso de puesta en práctica del programa, es decir, cómo se ejecuta, pone en práctica y funciona.

Evaluación de Resultados:

De *outputs* o productos: Analiza y valora la calidad de lo que “produce” un programa (actuaciones, actividades, materiales, etc.).

De *outcomes* o efectos: Analiza y valora los efectos que son producidos por los *outputs* del programa.

SEGUN EL MOMENTO DE LA EVALUACION

Ex-ante: Se realiza antes de que el programa se ponga en marcha.

Ex-postfacto: Se realiza durante o una vez el programa ya se ha ejecutado.

SEGUN EL AGENTE EVALUADOR

Interna: Se lleva a cabo por parte de los responsables o profesionales del programa o por una unidad interna de la entidad que financia, planifica y/o ejecuta el programa, siendo en realidad una (auto)evaluación.

Externa: Se realiza por parte de alguien ajeno al programa, y que no tiene nada que ver ni con su financiación, ni con su planificación ni con su ejecución.

Fuente: elaboración propia

2.5. Otros tipos de evaluación

Además de las cuatro clasificaciones básicas (por las que cualquier evaluación puede ser clasificada), se quiere señalar aquí otros dos tipos específicos de evaluación y que tienen que ver con los propios procesos evaluativos: la evaluación de la **evaluabilidad** y la **metaevaluación**.

La evaluación de la **evaluabilidad**, como su propio nombre indica, consiste en valorar si un programa, política o intervención es “evaluable”. Este tipo de evaluación se realiza, lógicamente, antes de realizar una evaluación y sirve precisamente para determinar la viabilidad y utilidad de realizarla. Tradicionalmente los criterios de la evaluabilidad de un programa o política eran, sobre todo, de carácter técnico: sobre todo se han centrado en si los objetivos estaban claramente definidos, especificado y eran medibles. Este concepto de evaluabilidad era, en nuestro parecer, muy limitado y, en cierto modo ingenuo. De hecho, si se tuviera que renunciar a hacer evaluaciones porque el programa o política a evaluar tiene unos objetivos ambiguos, poco claros o contradictorios, probablemente se harían muy pocas evaluaciones. La superación de esta ingenuidad y, sobre todo, la experiencia práctica de hacer evaluaciones, llevan a cuestionar los criterios exclusivamente técnicos de evaluabilidad. Así, Chelimsky, en un reciente artículo en el expone sus reflexiones a partir de su experiencia al frente del departamento de evaluación de la GAO (*General Accounting Office*) estadounidense (Chelimsky, 1998), propone una serie de criterios para decidir hacer o no hacer evaluaciones (es decir, criterios de evaluabilidad) que tienen un carácter más contextual y político que los criterios tradicionales de evaluabilidad (tabla 2.5.). Es más, precisamente el criterio de tener unos objetivos no medibles no debe ser, según Chelimsky, razón para rechazar una evaluación.

Tabla 2.5

CRITERIOS DE EVALUABILIDAD

Adaptado de Chelimsky, E. 1998. "The Role of Experience in Formulating Theories of Evaluation Practice" en *American Journal of Evaluation*, vol.19, nº 1, pp.35-55.

RAZONES PARA NO RECHAZAR UNA EVALUACION

1. Aunque el tema sea polémico
2. Aunque los objetivos del programa o política no sean medibles
3. Aunque se requieran expertos específicos para la evaluación
4. Aunque el uso de la evaluación pueda ser impedido por los defensores del programa u otros agentes

RAZONES PARA SI RECHAZAR UNA EVALUACION

1. Si no se va poder realizar de una forma convincente por una limitación razonable de recursos económicos.
2. Si hay poca o ninguna base de conocimiento sobre la que construir y el tema es nuevo para el equipo.
3. Si no se pueden recopilar datos originales y no hay datos ya existentes disponibles.
4. Si no puede ser terminada a tiempo para que resulte útil.
5. Si no va a añadir nada nuevo, ni para el programa ni para la práctica de la evaluación.
6. Si no hay una posición o postura respecto a la política o programa sobre la que ejecutar las recomendaciones potenciales.
7. Si el interés público no puede contrarrestar el resto de las consideraciones anteriores.
8. Si las preguntas propuestas fueran demasiado numerosas, amplias, triviales, sesgadas y *no* pudieran ser modificadas.

La **metaevaluación** es la evaluación de la evaluación, es decir, consiste en la valoración de uno o varios procesos de evaluación. No hay que confundirla con la síntesis de evaluaciones, la cual consiste en sintetizar los resultados de varias evaluaciones, aunque en algunas tipologías (Chelimsky, 1989) equiparan los dos significados. En el caso de la metaevaluación se valoran los procesos completos de evaluación, no sólo los resultados que arrojan dichas evaluaciones. El concepto de metaevaluación se desarrolla con mayor profundidad en el capítulo tres, ya que este tipo de evaluación es la base sobre la que se asienta la investigación que se presenta en esta tesis doctoral.

2.6. Otras clasificaciones

Ya se ha comentado que algunos autores han elaborado sus propias tipologías de evaluación que a veces sólo se diferencian en cuestiones de matiz. Normalmente estas tipologías se basan inicialmente en el criterio de las fases de una política o programa aunque suelen incluir dentro de una misma clasificación algún otro criterio.

Por ejemplo, Ballart (1992) distingue entre evaluación del diseño o conceptualización, de la implementación, de la eficacia o impacto y de la eficiencia. En realidad, en el concepto de eficacia (grado en que se cumplen los objetivos) se podrían incluir cuestiones de proceso de puesta en práctica (¿se está llevando a cabo el programa como se pretendía? ¿se han puesto en marcha las actividades previstas?) como de resultados (¿se han obtenido los resultados esperados?). Ocurriría lo mismo con el concepto de eficiencia (referido a una buena utilización de los recursos), preguntándose sobre la adecuada utilización de los recursos tanto en el proceso de puesta en práctica como en lo referente a los resultados y efectos conseguidos.

Alvira (1991) realiza una clasificación con nueve tipos de evaluación: de necesidades, de diseño y conceptualización, de la evaluabilidad, de la implementación, de la cobertura, monitorización y seguimiento del programa, de resultados (efectos sobre los destinatarios), de impacto (efectos sobre la comunidad en general) y, finalmente, la evaluación económica. La cobertura es un aspecto que también puede plantearse al menos, en las fases de proceso y resultados. El seguimiento de un programa o política, si se hace desde un punto de vista evaluativo, formaría parte de la evaluación del proceso; además, se puede seguir y monitorizar un programa sin que eso implique que se dé una evaluación. Por último, la evaluación económica puede ser un aspecto a tratar tanto en una evaluación de diseño, como en una de proceso o de resultados.

De alguna forma similar a la clasificación de Alvira, se encuentra una de las clasificaciones más extendidas y probablemente más utilizadas en el ámbito de las políticas públicas (Chelimsky, 1989, Caudle, 1989, Anderson y Ball, 1987). Esta clasificación fue realizada por la *Evaluation Research Society* (Evaluation Research Society Standards Committee, 1982), y en ella se identifican 6 tipos de aproximaciones a la evaluación de programas. Esta clasificación se realiza basándose en las prácticas de evaluación, principalmente en los EE.UU.¹² y se basa también en los criterios de las diferentes fases o aspectos que se evalúan, el momento en el que se realizan, la principal finalidad que cumplen y el tipo de decisiones al que asisten (cuestiones sobre la formulación o para la ejecución de políticas o para el rendimiento de cuentas). Lo interesante de esta clasificación es que los seis tipos de evaluación “constituyen el repertorio cotidiano de la evaluación de programas. Por supuesto están lejos de ser las únicas aproximaciones, pero sí sería justo decir que representan un amplio acuerdo en el campo en los que se refiere a la práctica común en evaluación de programas, y que

¹² La *Evaluation Research Society*, junto con el *Evaluation Network* se reconvierte en el año 1986 en la AEA (*American Evaluation Association*). En la asociación actual, aunque se declara internacional y hay asociados no norteamericanos y grupos de trabajo sobre temas internacionales y multiculturales, esto no deja de ser una nota meramente exótica. Al igual que la organización antecesora, su principal actividad científica, académica y profesional se desarrolla en EE.UU.

son relevantes, en diversos grados, para las necesidades de los administradores públicos” (Chelimsky, 1989: 269). Estos seis tipos son los siguientes:

Análisis previo (*Front-end analysis*): Este tipo de evaluación se suele hacer antes de la toma de decisiones sobre un nuevo programa o política. Tiene, pues, una orientación prospectiva y trata de estimar la viabilidad y los posibles efectos de los programas o políticas a evaluar. Considera la evaluación de necesidades y de costes y la capacidad operativa del programa o política y de la(s) agencia(s) protagonista(s) en su implantación. Para Chelimsky, este tipo de análisis intentan resolver cuestiones relacionadas con la formulación de las políticas. Sería equivalente, en parte, a una evaluación *ex ante*.

Evaluación de la evaluabilidad: En esta clasificación a este tipo de evaluación, además de darle la finalidad de comprobación de factibilidad y utilidad de procesos de evaluación posteriores, se le ha dado el sentido de valoración sobre si las premisas y teoría del programa son congruentes con los objetivos y si es probable que las actividades cumplan dichos objetivos (lo que equivaldría a una evaluación del diseño) y también a la descripción de las características reales del programa ejecutado y comparación con los objetivos (lo que equivaldría a la evaluación del proceso de puesta en práctica).

Evaluación del proceso: Para Chelimsky esta es una forma de evaluación que o bien se puede hacer sola o en combinación con otro tipo. Su finalidad es describir y analizar los procesos de las actividades en ejecución del programa, con la idea de mejorarlas. Para la autora, esta evaluación contesta sobre todo preguntas sobre la ejecución de la política pero ayuda también a las evaluaciones de efectividad a contestar preguntas relacionadas con el rendimiento de cuentas y a explicar los resultados de las mismas.

Evaluación de efectividad o impacto: También es retrospectiva, como la anterior y trata de averiguar si los cambios observados son resultado del programa o de otros factores. Se trata de contestar preguntas relacionadas, sobre todo, con el rendimiento de cuentas, aunque para la autora, este tipo de evaluación también es necesario para preguntarse sobre la formulación y la ejecución del programa.

Seguimiento o “monitorización” del programa o del problema: En esta clasificación se categoriza de igual manera el seguimiento de un programa o intervención (ir comprobando si se está llevando a cabo tal y como estaba previsto y corregir las desviaciones) con el seguimiento del problema (comprobación periódica del problema que dio origen a la intervención), que son dos cosas muy diferentes. El primero requiere un seguimiento relacionado con la gestión cotidiana del programa y el segundo, una recopilación, procesamiento y análisis periódico de información que requiere, más bien, una infraestructura de investigación (a modo de observatorio). Este tipo de evaluación es continuo. Para Chelimsky este tipo de evaluación puede contestar bien preguntas respecto a la formulación, bien a la ejecución de las políticas, aunque dice que también los sistemas de información que se generen para el seguimiento de un programa pueden ser utilizados para contestar preguntas relacionadas con el rendimiento de cuentas siempre que se haga un análisis de series temporales. En nuestra opinión, los sistemas de información y seguimiento, aunque

son una condición inmejorable y en muchos casos necesaria para la evaluación, no implican necesariamente la realización de un proceso evaluativo, menos aún la aplicación de un análisis de series temporales (que implicaría, a su vez, todo un diseño de evaluación previo en el que una parte fuera la aplicación de dicha técnica).

Metaevaluación o síntesis de evaluaciones: En esta clasificación se habla de forma indistinta de estos dos tipos de evaluación y se le otorga el significado de síntesis de evaluaciones, es decir, una forma de evaluación que reanaliza los resultados de una o varias evaluaciones para determinar que es lo que se ha aprendido sobre una política o programa.

Finalmente, para Chelimsky, si se tienen cuenta las cinco funciones básicas de la actividad de la administración pública:

1. Planificar y racionalizar un programa y su evaluación
2. Ejecutar y gestionar un programa
3. Justificar la efectividad de la ejecución y gestión de un programa
4. Demostrar la efectividad de un programa
5. Medir la evolución de un problema o un programa

Y si se cruzan con los seis tipos de evaluación, se obtiene que el Análisis Previo sirve para las funciones 1. Y 5. La evaluación de la evaluabilidad para el 2. (desde nuestro punto de vista, también y sobre todo para la 1.), la evaluación del proceso para la 2. Y la 3., la evaluación de impacto para la 3., la 4. Y la 5., el seguimiento para las cinco funciones y la síntesis de evaluaciones para la 1., la 3. Y la 4.

Estos tipos de evaluación no se practican en la misma medida, realizándose mayores esfuerzos en las evaluaciones de proceso y de resultados. También para Chelimsky, aunque el análisis previo, la monitorización y la metaevaluación son herramientas potentes, sobre todo, para la formulación de políticas, no se llevan a cabo muchas evaluaciones de este tipo en la rama ejecutiva (Chelimsky, 1989: 271).

3. LOS ACTORES DE LA EVALUACIÓN

Una de las características de los contextos evaluativos es la presencia de diferentes personas y colectivos que tienen algo que ver con el programa o política que se pretende evaluar. Aunque el **cliente** o comanditario de la evaluación, es decir, quien encarga la evaluación y decide que hay que realizarla porque tiene unas determinadas necesidades de información, sea sólo uno, los **destinatarios** de la evaluación, es decir a quién va dirigida y quiénes son los posibles usuarios de la evaluación, son muchos más. Esta realidad significa que existe un elenco amplio de potenciales agentes críticos en la evaluación.

Estas personas implicadas en el programa o política a evaluar reciben el nombre de *stakeholder*, término de difícil traducción al castellano y que vendría a ser algo así como "el que tiene algo que decir o tiene algún interés en". Según el *Joint Committee* la definición de *stakeholder* en evaluación sería "los individuos o grupos que pueden estar implicados en o afectados por una evaluación" (Joint Committee, 1994) y para Greene serían los "grupos de gente que pueden verse afectados por el programa y aquellos cuyas decisiones pueden afectar el futuro del mismo (Greene, 1988). Aunque sería más correcto traducirlos como implicados en y afectados por la evaluación, por razones de economía lingüística traducimos *stakeholders* como *agentes críticos* de los programas a evaluar y, por lo tanto, de su evaluación.

Todos los agentes críticos en un programa, por el hecho de serlo, serían susceptibles de tener interés en su evaluación y, por lo tanto, ser destinatarios y posibles usuarios de los procesos evaluativos. ¿Quiénes son o pueden ser estos agentes críticos? Se podrían establecer cuatro grandes categorías:

- Patrocinadores, formuladores de políticas y responsables políticos
- Directores de los programas, gestores públicos y profesionales en general
- Destinatarios directos e indirectos (por ejemplo, padres de niños en programas educativos), usuarios potenciales del servicio, programa o política evaluada y el público o comunidad en general
- Otros: profesionales y responsables políticos de programas similares, expertos del sector o política que se evalúa, científicos sociales en general.

Guba y Lincoln (1989) van más allá y hacen una definición curiosa y amplia de los posibles agentes críticos o *stakeholders* de un programa y que, por lo tanto, tendrían algún interés en su evaluación. Distinguen entre agentes, beneficiarios y "víctimas" de los programas o políticas (el evaluando). En la tabla siguiente (tabla 2.6) se expone su listado.

Tabla 2.6. Listado de Agentes Críticos (*Stakeholders*) según Guba y Lincoln

<p>Guba y Lincoln (1989) <i>Fourth Generation Evaluation</i></p> <p>Listado de posibles agentes críticos (<i>stakeholders</i>). Los autores distinguen 3 tipos con sus diferentes subtipos:</p> <p>1. Los <i>agentes</i> que son aquellas personas involucradas en la producción, uso y puesta en práctica del programa o política a evaluar (el evaluando). Incluyen:</p> <ul style="list-style-type: none"> - los que formulan el evaluando - los que financian, en los niveles local, regional y nacional - los asesores de necesidades locales, es decir quienes identifican la necesidad que el evaluando va a intentar mitigar o remover - los que toman las decisiones y deciden utilizar o desarrollar el evaluando de forma local - los proveedores de materiales, instrumental e instalaciones necesarias - el cliente de la evaluación (el que contrata) - el personal que ejecuta el evaluando, como por ejemplo, los profesores, los oficiales de policía, las enfermeras, etc. (profesionales en general) <p>2. Los <i>beneficiarios</i> que se benefician de alguna manera del uso del evaluando. Incluyen:</p> <ul style="list-style-type: none"> - los beneficiarios directos, el grupo diana, las personas para las que fue diseñado el evaluando - los beneficiarios indirectos, las personas cuya relación con los beneficiarios directos se ve mediatizada, facilitada, fomentada o positivamente influida - las personas que ganan algo por el hecho de que el evaluando esté en uso, como por ejemplo, las editoriales de materiales o textos, las empresas proveedoras de servicios, etc. <p>3. Las <i>víctimas</i>, personas o grupos afectados negativamente por el uso del evaluando (y que pueden incluir, debido a los fallos del evaluando, uno o más grupos supuestamente beneficiarios). Incluyen:</p> <ul style="list-style-type: none"> - los grupos sistemáticamente excluidos del uso del evaluando, como por ejemplo, los jóvenes “normales” excluidos de los programas para los superdotados - los grupos que sufren efectos colaterales negativos, como por ejemplo, los estudiantes, y sus padres, que son trasladados a un colegio distante de forma que los alumnos con necesidades especiales ocupan sus puestos en el colegio cercano - las personas que se ven en desventaja política, como pérdidas de poder, influencia o prestigio por el uso del evaluando - las personas que sufren costes de oportunidad, como por ejemplo personas que hubieran preferido dedicar ese dinero a otras cuestiones o, por ejemplo, editoriales rivales.
--

En cualquier caso, la elección de unos u otros agentes críticos como destinatarios principales de la evaluación depende del enfoque que se adopte. Para Greene (1997: 25-26) los enfoques experimentales en la tradición de Campbell tratan de contestar

preguntas sobre resultados, responsabilidad y coste-beneficio para los formuladores de las políticas; el enfoque orientado a la utilización de Patton trata de contestar a preguntas de usuarios comprometidos e identificados de la evaluación que normalmente suelen ser los gestores de programas y sus asesores. La evaluación sensible y centrada en el cliente de Stake pretende contestar a los agentes críticos más cercanos al programa evaluado, normalmente los propios directores de programas y sus profesionales. El enfoque orientado a los consumidores de Scriven trata de cubrir las necesidades de los consumidores de programas y servicios, etc. Es decir, dependiendo de la orientación, se va a hacer especial énfasis en un tipo de actores u otros.

En cualquier caso y, sin necesidad de ser tan exhaustivos como Guba y Lincoln, se cree conveniente insistir en la pluralidad de actores potenciales de un proceso de evaluación y en su necesaria participación e inclusión en el mismo, dependiendo de la situación evaluativa concreta. Este reconocimiento de la pluralidad de actores ha ido paralelo al auge de los enfoques participativos en evaluación. La fundamentación de estos enfoques se basa, como ya se ha visto, por un lado, en otorgar una mayor legitimidad al proceso evaluador y, por otro, en aumentar la utilidad del mismo. Lo que sí parece evidente es que a mayor participación de diferentes actores en la evaluación, mayor probabilidad de uso de la misma. Para Ballart (1993: 203), la implicación de los principales actores, incluyendo entre éstos a los mismos beneficiarios de la política y a los que tengan algún interés en la misma es condición para un cierto pluralismo y además también para la integración de todas las informaciones, interpretaciones y juicios pronunciados en el marco del proceso evaluador.

Eso sí, tampoco hay que ser ingenuos y pensar que todos los agentes críticos van a estar deseando participar abiertamente y sin problemas en los procesos de evaluación, en los que, lógicamente, cada persona o colectivo va a tener sus propios intereses y en los que es inevitable que surjan conflictos. Si se reconoce el ambiente y contexto político de cualquier situación evaluativa, también significa aceptar que no siempre estos actores críticos (*stakeholders*) van a querer negociar amablemente con los evaluadores (Chelimsky, 1998, 44-45). Para esta autora los evaluadores deben predecir los posibles ataques de aquellos para estar preparados a una respuesta valiente y adecuada.

4. EL PAPEL DEL EVALUADOR

Lógicamente, uno de los principales actores de los procesos evaluativos son los propios evaluadores. Dependiendo de la orientación y el tipo de evaluación a realizar, su papel puede variar sustancialmente. Weiss (1998: 34) menciona tres tipos de posibilidades para encargar una evaluación, basándose en la experiencia estadounidense:

1. Evaluador en el *staff* u organigrama: dentro de una organización hay alguien dedicado (en exclusiva o no) a realizar evaluaciones. Puede ser una persona o un equipo, habiendo instituciones que tiene una propia unidad de evaluación, dependiendo de las dimensiones de los programas o políticas a evaluar. Normalmente son, por lo tanto, evaluadores internos, pero también puede ocurrir que una organización contrate a un evaluador como extensión de su propio *staff*.
2. Organizaciones externas especializadas en estudios de evaluación: en este caso se contrata a una organización externa, que pueden incluir consultoras, institutos de investigación e incluso, departamentos de universidad. En el caso de países con poca tradición y cultura de evaluación, como es España, en muchos casos ocurre que las organizaciones o equipos contratados están más especializados en el sector al que pertenece el programa o política a evaluar que en evaluación.
3. Petición de proyectos de evaluación (*requests for proposals*). Esta en realidad es una leve variación de la anterior ya que en Estados Unidos, los programas gubernamentales no permiten la elección de una organización sin que exista competición¹³; es decir, se dan concursos en una supuesta igualdad de condiciones. De manera paralela ocurre en España si el que contrata es una administración pública y el contrato pasa de una determinada cantidad.

En cualquier caso, el evaluador trabaja en contextos en los que hay también mucha otra gente, como se ha visto en el punto anterior y, dependiendo, de la situación, orientación y tipo de evaluación, puede interactuar con dichos actores de forma diferenciada. Para Weiss (1998: 97 y ss.) caben dos posibilidades que se podrían situar en los extremos de un continuo:

- a. El evaluador distanciado o despegado: este es el papel que se le ha otorgado tradicionalmente al evaluador. Aquí el evaluador procura situarse por encima

¹³ Además de los contratos, en los que la institución pública contratante establece los parámetros del estudio (lo que sería el equivalente al pliego de condiciones técnicas en España) y selecciona al equipo basándose en esos parámetros, y que es el procedimiento más común, también existe en Estados Unidos la posibilidad de becas o proyectos de investigación (*grants*) que se dan a investigadores o evaluadores que son los que inician y manejan su propio estudio. Algunos estudios de evaluación se producen bajo esta segunda modalidad.

de la situación y mantener la máxima autonomía y objetividad posible. Este evaluador busca los hechos y pone toda su confianza en la metodología.

- b. El evaluador comprometido o apegado: el evaluador trata de responder de forma sensible e interactiva a las preocupaciones e intereses de los grupos de agentes críticos. En este caso, el evaluador interactúa de forma cotidiana durante el proceso de evaluación, emplea un tiempo considerable en el lugar en el que se ejecuta el programa a evaluar y se implica con el grupo de profesionales.

El segundo tipo de papel se ha reforzado mucho con los últimos enfoques en evaluación y a raíz del fracaso de muchas experiencias de evaluación que defendían a ultranza un papel absolutamente distanciado para el evaluador en aras de la objetividad. Por un lado, con un papel distanciado resulta mucho más difícil para el evaluador captar la realidad y las preocupaciones de las personas implicadas en un programa. Por otro, resulta mucho más difícil para un grupo de profesionales aceptar a alguien que sitúa “por encima”, emitiendo juicios sobre una realidad que, en muchos casos, se siente no se conoce realmente; realmente, en estas situaciones los desequilibrios de poder del evaluador respecto a la gente del programa pueden llegar a ser percibidos como grandes. Por otro lado, la propuesta de evaluadores más comprometidos ha hecho que se puedan superar algunas visiones estrechas de los programas y políticas y públicas y se pueda mostrar las diferentes percepciones que existen en torno a las mismas. También ha sido una manera de aumentar el uso de las evaluaciones y otorgarle legitimidad a las mismas en la comunidad de programa (aunque esto haya supuesto restarle “credibilidad científica”, al menos en los primeros años de propuesta). Por último, se ha comprobado que con este tipo de evaluaciones, por el mero hecho de poner en marcha un proceso de evaluación tiene una serie de efectos colaterales beneficiosos (Weiss, 1998: 103).

Esta tendencia de avalar un papel comprometido para el evaluador ha tenido su punto álgido recientemente cuando han salido a la luz una serie de escritos en el mundo anglosajón sobre la *advocacy evaluation* (Greene, 1997), que podríamos traducir por “evaluación comprometida”. La fundamentación de esta idea se encuentra en la inevitabilidad del compromiso en la evaluación; se supone que este compromiso no es equivalente a sesgo partidario sino un compromiso de valor a un ideal de pluralismo democrático. Esta idea también tiene algunos detractores: Chelimsky (1998) señala que no es lo mismo el reconocimiento de ciertos valores por parte de los evaluadores que un compromiso partidario.

Lo que sí es evidente es que los evaluadores se mueven en un contexto político en el que deben saber manejarse. Para Palumbo (1987: 20-21) “algunos evaluadores (...) se pueden sentir más cómodos siendo técnicos que agentes de cambio, y todos los evaluadores deben asumir el rol de técnico al menos parte del tiempo”. Esto es a veces difícil porque los evaluadores se encuentran con imperativos que “son opuestos, aunque no irreconciliables. En primer lugar, deben entender y mejorar los programas; en segundo lugar, deben descubrir los “hechos” sobre el programa, incluso cuando éstos son negativos y, por último deben aportar conocimientos sobre la evaluación y la gestión de programas.

Por último, si tenemos en cuenta este ambiente en el que deben moverse y al que deben responder los evaluadores, éstos, además de formación metodológica en evaluación, deben conocer los procesos políticos y teoría de la organización y de grupos. Para Torres, Preskill y Piontek (1996), en la última década se ha reconocido en el campo de la evaluación que, además de un componente **cognitivo** (relacionado con los aspectos sistemáticos, cognitivos y "comparativamente objetivos" de la investigación y la indagación), existe un componente **intuitivo** (relacionado con aspectos intuitivos y cargados de significado por parte del evaluador, sus clientes, los agentes críticos en la evaluación y las organizaciones en las que están inmersos). Ambos componentes deben ser integrados para una buena labor evaluativa.

Esto tiene que ver con lo que debe ser la **formación de los evaluadores** y la institucionalización de la evaluación de la profesión y actividad científica diferenciada. En el mundo anglosajón, los programas de formación especializados para evaluadores son una realidad desde hace años (Altschuld y Engle, 1994) y en los últimos 10-15 años se ha avanzado mucho en el terreno de la consolidación profesional, a través del desarrollo de asociaciones profesionales y la creación de códigos deontológicos para evaluadores (Bustelo, 1998). Incluso se comienza a señalar la necesidad de acreditación de los evaluadores profesionales en Estados Unidos.

5. LAS CUESTIONES ÉTICAS EN EVALUACIÓN

En los últimos años se le ha prestado mucha atención a las cuestiones éticas relacionadas con la evaluación. Esta preocupación, como es lógico, ha sido paralela al reconocimiento del carácter político de la evaluación y, por lo tanto, al poder que ésta puede tener y a la necesidad de que el evaluador interactúe en el contexto en el que debe intervenir. Muchos de los más recientes manuales en evaluación de programas y políticas contienen un capítulo específico en cuestiones éticas (Patton, 1997; Weiss, 1998; NORAD, 1997).

Otro indicador de esta preocupación ética son los códigos deontológicos elaborados por las asociaciones profesionales, que en algunos casos, tienen ese objetivo implícito incluso en el título que les han dado (“Orientaciones para una actuación ética” en el caso de las asociaciones canadiense y australasiática”). Como ya se ha señalado anteriormente, la existencia de códigos deontológicos es, a la vez, un indicador y un elemento generador o provocador de una mayor consolidación institucional y profesional de una disciplina. Profesiones fuertemente consolidadas desde hace tiempo (como son la medicina, la abogacía o la arquitectura) o más recientemente (como la psicología) han desarrollado este tipo de códigos. Además de fomentar un sentido de pertenencia e identificación entre los profesionales, acotan y regulan la práctica profesional entre sus respectivos campos.

Pero un aspecto de enorme trascendencia, probablemente su aspecto más característico, respecto a los códigos deontológicos es que permiten la introducción y desarrollo de los temas éticos, que resultan clave en la práctica de la evaluación (Bustelo, 1998). No se trata sólo de elaborar teóricamente sobre estos asuntos como parte intrínseca al propio campo científico, sino también de orientar ante los conflictos y dilemas éticos que se van a producir, inevitablemente, en cualquier situación evaluativa. En cierto modo, se trata de afrontar esta realidad y de reconocer explícitamente el contexto político en el que está inmersa la evaluación. Este reconocimiento explícito se refleja en los códigos que se han elaborado. Por ejemplo, el *Joint Committee*, además de plantearse, con la elaboración de sus Normas, orientar el diseño, empleo y crítica de las evaluaciones y de animar a la utilización de una variedad de métodos en evaluación, pretende ayudar a los evaluadores a identificar y afrontar la realidad política. Textualmente este comité afirma que "las agendas políticas y el dinero son fuentes de poder que pueden corromper la evaluación en cualquier circunstancia. La adhesión a las Normas ayudará a minimizar las posibilidades de una mala utilización de ese poder por parte de los evaluadores o sus clientes" (*Joint Committee*, 1994: 4).

El propio hecho de afrontar las cuestiones éticas en evaluación se reconoce no sólo por parte de las asociaciones profesionales, sino también por los propios académicos del campo de la evaluación. Scriven señala que “es verdaderamente anti-ético dejar a la ética fuera de la evaluación de programas” (Scriven, 1993: 30). Esto significa que el evaluador debe ser formado también en cuestiones éticas ya que necesita ciertas habilidades en ética aplicada y al menos, un entrenamiento básico en

las normas legales que se deben seguir en estudios de evaluación. En muchas situaciones, de esta formación e incluso de la orientación que puedan proporcionar los códigos éticos no se desprenderán respuestas claras y automáticas ante los dilemas éticos que un evaluador pueda encontrarse. Como dice House (1994: 28), "tenemos este tipo de problemas (dilemas éticos) y debemos hablar de ellos, incluso aunque no podamos resolverlos". Precisamente, en la nueva revista de la Asociación Americana de Evaluación *The American Journal of Evaluation* que sustituye a la anterior *Evaluation Practice*, se ha abierto una nueva sección en "Retos éticos", dirigida por Michael Morris, y que consiste en que, cada número se presenta una situación o dilema ético y se le pide a uno o dos profesionales de la evaluación que discutan sobre la misma y señalen qué harían ellos si se encontraran en una situación similar.

En el capítulo que dedica a cuestiones éticas ("Evaluar con integridad") en su manual de evaluación, Carol Weiss (1998: 320 y ss.) señala seis ideas fundamentales sobre cómo se debe realizar una evaluación ética en una sociedad democrática. Estas son las siguientes:

1. Conocer el programa o política que se va a evaluar y su contexto
2. Mantener la alta calidad técnica y la relevancia del estudio de evaluación: la calidad en investigación se define en términos metodológicos. En evaluación, además, se debe asegurar que el estudio es relevante para la situación y el contexto concreto.
3. Utilizar el equilibrio y el juicio: los evaluadores trabajan en un contexto en el que el programa es prioritario y las condiciones no siempre son buenas para el estudio sistemático. Las elecciones del evaluador deben tener en cuenta: el tipo de información que quiere el cliente, los deseos y puntos de vista de otros agentes críticos, la naturaleza de las premisas del programa, las posibilidades de eso de la evaluación y las oportunidades de que se incorporen los resultados al conocimiento profesional y a la opinión pública. También tiene que ser sensible a los cambios que se produzcan en el contexto a lo largo del proceso de evaluación.
4. Considerar el uso desde el principio: las conclusiones deben ser escuchadas, respetadas y aceptadas, ya que serán la base para la acción.
5. Comportarse éticamente a lo largo del estudio: a veces las obligaciones éticas pueden llegar a ser contradictorias o entrar en conflicto. Hay que procurar, sobre todo, no distorsionar y no dañar sensibilidades.
6. "Sobrevivir y cosechar recompensas": con esta máxima, Weiss pretende demostrar que, a pesar de todas las demandas que se requieren del evaluador, la evaluación merece la pena ya que es un trabajo que nunca es aburrido y que, sobre todo, genera comprensión y mayor calidad de vida, ya que pretende mejorar la forma de abordar problemas públicos.

Como se puede observar, a las cuestiones éticas a las que se refiere Weiss van más allá del "comportamiento ético" al que se refiere en el punto 5. Es decir, evaluar éticamente implica el propio concepto y concepción de la evaluación. Ocurre lo mismo

en el manual de Patton (1997), en el vincula las cuestiones éticas a la naturaleza política de la evaluación y a su poder. Patton (Patton, 1997: 350-1) plantea cuatro “máximas políticas” que deben seguir los evaluadores, según su enfoque de la evaluación orientada a la utilización: 1. No toda la información es útil; 2. No toda la gente es usuaria de información; 3. Es más probable que la información dirigida al uso alcance el objetivo; y 4. Sólo la información creíble es poderosa en última instancia. Es evidente que estas cuatro máximas tienen claras implicaciones éticas.

Finalmente, y como ejemplo de la importancia de las cuestiones éticas no sólo en la evaluación en general, sino también en los manuales o guías de evaluación que se han elaborado últimamente de forma específica, mencionamos el manual para evaluar proyectos de ayuda al desarrollo elaborado por la Agencia Noruega de Cooperación (NORAD, 1997: 12-13). El cuarto punto del manual, “La ética de la evaluación” expone diez principios éticos que se basan, a su vez, en las líneas directrices elaboradas por la Agencia Canadiense de Desarrollo Internacional. En este manual se señala que la evaluación puede tener consecuencias importantes para todos los implicados en la ejecución de un proyecto y, por lo tanto, el evaluador debe cumplir una serie de requisitos, tanto profesionales como éticos, y que son:

1. Sensibilidad cultural: estar familiarizados con las características culturales del país
2. Anonimato y confidencialidad: se debe respetar ese derecho de los informantes
3. Responsabilidad en las evaluaciones: el equipo evaluador no debe realizar ninguna afirmación a menos que todos sus miembros hayan tenido la oportunidad de aprobarla o manifestar su desacuerdo con ella
4. Consideración con los informantes: avisar con tiempo y hacer un uso eficiente de su tiempo
5. Valores fundamentales: cuando sea difícil encontrar un equilibrio entre las tradiciones culturales, las prioridades del país receptor y los propios valores del país donante, habrá que guiarse por la Declaración Universal de los Derechos Humanos.
6. Integridad y omisiones: el evaluador tiene la responsabilidad de sacar a la luz cuestiones que no hubieran sido reflejadas en los términos de referencia.
7. Evaluación de personas: la evaluación de proyectos no incluye normalmente la evaluación de personas; en todo caso, la función de las personas y no ellas mismas
8. Errores e infracciones: cuando se descubran errores habrá que evitar citar personas concretas en el informe
9. Convalidar la información: con los participantes e implicados, aunque la última palabra la tenga el equipo evaluador.
10. Compartir los resultados: se podrán hacer sesiones informativas y resúmenes de carácter no oficial, manteniendo la discreción respecto a los contenidos finales.

6. CONCLUSIONES: HACIA UN ENFOQUE INTEGRADOR DE EVALUACIÓN DE POLÍTICAS PÚBLICAS

Una vez hecho un repaso general de la teoría de la evaluación de programas y políticas públicas, estamos en condiciones de señalar, desde nuestro punto de vista, cuáles son las características en las que se debería insistir en un enfoque integrador de evaluación de las políticas públicas. De esta manera, destacaríamos una serie de características que creemos clave para este enfoque integrador: Estas serían:

- Concepción amplia y global de las políticas y de la evaluación, situándola en un contexto más general del proceso de políticas públicas y superando visiones estrechas o limitadas.
- Delimitación clara de QUE se quiere evaluar, PARA QUE y PARA QUIENES, además de hacerlo con visión estratégica.
- Reconocimiento del carácter político en todas sus dimensiones, lo que implica otorgarles un carácter específico a la actividad científica de evaluar programas y políticas públicas.
- Visión práctica y orientada a la acción, consiguiendo que los estudios de evaluación sean realmente útiles en la práctica
- Continuidad, que no significa estar continuamente evaluando sino tener presente la evaluación desde un principio y de forma continuada; ésta es la manera de aumentar las probabilidades de que sea una herramienta potente.
- Flexibilidad metodológica, tanto en la forma de afrontar la evaluación como en las técnicas de recopilación y tratamiento de la información.
- Sistemas de información adecuados: esta es una de las condiciones que, en la práctica, facilita o dificulta en mayor medida los procesos de evaluación.
- Temporalidad y oportunidad en el tiempo, como factor clave en los procesos de evaluación. Este es uno de los factores que más condiciona y hay que tener en cuenta en evaluación.
- Participación de los agentes críticos, no sólo para legitimar y aumentar las probabilidades de uso, sino para tener una visión lo más completa y menos sesgada posible de las políticas públicas.
- Importancia de la comunicación y difusión de los resultados, como clave para la utilización de los procesos evaluativos.

Además de estas características generales, querríamos añadir tres cuestiones que nos parecen clave para la evaluación de políticas públicas en España.

1. En primer lugar, creemos clave darle una mayor importancia a la evaluación del diseño de las políticas y a todo lo relacionado con la identificación de la “teoría de programa” en términos de Weiss. Esta es el tipo de evaluación menos practicada. Esta claro que, en teoría, se ha superado la visión de la evaluación como exclusiva medición de la eficacia (es decir, hasta qué punto se cumplen los objetivos propuestos), pero en la práctica sigue dominando esta visión. En las políticas que nos ocupan, es evidente que no sólo debería interesar si los planes de igualdad de género se han cumplido, sino los propios planes en sí mismos: ¿Son el instrumento adecuado?, ¿Cuáles son los mecanismos por los que generan igualdad?, ¿Están relativamente equilibrados respecto a los objetivos de la política?, ¿Qué concepto de igualdad están dejando entrever? Una reflexión profunda sobre estos extremos, hará a los poderes públicos más conscientes de lo que están haciendo, hacia dónde van y hacia dónde deben ir.
2. En España, tiene que aumentar la práctica de la evaluación a muchos niveles. Si existiera una mayor práctica de evaluación, esto produciría un aumento de la cultura de evaluación. Se deben evaluar programas y también se deben evaluar, globalmente, las políticas públicas¹⁴. La evaluación de políticas es más difícil en el mercado que la de programas, ya que, en muchos sectores, no suele estar identificado un núcleo con visión global; quizá el mundo académico –a través de analistas de políticas públicas- pueda aportar algo a este respecto. Por otro lado, la evaluación de políticas se hace más difícil si no se han evaluado medidas, programas, servicios, etc.
3. Por último, hay que realizar un esfuerzo por aumentar la consolidación e institucionalización del campo de la evaluación de programas y políticas públicas. Para esto hace falta la existencia de asociaciones académicas y profesionales específicas, de programas de post-grado específicos, de publicaciones y foros especializados.

¹⁴ Para una discusión más profunda sobre este aspecto, véase el capítulo 9 de este trabajo.

CAPÍTULO 3

LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO EN ESPAÑA EN LOS NIVELES CENTRAL Y AUTONÓMICO

INTRODUCCIÓN

El objetivo de este capítulo es ofrecer una visión general de las políticas públicas españolas de igualdad de género en los niveles de gobierno central y autonómico. Esta visión general resulta imprescindible para analizar las evaluaciones que sobre dichas políticas públicas se han realizado y para comprender la función de evaluación en torno a las mismas.

En primer lugar, se definen y fundamentan las políticas de igualdad de género como una de las posibles respuestas de los poderes públicos ante la desigualdad de género, explicando también el papel que han cumplido tanto el movimiento de mujeres como el feminismo institucional en la definición y la puesta en práctica de dichas políticas. También se revisan los estudios que sobre feminismo institucional y políticas públicas de igualdad de género ha habido en España y se señala la falta de estudios concretos, y la necesidad de los mismos, especialmente de las políticas autonómicas en este área.

En segundo lugar, se reconstruyen las políticas de igualdad de género en España en los niveles central y autonómico desde una perspectiva histórica, haciendo especial énfasis en su rápido desarrollo, ya que aunque la institucionalización de las políticas de género comenzó aproximadamente una década más tarde que en otros países occidentales, la situación y consolidación de dichas políticas en España es hoy en día equiparable a la de dichos países. También se tratan en este capítulo la entrada en agenda del problema de la desigualdad de género y la evolución que ha ido sufriendo la definición de dicho problema y las consecuencias de esa diferente definición en el tipo de respuestas que se ofrecen. De igual manera, se describe el marco institucional en el que se desarrollan las políticas de igualdad central y autonómicas, analizando el tipo de organismos específicos que se han ido creando, así como su posición en el organigrama gubernamental. Por último, se realiza un repaso de los actores críticos que contribuyen a la formulación y ejecución de las políticas públicas de igualdad de género españolas.

Finalmente, se analiza el instrumento básico y concreto con el que se han articulado las políticas de igualdad de género en estos dos niveles de gobierno: los planes de igualdad. Los planes de igualdad consisten en un conjunto de objetivos, medidas y actuaciones que afectan a diferentes áreas gubernamentales y que son aprobados por los ejecutivos de cada

gobierno. Además de ofrecer un panorama sobre su proliferación y generalización en los gobiernos autonómicos, también se analizan su significado y evolución, así como sus fortalezas y debilidades como principal y prácticamente exclusivo instrumento de las políticas de igualdad de género.

1. CONCEPTO E HISTORIA DE LAS POLÍTICAS DE IGUALDAD DE GÉNERO

Las políticas públicas de igualdad de género se definen, en este trabajo, como el **conjunto de intenciones y decisiones, objetivos y medidas adoptadas por los poderes públicos en torno a la promoción de la situación de la mujer y de la igualdad de género entre mujeres y hombres**¹. La adopción del término genérico de “igualdad de género” tiene que ver con una definición más actual, comprehensiva y aceptada internacionalmente del problema que pretenden resolver las políticas públicas en cuestión (la desigualdad entre hombres y mujeres)². En él, se incluyen todas las acciones impulsadas por los poderes públicos y que tienen por objetivo conseguir una mayor igualdad de oportunidades entre mujeres y hombres, incidiendo sobre las condiciones socio-económicas y culturales que impiden y obstaculizan dicha igualdad.

1.1. Las difíciles relaciones feminismo-Estado

El origen de estas políticas públicas de igualdad de género se encuentra en la idea o el **reconocimiento de que el Estado o los poderes públicos pueden y deben hacer algo para paliar las desigualdades de género** y la discriminación de la mujer. Aunque, como veremos más adelante, las relaciones de las mujeres con el Estado han sido casi siempre difíciles y, de alguna forma, contradictorias, el impulso esencial para que el tema de la desigualdad de género se convirtiera en una cuestión pública, entrara en la agenda de los poderes públicos y se comenzaran a formular medidas al respecto, se debe al movimiento feminista de segunda ola de la década de los años 60 y la de los 70³. Este **movimiento**

¹ Esta definición no obvia una visión más amplia como la que se refleja en la definición de Dye, ya que una política pública puede ser “todo aquello que los gobiernos deciden hacer y no hacer” (Dye, 1975: 1). En este sentido, también entendemos como política pública de igualdad de género cuando los poderes públicos no asumen un problema como público o deciden no actuar en determinadas esferas por considerarlas del ámbito privado.

² Véase el punto 2.2. de este capítulo, en el que se hace un análisis sobre la evolución del problema público de la desigualdad a lo largo de la existencia de las políticas públicas de igualdad de género.

³ De acuerdo con la literatura especializada en el movimiento feminista, han existido dos grandes olas de activismo político feminista en la mayoría de los países occidentales. La primera ola coincidiría con el principio de siglo y los movimientos sufragistas. La segunda empieza en la mayoría de los países industrializados a finales de la década de los años 60 y a principios de la de los 70 (en España será un poco más tarde, con un retraso de aproximadamente 10 años). Algunas autoras (Stetson y Mazur, 1995: 20)

feminista, que estuvo protagonizado, entre otros, por el feminismo radical, explica las desigualdades de género por la existencia de un sistema patriarcal. El nuevo concepto entonces de patriarcado o sistema patriarcal implica un sistema básico de dominación sexual, que pasa por encima del de raza o clase social y que configura un sistema social organizado a partir de la superioridad del género masculino sobre el femenino. Esta nueva noción de sistema patriarcal explicaba las razones por las que la igualdad formal o de derecho, reivindicada por las sufragistas que conformaron el feminismo de primera ola, y conseguida a “grosso modo” con el derecho al voto de las mujeres en la casi totalidad de los países occidentales, no había sido suficiente. La verdadera igualdad implicaría, entonces, cambios más grandes que la igualdad de hombres y mujeres ante la ley y que debían incidir en ese sistema organizado socialmente sobre una división del trabajo en función del sexo biológico, y a partir del cual se han construido social y culturalmente los géneros. Estas nuevas ideas, aunque generan rupturas en el movimiento de mujeres debido al diferente papel que se le otorga al Estado en la teoría feminista –una parte del feminismo radical considera que el Estado no puede hacer nada por las mujeres-, también provocan la propia intervención del Estado, a través de la formulación de políticas de promoción y de acciones positivas para las mujeres. Es en Estados Unidos de América en dónde surge y tiene un mayor auge este movimiento feminista radical de segunda ola, y también dónde habían comenzado a dar las primeras acciones positivas y los primeros indicios del llamado “feminismo institucional”, a través del *Women’s Bureau*, el cual se crea a principios del siglo XX (Stetson, 1995: 254-271).

Sin embargo, y como se ha comentado antes, la intervención del Estado a favor de la mujer no gozó nunca del apoyo de todas las feministas. Ya se ha dicho que las **relaciones de las mujeres con el Estado fueron siempre difíciles** y según la idea de cómo deban ser dichas relaciones se han defendido posturas muy distintas: en un extremo, estarían las feministas que han aceptado que el Estado sea el instrumento primordial para la mejora de la situación de la mujer y, por otro, las que han abogado por que el feminismo forme parte exclusiva de la sociedad civil, rechazando cualquier intervención del Estado, ya que éste contribuye a la dominación masculina. Efectivamente, junto con el tema de la participación de las feministas en las organizaciones políticas (es decir, si debía darse o no la “doble militancia”, en partidos políticos y en organizaciones feministas), el tema del análisis sobre el papel que el Estado debe tener en la corrección de las desigualdades de género, ha sido el que ha producido más debates en el feminismo y ha creado mayores conflictos en el movimiento. Si se acepta la idea del sistema patriarcal, se debe aceptar, necesariamente, que el Estado moderno también forma parte de dicho sistema y que ha desarrollado mecanismos e instituciones para mantenerlo y reproducirlo. Esto es así desde sus inicios, ya que la Revolución francesa negó la condición de ciudadanas a las mujeres. Para las historiadoras feministas, el surgimiento y desarrollo del Estado de derecho, pues, a diferencia de los hombres, no benefició a las mujeres deteriorando aún más su situación

reflejan en su escritos cómo los historiadores hablan de una tercera ola en la actualidad de algunos países (ej: EE.UU.). Otros textos hablan del “feminismo de tercera ola”, a partir de la década de los años 80, y que se centraría en la crítica al uso monolítico de la categoría “mujer”, reconociendo así la diversidad de las mujeres y que afecta a otras variables que interactúan con el género, como son el país, la raza, la etnicidad y la preferencia sexual (Miguel, 1995: 254-5).

política ya que, en el Antiguo Régimen, al menos, las relaciones familiares tenían gran importancia en la vida política. El Estado, entonces, ha tenido un papel importante en la creación y mantenimiento de la situación de las mujeres⁴. Sin embargo, y debido a las propias movilizaciones feministas, los gobiernos occidentales comienzan a aplicar medidas para corregir las desigualdades de género por lo que se produce, según algunas autoras, la paradoja de que algunas instituciones del Estado generan o apoyan directa o indirectamente la discriminación de género, mientras que otras intentan corregir la situación (Astelarra, 1990: 211).

El análisis que sobre el Estado hace la teoría feminista está marcado, precisamente, por la idea de Catherine MacKinnon de que **el feminismo no tiene una teoría del Estado** (MacKinnon, 1995: 177). Para esta autora, el feminismo tiene capacidad descriptiva sobre la forma en la que el Estado aborda las diferencias de género, pero le falta un análisis del Estado como “jerarquía de género”. De esta manera, “la práctica feminista habría oscilado entre la adopción de una teoría liberal del Estado por una lado, y una teoría del Estado ‘de izquierdas’ (marxista y postmarxista) por otro” (Threlfall, 1990: 215).

Siguiendo en esta línea, Stetson y Mazur (1995: 6-10) dividen las críticas feministas al Estado en dos tipos o corrientes: al Estado de Bienestar y al Estado liberal. Una parte de las **críticas** de la literatura feminista se concentran, pues, en el **Estado de Bienestar**. Así, las feministas marxistas británicas (McInstosh, 1978; Wilson, 1977) reclaman que el Estado de bienestar es un instrumento creado por el Estado capitalista que oprime particularmente a la mujer; de esta manera, la dependencia de las mujeres de sus maridos se sustituye por una dependencia del Estado capitalista. Las feministas escandinavas (Hernes, 1990; Borchost and Siim, 1987) reconocen que algunas políticas de bienestar son esfuerzos bien intencionados para ofrecer beneficios sociales a los trabajadores y sus familias y que el énfasis en las madres trabajadoras ha aumentado las opciones y la independencia económica de las mujeres; sin embargo, estas políticas han institucionalizado el papel de las mujeres como madres y trabajadoras, lo cual contribuye a la doble carga ya que no se han alterado los papeles tradicionales de los hombres. Sus conclusiones son que las mujeres se convierten en usuarias sin haber alcanzado realmente el estatus de ciudadanas; por lo tanto, aunque las mujeres hayan alcanzado libertades respecto a las estructuras familiares patriarcales, el Estado de bienestar ha fortalecido su subordinación a la esfera pública. De esta manera, “la vida de la mujer depende más de la política de los Estados que la de los hombres, y está determinada por la misma” (Hernes, 1990: 47). Threlfall señala que existe una segunda etapa en la crítica feminista al Estado de bienestar que consiste en una nueva reflexión hecha a partir de los años de crisis económica y de políticas de recorte del gasto público coincidente especialmente con los años de thatcherismo inglés: se reconoce

⁴ Existe una línea de trabajo sobre género y ciudadanía, de la que es un buen exponente Elisabeth Meehan, quien realiza una reconceptualización de la noción de ciudadano a la luz de las nuevas realidades europeas. Según Meehan, la ciudadanía es un concepto que ha evolucionado a lo largo de la historia, pero todavía es una categoría que excluye. Viejas ideas como la incapacidad de las mujeres de pensar en el interés público y como la división del trabajo en las esferas pública y privada tienen todavía influencia en la política europea y afectan a las posibilidades de las mujeres para actuar como ciudadanas (Meehan, 1993: 24). Sobre ciudadanía y mujeres, véanse también Sassoon (1998), Woodward (1998), Gallego (1998), y todos los artículos en la obra colectiva editada por Paloma de Villota (1998).

entonces que, aún siendo discriminatorios en su concepción y puesta en marcha, el recorte en los servicios sociales públicos afecta especialmente a las mujeres, debido a que son éstas las que más los utilizan y, por lo tanto, dichos recortes son un duro golpe al bienestar de las mujeres (Threlfall, 1990: 219).

Otra parte importante de las **críticas** hacia el Estado por parte de la literatura feminista, son las referidas **al Estado liberal**. Estas críticas van dirigidas no sólo hacia el Estado liberal sino también hacia el feminismo liberal que es el que ha reclamado, desde un principio, cambios en las leyes y la integración de las mujeres en el ámbito público. Según estas críticas, subyacente a esta concepción liberal, se halla la idea de que el Estado es neutral y puede responder ante cualquier grupo que desarrolle recursos políticos. Y esto, según Ferguson (1984) y MacKinnon (1995), es un mito que sirve, además, a los intereses de los más poderosos. La idea es que, independientemente de las leyes que se aprueben o de las mujeres que accedan a puestos públicos, el Estado refleja, promueve, sostiene y responde ante una jerarquía de dominación masculina y subordinación femenina. Para estas autoras, las burocracias son la misma esencia de la desigualdad; por eso las políticas lanzadas por dichas burocracias no pueden producir igualdad. Además, el discurso burocrático en sí mismo produce clientes y no participantes, haciendo que las mujeres acaben subordinadas al Estado⁵.

Otras autoras establecen diferentes tipologías que incluyen las distintas aproximaciones que se pueden encontrar en la teoría política feminista y que implican, lógicamente, diferentes formas de definir el problema de la desigualdad de género y, por lo tanto, de concebir el papel que los poderes públicos deben adoptar respecto al mismo⁶. Joni Lovenduski (1995: 202-203) establece una de las tipologías más completas, con cinco categorías distintas y corrobora, además, la idea de MacKinnon de que hasta la fecha no existe una teoría feminista del Estado. Para la autora, se han desarrollado **diferentes perspectivas** dentro del movimiento, ya que se han fijado en elementos diferentes del Estado; ella señala que se puede hablar de **cinco paradigmas** distintos en la literatura feminista, a saber: 1. Según la *visión liberal* el Estado es neutral y sólo se le puede reprochar desigualdad entre los sexos si se ponen obstáculos legales específicos para la igualdad de las mujeres (como, por ejemplo, ausencia de sufragio femenino). 2. Según el *paradigma radical* el Estado es masculino y nunca podrá dar poder a las mujeres, por su

⁵ Las críticas feministas a la ciencia de la administración son también claras y se refieren a las teorías de Max Weber y también a las críticas a la profesionalización, liderazgo e institucionalización de la autoridad. Lovenduski (1997: 208-209) menciona las teorías de Jones (1993), Ferguson (1984) y Bologh (1991). Según estas autoras, el burócrata racional de Weber es, al fin y al cabo, una constructo de una masculinidad concreta idealizada; en su mundo, las mujeres se mantienen en la esfera privada bajo la protección de los hombres. En definitiva, las organizaciones burocráticas institucionalizan la preponderancia de masculinidades concretas, potenciando y otorgando ventajas de este modo a algunos hombres sobre todas (o casi todas) las mujeres y otros hombres.

⁶ Castells (1996: 21-25) establece tres grandes perspectivas de la teoría política feminista: la teoría feminista liberal, la teoría feminista marxista y socialista y la teoría feminista radical. Para ello se basa en la clasificación de Alison Jaggar (Jaggar, 88) que identifica cuatro tendencias separando la teoría marxista de la socialista (liberal, marxista, socialista y radical). Para Castells la distinción entre teoría feminista marxista y socialista resulta poco clara.

propia naturaleza, ya que está pensado para el poder de los hombres. Este paradigma rechaza al Estado por masculino. 3. Según el *paradigma marxista feminista* el Estado no oprime específicamente a las mujeres, ya que éstas están oprimidas como consecuencia de su posición de clase; es decir, las mujeres de clases dominantes no están oprimidas. El Estado es capitalista, pero no patriarcal, siendo el patriarcado un sistema de dominación masculina sobre la mayoría de las mujeres y algunos hombres. 4. Según *los sistemas duales*, que son una mezcla entre el feminismo radical y el marxista, el capitalismo y la dominación masculina están combinados en el Estado para oprimir a las mujeres. Y, por último, 5. La *visión postestructuralista* aboga por que no se puede asumir lo que hace el Estado por las mujeres sin analizar sus diferentes partes y los efectos que cada una produce en las mujeres, bien sean positivos o negativos. El Estado y, por lo tanto, las organizaciones estatales, son espacios de carácter cambiante en el que compiten discursos tanto internos como externos. Para Lovenduski, ninguna de estas teorías explica adecuadamente el papel del Estado en las políticas de género y, cada una, tiene sus ventajas e inconvenientes. La teoría feminista liberal nos muestra la importancia de las leyes para las mujeres aunque no tiene en cuenta cómo funcionan realmente. Si se asumen que todos somos iguales ante la ley se pueden institucionalizar las desigualdades. La teoría radical, a pesar de ser simple y no ofrecer un marco para la comprensión de por qué las mujeres se benefician a veces de la acción del Estado, nos alerta del modo en el que las normas masculinas se introducen en las instituciones para gobernarnos. El paradigma marxista muestra lo importante de las relaciones económicas y los sistemas duales, cómo los sistemas económicos y de género se combinan para afectar la posición de las mujeres. Sin embargo, la teoría marxista ignora dicha interacción y también la importancia de las circunstancias familiares para su estatus de clase. Tampoco explican los sistemas duales por qué algunas mujeres se benefician a veces de la acción del Estado, cuestión que sí permite la visión postestructuralista que, sin embargo, presenta muy pocas posibilidades para la generalización.

En cualquier caso, parte de la **literatura feminista más reciente** es algo menos radical respecto al posible papel del Estado. Actualmente hay posturas que ven al Estado de forma menos monolítica y como un aliado potencial para los diferentes y diversos intereses de las mujeres (Stetson y Mazur, 1995: 6). En palabras de Lovenduski: “El problema es que es imposible evitar el Estado porque está muy presente en la mayoría de las áreas importantes de la vida moderna. En este sentido, sostenemos que las mujeres necesitan del Estado y que es importante para el movimiento feminista entender esto” (Lovenduski, 1997: 202). Incluso se reconoce la necesidad de una visión más amplia, desde una perspectiva del análisis de políticas públicas, en la que no se mire hacia el Estado analizando exclusivamente su papel sino que se contemple, de forma más global, la relación entre el Estado y la sociedad. Para Astelarra (1998: 223), por ejemplo, las demandas que hacen las feministas requieren la concertación Estado-sociedad ya que implican revisar el entramado comunitario y de la familia y sus relaciones sociales; por lo tanto, las feministas no pueden dirigirse sólo al Estado, sino que deben revisar la relación Estado-sociedad.

1.2. El Feminismo institucional

A pesar de la debilidad organizativa del movimiento feminista de las décadas de los años sesenta y setenta, y de la fragmentación del mismo debido a las diferentes concepciones sobre cuáles debían ser las relaciones del movimiento con el Estado y si éste debía intervenir o no en pro de las mujeres, el impacto en las agendas políticas de las reivindicaciones feministas fue grande, especialmente si se tiene en cuenta la buena acogida que las nuevas ideas de intervención a favor de las mujeres tuvieron en los organismos internacionales, especialmente en la Organización de las Naciones Unidas (ONU), a partir del año 1975⁷. Esto se tradujo en un “contagio” a otras organizaciones internacionales, y esta nueva preocupación en el nivel internacional tuvo su reflejo, a su vez, en las diferentes naciones, influyendo, definitivamente, en la formulación de políticas de igualdad de género en la mayoría de los países desarrollados. Es decir, las feministas fueron especialmente eficaces como **“introductoras de agenda” o *agenda setters***⁸ del problema de las desigualdades de género en las diferentes políticas nacionales. Compartimos aquí la idea de Granados (1999: 159) de que, en general, “tanto las definiciones del problema de la desigualdad, como su entrada en agenda y la toma de decisiones al respecto no han tenido *historias nacionales*. Ha sido el movimiento feminista el que ha presionado por que la desigualdad entre los géneros fuera tenida en cuenta por los poderes públicos, y el que ha conseguido que, primero en países aislados, Estados Unidos casi siempre, y después generalizado a través de la actuación de determinadas organizaciones internacionales, el problema de la desigualdad entre los géneros fuera incluido en los procesos de toma de decisiones”.

Una de las posibles respuestas⁹ de los poderes públicos de las diferentes naciones ante el problema de la desigualdad apuntado y destacado por el movimiento feminista y las organizaciones internacionales, es lo que se ha venido a denominar **“feminismo institucional”** (Lovenduski, 1997; Granados, 1999), **“feminismo de Estado”** (Stetson y Mazur, 1995; Valiente, 1994) o **“femocracias”** (Threlfall, 1990). Este feminismo institucional o de Estado, que se generaliza a partir de la década de los años 70 y la de los 80 en la mayoría de los países occidentales, puede definirse como el “establecimiento de

⁷ Este año fue declarado el Año internacional de la Mujer y se celebró en México la I Conferencia Mundial sobre las Mujeres, y fue seguido del Decenio de las Naciones Unidas para las Mujeres (1976-1985).

⁸ Por este término se entiende los colectivos o movimientos que consiguen llamar la atención suficientemente sobre un tema en cuestión (en este caso, la desigualdad de género) como para que adquiriera relevancia pública y sea considerado por parte de los poderes públicos como susceptible de ser tratado y abordado como problema público, es decir que entre en la agenda de actuación de los poderes públicos y, por lo tanto, éstos consideren que se debe hacer algo al respecto.

⁹ Para Threlfall (90: 226) se han dado dos tipos de estrategia feminista con efectos transformadores sobre las administraciones públicas: una es la creación de “femocracias”, que haría referencia a lo que aquí denominamos feminismo institucional o de Estado y que consiste en la creación de organismos dedicados a la gestión de programas de apoyo a la mujer y la otra sería el cambio del personal mismo del Estado y de los organismos públicos y semipúblicos a través de la política de igualdad de oportunidades en el acceso a los cargos en la Administración. Esta segunda estrategia está menos generalizada, aunque la autora menciona las experiencias en Reino Unido y Australia en los niveles local y regional respectivamente.

instituciones cuyo principal cometido es impulsar la formulación y puesta en práctica de medidas para alcanzar un mayor grado de igualdad entre las mujeres y los hombres" (Valiente, 1994: 1), o como mencionan Stetson y Mazur, dándole una dirección más concreta al término, es el fenómeno de institucionalización del feminismo en organismos públicos que promueven una agenda de igualdad de la mujer, es decir, utilizan el concepto para denominar "la institucionalización de intereses feministas" (Stetson y Mazur, 1995: 10). También Lovenduski (1997: 207) considera que "el feminismo institucional consiste en las estructuras, procesos y prácticas que resultan de la institucionalización del feminismo en el gobierno"¹⁰. Este feminismo institucional ha sido en más de una ocasión controvertido y criticado por las feministas radicales, debido a la idea de que las estructuras gubernamentales eran un componente importante del sistema patriarcal y ser parte de las mismas significaba aliarse con el poder masculino y por la necesidad de diferenciarlo del propio movimiento feminista. Hernes, por ejemplo, habla del "feminismo desde arriba" como respuesta oficial, en la forma de feminismo estatal, a la agitación feminista o "feminismo desde abajo" (Hernes, 1990: 23). El término "**femócrata**" para designar a las personas, casi todas mujeres, que dirigen y trabajan en los organismos de igualdad, aparece por primera vez en el movimiento feminista australiano, en la década de los años 70, y se utilizaba para designar peyorativamente a las feministas que se habían dejado cooptar, ocupando puestos en los organismos de igualdad. Este término que originariamente tenía cierto carácter irónico y de autodesprecio, se ha convertido en Australia en una imagen del experimento australiano de intervención feminista en el Estado (Sawer, 1995: 22). Por lo tanto, este carácter peyorativo inicial ha ido desapareciendo y actualmente se utiliza, en general, para las personas que trabajan en los organismos de igualdad, siendo un término que utilizan las propias "feministas de Estado" para referirse a ellas mismas (Valiente, 1996).

En cualquier caso, y como ya se ha mencionado anteriormente, debido en parte a la fuerza de la segunda ola del movimiento feminista y a la influencia de los organismos internacionales, unido a la concepción generalizada de una necesaria intervención pública en los problemas sociales, a partir de las décadas de los años 60 y 70 en la mayoría de las democracias occidentales y, a partir de la década de los 80 en España, se implantan y desarrollan diferentes **instituciones públicas** con el fin de promocionar la situación de la mujer y promover la igualdad entre los sexos, y esto es lo que se denomina feminismo institucional. Estas instituciones varían en tamaño, recursos, estructura orgánica y posición en el organigrama gubernamental, pudiendo tomar la forma de ministerios, secretarías de Estado, comisiones interministeriales, agencias, organismos autónomos, departamentos o unidades administrativas dentro de ministerios o departamentos, etc.¹¹. En la mayoría de los

¹⁰ Para Lovenduski existe una relación entre el feminismo institucional y los sistemas de representación democrática. Para esta autora "un indicador importante del lugar que ocupan las mujeres en un sistema de representación democrática es el grado de institucionalización de los conflictos de género. El feminismo institucional es la manifestación organizativa de su institucionalización. Las interesantes cuestiones que plantea este concepto tienen que ver con las políticas de legitimidad, representación y organizaciones (Lovenduski, 1997: 208).

¹¹ Véanse el listado de organismos de igualdad en Europa occidental consignados en el cuadro 10.1 en Lovenduski, 1997: 206. En el se reflejan los nombres de tres organismos establecidos antes de 1970 (en

países occidentales se da este feminismo institucional, además de en el nivel nacional, en los otros niveles gubernamentales (regional y local). Para algunas autoras, este feminismo institucional no sólo consiste en los organismos de igualdad que trabajan en nombre de los intereses de las mujeres, sino que también incluyen los pactos de mujeres que se han dado en los países nórdicos y la formación de grupos de presión o *lobbies* al estilo americano (Miguel, 1995: 253). Lo que está claro es que **este feminismo institucional ha ido adquiriendo progresivamente fuerza en estos últimos años**, de forma paralelamente inversa al dinamismo del movimiento feminista de segunda ola (Lovenduski, 1997: 204; Granados, 1999: 190; Miguel, 1995: 253). Para Lovenduski, desde 1970 se ha dado un continuo crecimiento de la presencia de mujeres en la dirección de instituciones políticas, un cambio en las agendas políticas para incluir temas de mujeres y el desarrollo del feminismo institucional; y estos cambios han sido el “resultado de acciones deliberadas de las mujeres para influir en política” (Lovenduski, 1997: 204). Como hemos visto antes, esto no significa que el feminismo institucional y su decidido abandono de la apuesta por situarse fuera del sistema no haya recibido críticas de otros sectores del movimiento feminista (Miguel, 1995: 253). De hecho, y como se verá más adelante, en España las relaciones entre femócratas y feministas no han sido especialmente fáciles (Valiente, 1996) y un indicador de ello puede ser la inicial confusión y la posterior valoración que se realizaba del papel del IM en el proceso preparatorio de la IV Conferencia Mundial de la Mujer en Pekín, por parte del movimiento feminista (Solana et al., 1997).

Este crecimiento del feminismo institucional genera un especial interés por las instituciones estatales en los estudios de género desde el área de la ciencia política. A partir de los años 90, comienzan a aparecer investigaciones cuyo objeto de estudio es este feminismo institucional en el mundo occidental, el papel que dicho feminismo cumple respecto a las políticas públicas de igualdad y sus relaciones con el movimiento feminista (Meehan y Sevenhuijsen, 1991; Stetson y Mazur, 1995; Gardiner, 1997). Una de las obras emblemáticas de esta nueva perspectiva del “feminismo de Estado” es *Comparative State Feminism*, de Stetson y Mazur (1995), en la que realizan un análisis comparativo del feminismo institucional en 14 países. Esta mirada hacia las agencias o estructuras dedicadas a la mujer, rompiendo en cierto modo la visión anterior más monolítica y global del Estado, ofrece una nueva perspectiva de análisis del Estado y de su capacidad para conseguir objetivos feministas. Como mencionan algunas autoras, para desarrollar una teoría feminista del Estado, se necesita menos abstracción y más estudios sobre "el alcance y el contexto de la acción gubernamental y sus consecuencias en la posición de las mujeres" (Dahlerup, 1987: 108). El trabajo empírico de Stetson y Mazur es una referencia obligada en este sentido. Estas autoras hacen un análisis comparativo a partir de 14 estudios de caso, entre los que se encuentra España, realizados por investigadoras que ya habían realizado previamente algún tipo de estudio relacionado con el feminismo de Estado en su país. El objetivo de la investigación era analizar hasta qué punto las estructuras

Austria y Reino Unido), veinticinco entre 1970 y 1980 y otros diez entre 1981 y 1990. Curiosamente, no consigna la creación del IM en España en 1983.

gubernamentales dedicadas a la mujer (*women's policy machinery*)¹² pueden ser consideradas como verdadero feminismo de Estado, es decir, hasta qué punto han influido en la formulación e implantación de políticas feministas (que promocionan la situación de las mujeres y/o socavan los patrones de la jerarquía de género) y hasta qué punto han permitido el acceso de los grupos no gubernamentales de mujeres para ejercer influencia en dichas políticas. Es decir, la existencia de estructuras gubernamentales dedicadas a la mujer no implica la existencia de un verdadero feminismo institucional o de Estado. Las autoras concluyen que todos los países, excepto Polonia, muestran evidencia de feminismo de Estado, es decir, todas las estructuras gubernamentales dedicadas a las mujeres son feministas en mayor o menor grado. Establecen una tipología de feminismo de Estado basándose en dos categorías (nivel de influencia en las políticas y nivel de acceso de grupos de mujeres a las políticas) y concluyen que hay mayor grado de feminismo de Estado cuando ocurren dos condiciones: cuando el Estado se define como un espacio de justicia social y tiene la capacidad estructural de institucionalizar las nuevas demandas de igualdad, y por otro lado, cuando la sociedad apoya ampliamente organizaciones feministas que retan las jerarquías de sexo a través tanto de movimientos radicales como de movimientos más moderados y reformistas dentro de los partidos y sindicatos. Es decir, los factores que contribuyen a explicar el carácter feminista de los organismos de igualdad tienen que ver con la forma en la que se estableció el organismo, su estructura y forma organizativa, la cultura política de la sociedad (sobre todo en lo relacionado a cómo se concibe el Estado) y la política del movimiento de mujeres.

En el análisis comparativo de Stetson y Mazur, y según el cruce de los dos parámetros empleados, Australia, Dinamarca, los Países Bajos y Noruega son países en los que el feminismo institucional tienen un fuerte arraigo, Irlanda e Italia se encuentran en niveles bajos de feminismo institucional y el resto de los países se encuentran en situaciones intermedias. España y Polonia son los dos casos elegidos para el contraste, ya que representan recientes cambios políticos (el resto de los casos fueron seleccionados a través del mecanismo de comparación de casos más similares y escogiendo países con sistemas democráticos estables desde 1950). El **análisis del caso español**, basado en el estudio de caso realizado por Celia Valiente sobre el feminismo de Estado español en el nivel central es clasificado por Stetson y Mazur en la tipología de aquellos países que tienen un alto nivel de influencia en las políticas pero un bajo nivel de acceso a los movimientos de mujeres¹³. Es decir, el caso español, aunque más tardío en su desarrollo, es comparable con el resto de los países industrializados democráticos y es clasificado en la misma tipología junto con Suecia, Gran Bretaña y Francia.

¹² El término *women's policy machinery* es utilizado por Stetson y Mazur para describir "cualquier estructura establecida por el gobierno y cuyo propósito principal es el de la mejora de la situación social de las mujeres" (p.3).

¹³ Australia, Holanda, Noruega y Dinamarca se clasifican en el tipo 1, en el que tienen tanto alto nivel de influencia en las políticas como alto nivel de acceso a las organizaciones de mujeres. En el tipo 2 y junto a España, se encuentran Suecia, Gran Bretaña y Francia; en el tipo 3, en el que tienen un alto nivel de acceso pero un bajo nivel de influencia en las políticas, se encuentran Alemania, Estados Unidos y Canadá y en el tipo 4 (bajo nivel en las dos categorías) estarían Irlanda e Italia. Polonia es el otro caso para el contraste en el que no se encuentran signos de feminismo de estado.

Por lo tanto, se puede decir que el estudio del **feminismo institucional ha estado inextricablemente unido al análisis de las políticas públicas de igualdad de género** y al papel del movimiento feminista en dichas estructuras y políticas. Se podría decir, pues, que la existencia de algún tipo de estructura gubernamental de igualdad es condición necesaria para la formulación y ejecución de lo que denominamos políticas de igualdad de género, entendidas éstas como se han definido al inicio de este capítulo (como el conjunto de intenciones y decisiones, objetivos y medidas adoptadas por los poderes públicos en torno a la promoción de la situación de la mujer y de la igualdad de género entre mujeres y hombres). El objeto de este capítulo es realizar un análisis de las políticas de igualdad en España en el nivel central y autonómico.

El análisis tanto del feminismo de Estado y su impacto como de las propias políticas de igualdad **se ha centrado en España, sobre todo, en el nivel central** (Astelarra, 1990; Threllfall, 1994; Valiente, 1995, 1996 y 1997; Gil Ruiz, 1996; Martínez, 1997; Sensat y Varella, 1998). Se han realizado algunos estudios autonómicos en el caso andaluz y vasco por parte de jóvenes politólogas (Granados, 1997 y 1999; Elizondo y Martínez, 1995), pero son todavía a todas luces insuficientes¹⁴ y son prácticamente inexistentes los estudios del feminismo institucional en el nivel local, con alguna pequeña excepción (Valiente, 1999). Sin embargo, dada la actual estructura del Estado español y su desarrollo en los últimos 25 años, no se puede obviar el nivel autonómico en el análisis del impacto de las políticas públicas de igualdad de género y de las estructuras gubernamentales dedicadas a las mujeres. Especialmente, si tenemos en cuenta que las competencias relacionadas con la “promoción de la mujer” iban en el paquete inicial de competencias cedidas a las comunidades autónomas desde el inicio en su misma creación, incluso para las que accedieron a la autonomía por la “vía lenta” o a través del artículo 143 de la Constitución. Es decir, la competencia en temas de mujer ha sido característicamente autonómica desde el comienzo del Estado de las autonomías.

En el **nivel internacional** también existe una tendencia al estudio del feminismo institucional a niveles nacionales, renunciando a profundizar en las estructuras regionales y locales a cambio de obtener análisis globales en cada una de las naciones, especialmente si los estudios son de carácter comparado. En el análisis de Stetson y Mazur, por ejemplo, los tres casos que se incluyen en el tipo 3 (baja influencia en las políticas, alto acceso de los grupos de mujeres a dichas políticas) son Alemania, Estados Unidos y Canadá, precisamente tres estados federales. En el caso de Alemania, por ejemplo, existen una gran cantidad de estructuras gubernamentales dedicadas a las mujeres al nivel regional (de los *Länder*) y al nivel local (aproximadamente 1.100 después de la unificación). Esto hace que haya una gran variedad de estructuras y que sea "imposible dar una única valoración del impacto de dichas estructuras" (Ferree, 1995: 105-106). En definitiva, lo que queremos aquí señalar es que el hecho de que no se pueda hablar de una alta influencia en las políticas en el nivel nacional, no tiene por qué decir que no la haya en el nivel regional o en el local, aunque haya una alta variabilidad y no se pueda generalizar.

¹⁴ Para tener una mejor idea del estado actual de los Estudios de género o *Women's studies* en ciencia política en España y la evolución que han tenido, consultar Granados, 1999: 74-105; Uriarte, 1997 y Valiente, 1998.

En España, la creación, en 1983, del **Instituto de la Mujer** (a partir de ahora IM), se considera como el **punto de partida del feminismo de Estado y de las políticas públicas de igualdad de género en España** (Gil, 1996). Este feminismo de Estado español no sólo ha tenido un impacto significativo en la formulación de las políticas públicas de igualdad (Valiente, 1997: 127), sino que ha sido el principal protagonista e impulsor de las mismas. El nivel central es, por lo tanto, el nivel gubernamental de referencia en los inicios de las políticas de igualdad en España, aunque esta característica se va desdibujando con los años el desarrollo del Estado de las autonomías. Como veremos con más detalle más adelante, a partir de finales de la década de los años 80, en España se van desarrollando estructuras gubernamentales similares al IM en las comunidades autónomas de forma generalizada y en el nivel local de forma más irregular. En la actualidad, por lo tanto, parece especialmente interesante incluir también en el análisis del feminismo de Estado español, al menos, el nivel autonómico.

2. LAS POLÍTICAS DE IGUALDAD DE GÉNERO EN ESPAÑA DESDE UNA PERSPECTIVA HISTÓRICA

2.1. Un crecimiento rápido

En 1977, estando la Unión de Centro Democrático en el Gobierno, se crea la Subdirección General de la Condición Femenina en el entonces Ministerio de Cultura y Bienestar Social, suprimiendo la Sección femenina heredada de la época franquista. Al frente de esta nueva subdirección se puso a un grupo de mujeres progresistas no vinculadas a ningún partido político y que realizaron una labor no demasiado fácil, ya que su voluntad reformadora se entendía como un ataque a la familia y a la sociedad por parte de los grupos más conservadores pero no era suficiente para los grupos feministas¹⁵. Su línea de trabajo incluía el apoyo a asociaciones a través de subvenciones, la creación de tres Centros de Información a la Mujer y la revisión de la legislación discriminatoria¹⁶. En 1980 pasaría a denominarse Subdirección General de la Mujer, dependiente de la Dirección General de Juventud y Promoción Sociocultural, del Ministerio de Cultura. Con este cambio no aumenta su rango político-administrativo ni deja de ser una pequeña unidad administrativa con poco presupuesto y dentro de un departamento ministerial no central.

Es decir, aunque se pueden analizar las estructuras anteriores como las primeras estructuras gubernamentales dedicadas a la mujer, e incluso valorar el papel de “transición” que cumplieron, habrá que esperar hasta la creación del IM en 1983 para hablar de un feminismo institucional más significativo¹⁷. Con la creación del IM se aumenta el rango político administrativo de las estructuras gubernamentales dedicadas a la mujer, se dispone de un presupuesto propio y es entonces cuando las políticas públicas de igualdad de género comienzan a formar parte de las políticas de gobierno.

¹⁵ El movimiento feminista mantuvo contradicciones al plantearse las relaciones con la Subdirección General de la Condición Femenina. Por un lado, era el único organismo democrático que defendía los intereses de las mujeres y que podía facilitar apoyo económico a las asociaciones y, por otro se representaba una institución estatal en un gobierno percibido como de derechas (Alberdi, Escario y López-Acotto, 1996: 331).

¹⁶ En aquellos años se consiguieron algunos éxitos importantes, como la inclusión de la prohibición de las discriminaciones por razón de sexo en el trabajo en el estatuto de los Trabajadores, y la reforma del Código Civil en 1981, que legalizaba el divorcio y adaptaba las leyes relativas al matrimonio al principio de igualdad establecido en la Constitución (Granados, 1999: 263-264).

¹⁷ Queda pendiente la realización de un estudio en profundidad sobre el verdadero papel que cumplió, durante la transición democrática, la Subdirección General de la Condición Femenina y que, posteriormente se denominaría Subdirección General de la Mujer en la creación y aparición del feminismo institucional en España.

Este retraso español de aproximadamente dos lustros respecto a otros países del ámbito occidental, que habían desarrollado el feminismo de Estado en la década de los setenta, tiene que ver sobre todo con el tardío desarrollo democrático español debido a la dictadura franquista de cuarenta años de duración.

A finales del año 1983, se crea el IM en el nivel nacional como consecuencia de las presiones de algunos sectores del movimiento feminista, la influencia internacional y el ejemplo de otros Estados que ya contaban con organismos similares y, sobre todo, la subida del PSOE al poder en 1982 y las presiones de las mujeres militantes dentro del propio partido socialista (Valiente, 1996).

Cinco años más tarde se lanza el I Plan de Igualdad de Oportunidades de las Mujeres (1988-1990) en el nivel nacional al que le siguen el II y el III Plan en los periodos 93-95 y 97-2000 respectivamente. A partir del año 1988 se comienza un proceso paralelo en la creación de organismos específicos¹⁸ y la aprobación de planes en las comunidades autónomas (a partir del año 1989). Aunque este proceso ha tenido ritmos distintos en las diferentes Autonomías, actualmente, todas las comunidades autónomas (exceptuando las ciudades autónomas de Ceuta y Melilla) cuentan o han contado con un Plan que estructura las políticas de igualdad de género en cada uno de sus territorios.

Como se puede observar en la tabla 3.1, se puede considerar el año 1988 como el punto de partida del Feminismo institucional o de Estado en el nivel autonómico. Aunque se habían creado poco antes algunas estructuras dedicadas específicamente a la mujer, éstas eran unidades administrativas inferiores y subsumidas dentro de una Dirección General no específica (Cantabria, Agencia de Promoción de la Mujer, 1986) o estructuras intermedias previas a la creación de una estructura más estable (Cataluña, Comisión Interdepartamental para la Promoción de la Mujer, 1987). En 1988 se crean los institutos de Andalucía, Comunidad Valenciana y País Vasco, además de unidades administrativas inferiores en Extremadura y La Rioja, que junto con Cantabria y Cataluña, hacen que ya este primer año sean más de la tercera parte del total de comunidades autónomas (17) las que cuenten con algún tipo de estructura gubernamental dedicada a la mujer. En 1989 se añaden Asturias y Madrid, además de crearse el Institut Català de la Dona en Cataluña. En 1990 se crean estructuras en Baleares y Castilla-La Mancha, lo que hace que casi dos tercios de las comunidades autónomas tengan este tipo de estructuras. Las restantes comunidades autónomas van creando sus estructuras en los primeros años 90 (Murcia, Galicia y Castilla-León), y en los últimos años, se crean sendos institutos en las tres comunidades autónomas que faltaban (Instituto Aragonés de la Mujer, 1993; Instituto Canario de la Mujer, 1994; Instituto Navarro de la Mujer, 1995)¹⁹. Es decir, y como se verá más adelante cuando

¹⁸ El término "organismo específico" se utiliza aquí con un significado similar al de "estructuras gubernamentales dedicadas a la mujer" (*women's policy machinery*). También se utiliza, a lo largo del texto, la expresión "organismos de igualdad" con el mismo significado.

¹⁹ En 1999 se crea el Instituto Asturiano de la Mujer, que no tiene rango de organismo autónomo ni está creado por ley, sino que es un órgano desconcentrado dependiente de la Consejería de la Presidencia, establecido por decreto (véanse Decretos 61/99 de 10 de Agosto y 137/99 de 16 de Septiembre, publicados en el B.O.P.A. del 12 de Agosto y 18 de Septiembre, respectivamente).

analicemos el marco institucional, aunque estas estructuras son muy diferentes y su poder y grado de institucionalización y consolidación difiere mucho entre sí, vemos que en un espacio de no más de siete años se generalizan las estructuras gubernamentales dedicadas a la mujer en el nivel autonómico.

TABLA 3.1
Estructuras gubernamentales para la igualdad de género en las
comunidades autónomas

Comunidad Autónoma	Organismo de igualdad y año de creación	Planes de Igualdad
ANDALUCIA	Instituto Andaluz de la Mujer (88)	I Plan 1990-92 II Plan 1995-97
ARAGON	Instituto Aragonés de la Mujer (93)	I Plan 1994-96 II Plan 97-2000
ASTURIAS	Secretaría de la Mujer (89) Dirección Regional de la Mujer (93) Secretaría Regional de la Mujer (95) Instituto Asturiano de la Mujer (99)	I Plan 1989-91 II Plan 1993-95 III Plan 1996-2000
BALEARES	Comisión Interdepartamental de la Mujer (90) Instituto Balear de la Mujer (2000)	I Plan 1991-93 II Plan 1996-99
CANARIAS	Instituto Canario de la Mujer (94)	I Plan 1995-96 II Plan 97-2000
CANTABRIA	Agencia de Promoción de la Mujer (86)* Dirección General Mujer (97)	I Plan 1991-93 II Plan 1998-2001
CASTILLA-LA MANCHA	Dirección General de la Mujer (90)	I Plan 1990-94 II Plan 1995-99 III Plan 1999-2003
CASTILLA-LEON	Servicio Regional de la Mujer* Secretaría Regional de la Mujer (94) Dirección General de la Mujer e Igualdad de Oportunidades	I Plan 1994-96 II Plan 1997-2000
CATALUÑA	Comisión Interdepartamental de Promoción de la Mujer (87) <i>Institut Catalá de la Dona</i> (89)	I Plan 1989-92 II Plan 1994-96 III Plan 1998-2000
EXTREMADURA	Asesoría Ejecutiva de la Mujer (88)* Dirección General de la Mujer (91)	I Plan 1991
GALICIA	<i>Servicio Galego de Promoción da Igualdade del Home e la Muller</i> (91)	I Plan 1992-94 II Plan 1995-97 III Plan 1998-2001
MADRID	Dirección General de la Mujer (89)	I Plan 1989-91 II Plan 1993-95 III Plan 1997-2000
MURCIA	Dirección General de la Mujer (91) Dirección General de Política Social y Familia (95) Dirección General de Juventud, Mujer y Familia (97) Secretaría Sectorial de Mujer y de Juventud	I Plan 1993-95 II Plan 1997-2000
NAVARRA	Subdirección General de la Mujer (91) Instituto Navarro de la Mujer (95)	I Plan 1998-2000
PAIS VASCO	Instituto Vasco de la Mujer (<i>Emakunde</i>) (88)	I Plan 91-94 II Plan 95-98 III Plan 1999-2002
LA RIOJA	Centro Asesor de la Mujer* (Dirección General de Bienestar Social -88)	I Plan 91-95 II Plan 96-99
COMUNIDAD VALENCIANA	<i>Institut Valenciá Dona</i> (88) Dirección General Mujer (97)	I Plan 89-91 II Plan 1997-2000

* Unidades administrativas inferiores al rango de Dirección General y dependientes de una Dirección General no específica de temas de mujer.

Desde el momento en el que empiezan a surgir las estructuras gubernamentales dedicadas a la mujer en el nivel autonómico, comienzan también a desarrollarse las políticas autonómicas de igualdad de género. Es decir, con la misma rapidez que surgen dichas estructuras, se van consolidando también unas políticas de igualdad de género, articuladas, como veremos más adelante, a través de planes de igualdad. Como señala Celia Valiente (1995, 1996), aunque el feminismo de Estado aparece en España aproximadamente diez años más tarde que en otros países occidentales, en el año 1994, el IM español es comparable en términos de presupuesto, personal y objetivos a las instituciones de igualdad de dichos países. Si además tenemos en cuenta el desarrollo tanto de las políticas como de los organismos de igualdad en el nivel autonómico, se puede concluir que España es uno de los países europeos con un crecimiento mayor y más rápido en este terreno en los últimos diez años.

Este desarrollo debe ser entendido, además, teniendo en cuenta dos factores del contexto socio-político español que son importantes para el análisis del desarrollo y el papel de las políticas públicas de igualdad de género en España en general, y en el nivel autonómico, en particular. En primer lugar, este país pasa de la dictadura franquista a un sistema democrático integrado en Europa en un período aproximado de 10 años. A esta situación política hay que añadirle los profundos cambios sociales, incluyendo el papel de las mujeres, que ha experimentado la sociedad española. En segundo lugar, y de forma paralela, hay que tener en cuenta el desarrollo del Estado de las autonomías. En menos de 20 años hemos pasado de un país fuertemente centralizado a uno de los más descentralizados de Europa, desarrollando un modelo abierto, todavía en desarrollo, con claros caracteres federales, aunque sólo podría ser adjetivado como "cuasi-federal o federalizante" dada la atrofia del Senado como verdadera cámara territorial y cuya reforma se está estudiando (Moreno, 1997:141).

En este contexto, las políticas de igualdad de género en el nivel autonómico han ido adquiriendo una importancia creciente, de tal manera que resulta imprescindible su análisis para obtener una idea general de estas políticas en el Estado español. Esta generalización y descentralización de las políticas de igualdad de género también ha tenido sus efectos en el nivel local aunque, como se ha comentado anteriormente, este desarrollo ha sido mucho más irregular. El estudio sistemático de las políticas de igualdad de género en el nivel local (incluyendo políticas de promoción de la mujer, que es lo que, en muchos casos, domina en este ámbito), que excede este trabajo de investigación, permanece como un campo de estudio todavía sin explotar y que, creemos, adquirirá deseablemente un gran desarrollo en los próximos años.

2.2. La evolución en la definición del problema

El tipo de respuestas que se den y el tipo de acciones públicas que se lleven a cabo al amparo de una política pública dependen, en gran medida, de cómo se defina el problema público que se pretende resolver. Es decir, en este caso es importante saber cómo se define el concepto de (des)igualdad de género. Más allá, resulta interesante analizar si ha habido algún cambio en la definición del problema desde que éste se hiciera público y ver cuál ha sido su evolución desde el inicio de las políticas públicas de igualdad de género en España.

A principios de la década de los años 80, las reivindicaciones de las socialistas feministas consistían, sobre todo, en desarrollar el derecho de igualdad establecido en la Constitución española de 1978, impulsando legislación no discriminatoria y desarrollando acciones que promovieran la igualdad entre los sexos. De hecho, el objetivo del IM, expresado en su ley de creación es “el cumplimiento y desarrollo de los principios constitucionales, la promoción y fomento de las condiciones que hacen posible la igualdad social entre los sexos y la participación de la mujer en la vida política, cultural, económica y social”²⁰.

Es decir, por un lado, se reivindicaba en primer lugar, la igualdad legal o de derecho. Aunque la Constitución de diciembre de 1978 consagraba el principio de igualdad en su artículo 14, todavía quedaban, en el momento de creación del IM, algunos restos de una legislación claramente desigualitaria y machista. Una buena parte de las medidas del I Plan de Igualdad en el nivel central tenían que ver con los cambios legislativos necesarios como condición previa para comenzar a abordar los problemas de desigualdad de las mujeres. Por lo tanto, muchas de aquellas primeras energías se emplearon en conseguir la **igualdad de derecho**. Ya por entonces, se era consciente de que esta igualdad de derecho, aunque necesaria, no conllevaba la **igualdad de hecho**. De esta manera, se trataba de promover las condiciones necesarias para que las mujeres pudieran tener las mismas oportunidades que los hombres en toda una serie de ámbitos (haciendo especial hincapié en el laboral). Sin embargo, la verdadera igualdad de oportunidades es difícil de conseguir cuando la situación de partida de ambos sexos ha sido muy distinta y cuando las causas profundas de la desigualdad se hallan en las propias estructuras sociales que apoyan y perpetúan dicha desigualdad. Esta es la razón por la que se ha comenzado a formular la idea de la igualdad de hecho en torno a los **resultados** y no sólo en torno a las **oportunidades**. Es decir, no sólo se trata de poner el punto de mira en que todas las personas, independientemente de su sexo, tengan las mismas oportunidades, sino de empeñarse en conseguir que en muy diferentes ámbitos, y que se reflejan en datos y estadísticas (salarios, tasa de desempleo, tipo de estudios, niveles en la escala laboral, usos del tiempo, etc.) no sea la variable sexo la que explique las mayores diferencias.

²⁰ (Ley de creación del Organismo Autónomo Instituto de la Mujer. Ley 16/1983, de 24 de Octubre).

Por otro lado, el problema se definía, en los comienzos de feminismo de Estado en España, en términos de **desigualdad entre los sexos**, refiriéndose especialmente a la falta de participación de las mujeres en la sociedad. Sensat y Varella (1998) analizan el primer y segundo planes en el nivel central y concluyen que las políticas de igualdad de género en España han tratado de luchas contra la desigualdad de sexo y no de género. Entendemos, como señala Lovenduski (1997), que **sexo** se refiere al sexo biológico y **género** al modo en que las relaciones entre sexos se producen e institucionalizan, esto es, el género se refiere a las consecuencias sociales de las diferencias biológicas entre hombres y mujeres²¹. Siguiendo a aquellas autoras, que definen el género como “una categoría relacional que tiene que ver con las relaciones entre hombres y mujeres que se establecen asimétricamente”, aunque la categoría de género se reconoce en el II Plan central (93-95), esto se ha hecho sólo al nivel del discurso. Estas mismas autoras también señalan que los dos primeros planes centrales se han articulado alrededor de la arena productiva (y no la reproductiva). La arena productiva tiene que ver con las actividades realizadas dentro del mercado o esfera pública y que se hacen a cambio de un salario; estas actividades están reconocidas social y económicamente. Por el contrario, la arena reproductiva tiene que ver con la esfera privada y con las actividades de cuidado dentro del hogar y la familia; estas actividades son las que mantienen, cotidianamente, la fuerza de trabajo dentro de la unidad doméstica. Pero no están reconocidas ni social ni económicamente; son invisibles. Es decir, por un lado, aunque hay una tendencia hacia el reconocimiento de la categoría de género - superando la categoría sexo-, ésta no llega a ser asumida plenamente. Por otro lado, se hace hincapié en la igualdad de la mujer hacia el hombre (en el ámbito productivo) y no tanto en la igualdad del hombre hacia la mujer (ámbito reproductivo).

También basándose en el dilema público-privado, Gil (1996) señala diferencias entre el primero y el segundo plan central. El I Plan (1988-1990) estaba orientado casi exclusivamente hacia el ámbito del trabajo, no mencionándose el dualismo público-privado o la necesidad de promocionar un reparto más igualitario de las responsabilidades domésticas. El II Plan (1993-1995), sin embargo, mencionaba la necesidad de nuevos balances entre las esferas pública y privada. Este II Plan menciona ya algunas acciones relacionadas con la esfera “privada”. Estas diferencias son un indicador de cómo ha ido evolucionando la definición del problema de la igualdad. A lo largo del tiempo, las mujeres se han dado cuenta de que no es suficiente conseguir la igualdad formal; ni siquiera es suficiente tener mayor acceso al mercado laboral porque las estructuras sociales, económicas y culturales no permiten que este acceso se haga en verdadera igualdad de condiciones. Las mujeres han conseguido tener un cierto acceso al mundo productivo (principalmente masculino) pero todavía tienen, de forma casi exclusiva y sin conseguir

²¹ El concepto de género, además, pretende romper en cierto modo las categorías monolíticas de mujeres y hombres, desplazando el foco de “mujer” hacia el de “género”, explicando la desigualdad no tanto por las diferencias genéricas entre hombres y mujeres, sino por los conflictos en las relaciones de género. Para Lovenduski, los “conflictos de género (...) se definen como conflictos contruidos en función de los roles opuestos de hombres y mujeres. Al contrario de conceptos como ‘temas de mujeres’, ‘derechos de mujeres’ o ‘necesidades de mujeres’, el conflicto de género no presupone intereses comunes de todas las mujeres y todos los hombres. El conflicto de género no requiere una identidad social amplia, sino que requiere que la disputa se construya en función de diferencias entre mujeres y hombres concretos” (Lovenduski, 1997: 207).

repartirlo²², la responsabilidad del mundo reproductivo (principalmente femenino). De esta manera, aunque todavía es importante conseguir un acceso más igualitario de las mujeres a la esfera productiva, va adquiriendo importancia creciente la combinación de estrategias dirigidas al ámbito reproductivo.

Una de las consecuencias de esta evolución en la definición del problema es la creciente, aunque todavía débil, (re)consideración de algunos aspectos tradicionalmente considerados *privados* (relacionados con el ámbito reproductivo y doméstico) como asuntos públicos (algo que se considera para la acción pública). Esto también ha llevado a nuevas definiciones del problema de igualdad: del concepto de “**igualdad a**” al de “**igualdad entre**”; de la idea de la participación de la mujer en un mundo de hombres a la de cambiar el mundo para que en él quepan hombres y mujeres. Un indicador simbólico de esta tendencia son los cambios en los títulos de los sucesivos planes centrales. Los planes primero y segundo se titulaban “Planes de la Igualdad de Oportunidades *de las Mujeres*”. El III Plan (1997-2000) es denominado “Plan para la Igualdad de Oportunidades *entre Mujeres y Hombres*”. Además, en la introducción de este III Plan central se dice que “la promoción de la igualdad no requiere únicamente la adopción de medidas positivas dirigidas las mujeres (...). También requiere medidas destinadas a adaptar la organización de la sociedad hacia una distribución más justa de roles”²³. Sin embargo, creemos, de acuerdo con Sensat y Varella, que esta nueva definición se reconoce en el nivel del discurso pero no se “traduce” a la práctica de las políticas públicas de igualdad de género²⁴.

Otra tendencia evolutiva en la definición del concepto de igualdad es su reivindicación no sólo en términos de derechos, sino también en términos de sus **efectos positivos**, es decir, la idea de que la igualdad no es sólo buena para las mujeres sino para las hombres y la sociedad en general también. Este es el caso, por ejemplo, del lema de la

²² Los últimos datos sobre “usos del tiempo” son que en 1996 las mujeres españolas emplean una media de 7h35’ en actividades domésticas, mientras los hombres lo hacen durante 3h05’. Estos datos son algo mejores que los de 1993 (7h.58’ frente a 2h30’), y esto significa que las cosas van cambiando. Pero las diferencias siguen siendo enormes (Instituto de la mujer. *Mujeres en cifras*).

²³ III Plan para la Igualdad de Oportunidades entre Hombres y Mujeres 1997-2000. Pág. 10.

²⁴ Tanto Gil (1996) como Sensat y Varella(1998) basan sus análisis en el I y II Plan exclusivamente. En nuestro breve análisis del III Plan central, encontramos que, en estos temas, parece haber incluso algún retroceso. Mientras en el II Plan había un área específica sobre “Promoción de una distribución más equitativa de las responsabilidades domésticas”, en el III sólo hemos encontrado un objetivo dentro del área de “Economía y Empleo” que tiene que ver con la “armonización de la vida familiar y laboral de las mujeres y los hombres”. Es significativo no sólo que ese objetivo sea formulado desde el área de economía y empleo (área productiva), sino que es menos preciso y ambicioso que el del II Plan. Así, por ejemplo, podemos comparar la acción 3.5.4. del III Plan con la 5.1.5. del II. La primera es “sensibilizar a la opinión pública sobre la importancia de difundir una imagen de las mujeres y de los hombres que afiance y consolide una distribución de las tareas no sexistas”. La segunda es “realizar campañas, a través de publicidad y material divulgativo, sobre la práctica cotidiana del reparto de papeles y responsabilidades, a fin de suscitar una toma de conciencia en dos vertientes: una, sobre la igualdad de oportunidades, y otra, sobre los efectos positivos que para las niñas y niños supone el ser atendidos tanto por parte de la madre como del padre”. (Curiosamente, aunque esta última acción se fundamenta con los efectos positivos que la misma puede tener no sólo sobre la mujer madre, sino también sobre los hijos -y esto es un paso importante-, no se menciona en ningún momento los posibles efectos positivos posibles para el hombre en su labor de padre).

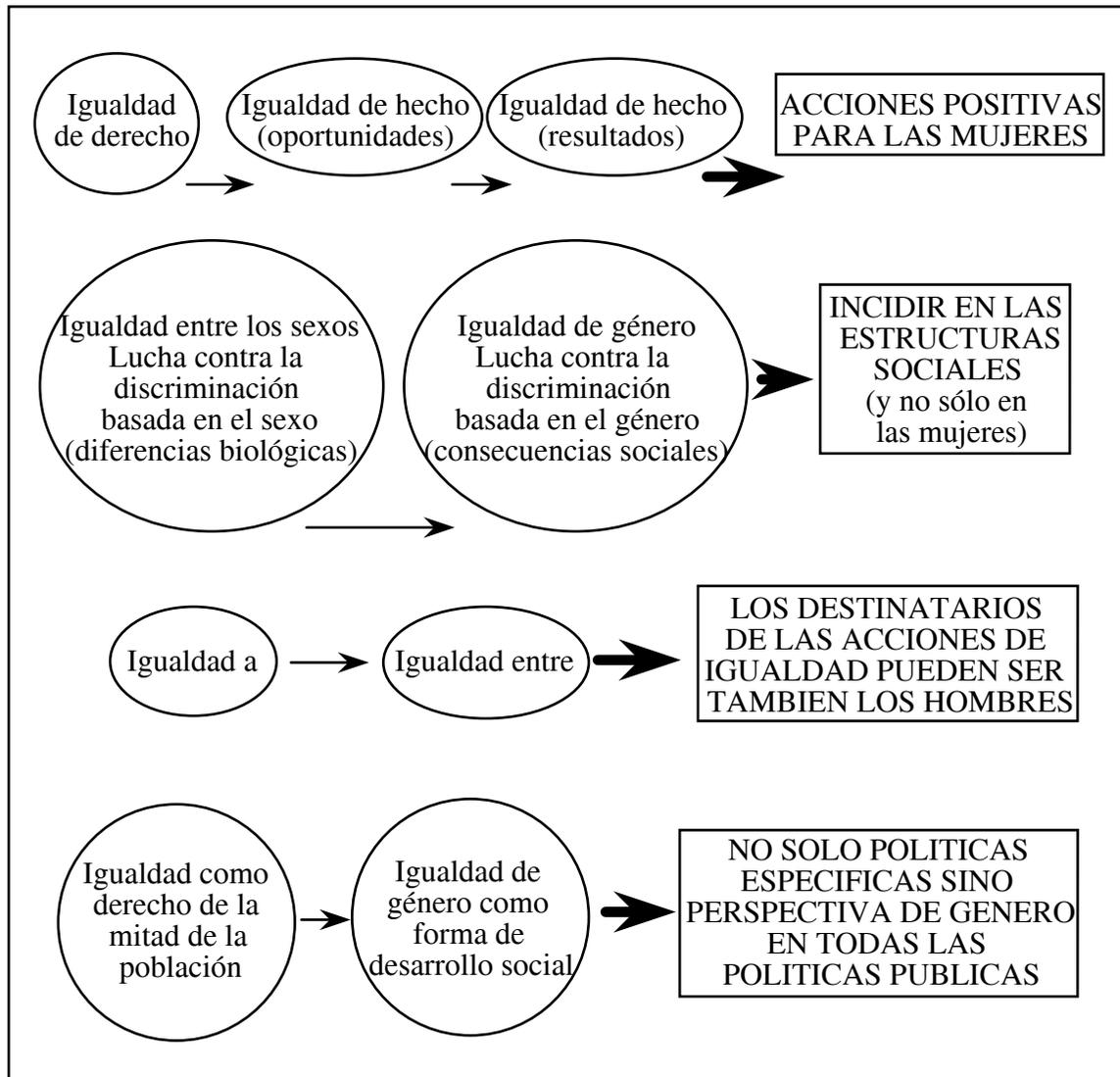
campaña de 1998 de la Comunidad de Madrid (“*La igualdad nos interesa a todos y no sólo a las mujeres. Aplica la igualdad en tu vida cotidiana*”). Otro ejemplo es el del programa “OPTIMA” (lanzado por el IM después del II Plan, a través del IV Programa de Acción de la UE) y que trata de estimular a las empresas para que establezcan compromisos de igualdad en sus organizaciones²⁵. Esta tendencia podría llevar a utilizar también estrategias de “marketing” o “venta”, además de las tradicionales de reivindicación que, por supuesto, siguen estando en la esencia de cualquier acción pública dirigida a conseguir la igualdad. Puede ser bien para “vender” la imagen de la mujer para el mundo productivo (por ejemplo, “las mujeres pueden ser muy buenas y diferentes directivas en una organización”) o bien para “vender” las actividades reproductivas a los hombres (por ejemplo, “es importante una presencia fuerte del padre en los primeros años de crianza” o “la práctica de actividades domésticas y de cuidado de los demás son importantes para el desarrollo y maduración afectiva y emocional del ser humano”).

En el *Institut Catalá de la Dona*, para realizar la evaluación de II Plan de Igualdad de Oportunidades (1994-1996), establecen una categorización del tipo de acciones contenidas en ese plan que reflejan, en cierto modo, la tendencia evolutiva en la definición del problema de la igualdad. Hablan de acciones de "**promoción de la mujer**" (tienen por objetivo la prestación de servicios a las mujeres), de "**igualdad**" (tienen por objetivo establecer los condicionantes que mujeres y hombres tengan las mismas oportunidades) y de "**paridad**" (encaminadas a potenciar la modificación de las pautas de relación social, de forma que mujeres y hombres compartan responsabilidades tanto en el ámbito público como en el privado) encontrando que se habían ejecutado muchas más de la primera categoría, menos de la segunda y muy pocas de la tercera. Una de sus conclusiones es que hay que ir equilibrando estas tres categorías de acciones, es decir, hay que tender hacia la paridad.

En la tabla 3.2. se expresan gráficamente las tendencias en la evolución del problema de la igualdad y algunas de sus consecuencias sobre su abordaje en las políticas públicas de igualdad de género.

²⁵En el II Plan, la acción 3.8.3. mencionaba, entre otras medidas, campañas con el lema “la igualdad de oportunidades es una buena oportunidad para las empresas”.

TABLA 3.2. Evolución del problema de la igualdad



Todas las tendencias evolutivas expresadas en la tabla 3.2. no representan conceptos mutuamente excluyentes; es decir, la evolución en la definición del problema no significa que se abandone la antigua formulación, sino que se enriquece con la nueva. Así, la igualdad de derecho como reivindicación de la mitad de la población sigue siendo básica en la definición del problema de la igualdad.

Una de las consecuencias de poner en el punto de mira de las acciones públicas no sólo en la igualdad de oportunidades, tanto legal como "de facto", sino en la igualdad de resultados implica reafirmar la necesidad de **acciones positivas** para las mujeres. Es decir, para que la igualdad de oportunidades sea real hay que "colocar" a las mujeres en el mismo punto de partida removiendo los obstáculos para ello y estableciendo medidas especiales

que animen a superarlos. Y si lo que se pretende es la igualdad de resultados, hay que intensificar dichas acciones positivas, incluyendo incluso, medidas de discriminación positiva.²⁶ Las tendencias hacia el concepto de “género” y el de “igualdad entre” han hecho replantear el que haya que cargar todas las tintas en las mujeres como únicas destinatarias de las políticas de igualdad; hay que plantearse también el cambio de las estructuras sociales en general e incluso la idea de que los **hombres también son destinatarios** directos de las políticas de igualdad de género. El foco se traslada a las diferencias basadas en el género, en ambas direcciones, aunque como es evidente las discriminaciones de género afectan en mayor medida a las mujeres que a los hombres. Por último, contemplar la igualdad no sólo como derecho sino **como algo deseable para todos** los seres humanos y en todos los aspectos, hace que, más que acciones públicas sectoriales para la igualdad, se planteen acciones públicas conducentes a impregnar de "**perspectiva de género**" todas las políticas públicas, es decir, que todas las actuaciones y políticas públicas se formulen, ejecuten y evalúen teniendo en cuenta a las mujeres y las consecuencias diferenciales que sobre ellas puedan tener dichas actuaciones. Para esto, lógicamente, habrá que haber analizado la realidad también con dicha perspectiva, “visibilizando” a las mujeres en dicho análisis.

Aunque parece que la definición de igualdad apunta en las direcciones arriba indicadas, como ya se ha mencionado antes, estas tendencias están, por lo general, sólo reconocidas en el nivel discursivo. En general, el feminismo institucional español ha tenido mucho éxito en la formulación del discurso subyacente a las políticas públicas de igualdad pero no tanto en la traducción práctica de dicho discurso en actuaciones verdaderamente coherentes con el mismo. Por un lado, este discurso todavía contiene algunas contradicciones (hablamos de género, pero lo hacemos desde institutos de la “mujer”; hablamos de igualdad y nos vamos dando cuenta de que para conseguirla debemos establecer un trato diferencial y no obviar una supuesta igualdad); por otro, el discurso no se traduce en la práctica y ejecución de las políticas. Algunos ejemplos e indicadores de esto podrían ser:

1. En las evaluaciones hablan exclusivamente de las “destinatarias” de los planes de igualdad. Aunque incluso los títulos de los planes incluyan a la otra mitad de la población -los hombres-, en los diseños de evaluación, cuando éstos incluyen recopilar información de la población destinataria, el universo sobre el que se trabaja es el de las mujeres.
2. Cuando se propone algún tipo de medida relacionada con el reparto de las responsabilidades del ámbito reproductivo, se hace desde el marco de la “conciliación de la vida laboral y doméstica”, es decir se insiste desde la perspectiva de que el

²⁶ Nosotros no estamos de acuerdo en que la frontera entre una acción y una discriminación positiva esté claramente delimitada; creemos, además, que la utilización del término “discriminación” puede acentuar el carácter peyorativo de este tipo de medidas. Lo que está claro es que, aunque estas medidas han sido muy criticadas y tienen una altísima probabilidad de ser polémicas, sí se reconocen como efectivas (por ejemplo, el sistema de cuotas en los partidos políticos). Para un estudio en profundidad la fundamentación jurídica de dichas medidas y de las críticas posibles a las mismas, véanse Ruiz Miguel (1996) y Pérez del Río (1996).

hombre debe asumir responsabilidades para que la mujer pueda trabajar y no soportar una doble carga (es decir, en función del ámbito productivo y por una cuestión de justicia social, para que las mujeres tengan la oportunidad de ser más “productivas”), pero se insiste menos en una justificación del reparto del ámbito reproductivo en sí mismo, reconociéndolo, valorizándolo e intentándolo repartir con los hombres, porque también puede ser bueno para ellos y para la sociedad en general.²⁷

3. La “perspectiva de género” se asume como discurso, pero en muchos casos se vacía de contenido y se traduce en una simple contabilización de cuántas mujeres han sido beneficiadas por cada actuación, proyecto o política.

Como se puede observar, a pesar de todo lo avanzado, existen todavía bastantes contradicciones, que resultan interesantes para el análisis. Al tratarse de políticas públicas de un alto carácter simbólico, es relativamente fácil cubrirlas de un discurso en el nivel formal que “cubra el expediente”, a pesar de que las políticas simbólicas tengan su importancia (Valiente, 1999: 485-490). Precisamente, el análisis de las acciones y planes públicos de igualdad desde esta perspectiva nos parece especialmente interesante, incluida la fase de la evaluación. ¿Cuál es el concepto de igualdad al que verdaderamente están respondiendo los planes públicos? ¿Existen contradicciones entre el discurso y las acciones y la perspectiva desde la que se realizan?

2.3. La entrada en la agenda política

El proceso de formación o construcción de la agenda política es una fase determinante en el análisis de políticas públicas y ha recibido especial atención por parte de los politólogos, ya que intenta responder a interrogantes clave en el estudio de las políticas públicas. Dentro de esta fase se analizan las razones de por qué un determinado tema adquiere relevancia pública y otros no, qué es lo que hace que los poderes públicos se encarguen de unos temas y no de otros, quiénes influyen en esa entrada en la agenda. Posteriormente habrá que estudiar si es suficiente la entrada en agenda para que los poderes públicos actúen sobre el tema en cuestión. La igualdad de género resulta especialmente paradigmática en el análisis de la construcción de agenda ya que, entre otras cosas, los poderes públicos han demostrado interés por él de forma reciente porque la mayoría de las

²⁷ Aquí nos encontramos también con la dualidad público-privado. Lo que se ha definido como perteneciente al ámbito privado es un terreno en el que los poderes públicos no deben entrar. En el corazón de la cuestión está, pues, qué es lo que se considera de ámbito público y qué se considera del ámbito privado. La historia de las políticas públicas de igualdad es, en cierto modo, cómo los movimientos de mujeres han ido consiguiendo, a lo largo de la historia, que cuestiones consideradas tradicionalmente como privadas hayan sido consideradas cuestiones públicas y hayan entrado en la agenda de actuación de los poderes públicos. Uno de los ejemplos más recientes ha sido el caso de la violencia doméstica y el maltrato a las mujeres.

cuestiones aparecían vinculadas tradicionalmente al ámbito de lo privado (Martínez, 1998: 149).

Se ha argumentado que la entrada en la agenda política es un paso particularmente difícil para las políticas públicas de igualdad de género (Martínez, 1997; Threlffal, 1990; Ousthoorn, 1991). Efectivamente, estas políticas se deben definir como un problema público y la cuestión de la igualdad debe ser suficientemente reconocida para que llegue a discutirse. Sin embargo, y como señala Celia Valiente (1997), las femócratas²⁸ españolas han tenido un cierto éxito en la definición de problemas y en la introducción de los mismos en la agenda política. La introducción de los temas de igualdad de género en la agenda política está, sin duda, relacionada con la utilización de planes transversales, es decir que, al menos en la teoría, acata todo el Gobierno (en su mayoría, denominados planes de igualdad de oportunidades). Estos planes son actualmente el instrumento principal para la articulación de las políticas públicas de igualdad de género, tanto en el nivel central como en el autonómico²⁹. Así, las femócratas españolas han conseguido introducir la igualdad de género en muchas de las agendas políticas de diferentes sectores gubernamentales (aunque en algunos con más éxito que en otros). Sin embargo, y como veremos más adelante, las femócratas españolas no han tenido tanto éxito en el control de la formulación y ejecución de dichas políticas. Entre otras cosas, la mera inclusión de temas en las agendas de los poderes públicos no garantiza necesariamente la implementación de políticas públicas al respecto (Martínez, 1998: 149).

La entrada en la agenda política de la igualdad de género ya ha sido analizada por Celia Valiente (1995) en el nivel central. Entre otros factores, Valiente destaca el **compromiso del Partido Socialista Obrero Español** (PSOE) impulsado, a su vez, por las mujeres feministas dentro de dicho partido. En 1981, se crea "Mujer y Socialismo", dentro del PSOE, y esto representa un factor fundamental para que el Gobierno Socialista de 1982 introduzca el tema de la desigualdad de género en su agenda política. De hecho, en el programa electoral de 1982 de dicho partido se hablaba de crear una "comisión para la igualdad que garantice la no discriminación entre los dos sexos" (PSOE 1982: 29). Esto hace que, una vez ganadas las elecciones generales de Octubre de 1982 con mayoría absoluta, y con la presión de "Mujer y Socialismo" a sus compañeros de partido, entrara en la agenda de los poderes públicos la creación del IM, que se realiza formalmente a finales de 1983.

El principio de igualdad ya había quedado consagrado en la **Constitución** de 1978, como valor fundamental (a través del precepto 1.1), como principio jurídico (a través de su artículo 14) y como principio político (a través del precepto 9.2). Las primeras femócratas españolas insisten en la necesidad de adecuar la legislación a dicho principio constitucional

²⁸ Por femócratas entendemos las personas responsables de las instituciones públicas que tienen por objetivo la promoción e igualdad de la mujer. Este término, junto con el de *feminista de Estado* se ha utilizado también para denominar a cualquier mujer feminista en la administración Pública (Siim, 1991; Sawyer, 1990)

²⁹ Aunque en el nivel local también se utiliza en ocasiones este mismo instrumento (sobre todo en las ciudades grandes y medianas), su uso no está tan extendido y generalizado como en los otros dos niveles gubernamentales.

y se apoyan especialmente en el precepto 9.2, que establece que “corresponde a los poderes públicos promover las condiciones para que la libertad y la igualdad del individuo y de los grupos en que se integra sean reales y efectivas; remover los obstáculos que impidan o dificulten su plenitud y facilitar la participación de todos los ciudadanos en la vida política, económica, cultural y social”, para fundamentar la necesidad de su intervención para la aplicación real de dicho principio de igualdad. Estos fundamentos jurídicos deben también estudiarse en los **Estatutos de autonomía**, normas fundamentales de las comunidades autónomas. En todos estos estatutos se han incluido preceptos similares al 9.2 de la Constitución, algunos de forma más específica que la propia Constitución y los otros Estatutos. Este es el caso del precepto 12.2 del Estatuto de Andalucía que establece que “la Comunidad Autónoma propiciará la efectiva igualdad del hombre y de la mujer andaluces, promoviendo la plena incorporación de ésta en la vida social y superando cualquier discriminación laboral, cultural, económica y política”³⁰. En cualquier caso, la Constitución Española (CE) de 1978 efectúa una descentralización territorial del poder, y como ya se ha visto anteriormente, las materias de promoción de la igualdad de género, así como casi toda la política social van a pasar a depender desde un principio de las comunidades autónomas, por lo que los fundamentos jurídicos de las políticas de igualdad de género también deben contemplarse desde la perspectiva autonómica.

Valiente también menciona como factor determinante en la entrada en agenda, tanto del tema de la igualdad en general como de la creación de organismos específicos para su consecución, la **influencia de los organismos internacionales**. El año 1975 fue declarado Año Internacional de la Mujer por la Organización de las Naciones Unidas y fue seguido del Decenio de las Naciones Unidas de la Mujer que comienza en 1976 y termina en 1985. Las Conferencias Mundiales sobre la Mujer (México, 1975; Copenhague, 1980; Nairobi, 1985 y Pekín, 1995) han ido marcando hitos y arrancando compromisos a los diferentes Estados, reflejados en el Plan de Acción Mundial que se aprobó en México (1975), las Estrategias orientadas hacia el futuro para el adelanto de la Mujer en Nairobi (1985) y la Plataforma de Acción de Pekín (1995). Las propias estructuras para la igualdad de la Organización de Naciones Unidas se desarrollaron enormemente durante aquellos años: para el final del Decenio, se habían establecido y consolidado tres de las principales instituciones: La División para el Adelanto de la Mujer (DAW) que se centra en la política, el Instituto Internacional de Investigaciones y Capacitación para la Promoción de la Mujer (INSTRAW), que se centra en la investigación y, finalmente, el Fondo de Desarrollo de las Naciones Unidas para la Mujer (UNIFEM), que se centra en las actividades operativas. Paralelamente, en el año 1979, la Asamblea General de la ONU aprobó la Convención sobre la Eliminación de todas las Formas de Discriminación Contra la Mujer que comprende derechos de la mujer aceptados internacionalmente y que son aplicables a todas las mujeres y en todos los campos. Esta Convención, cuya norma básica es la prohibición de todas las formas de discriminación contra la mujer y que establece medidas que deben tomarse para asegurar que las mujeres disfruten de los derechos que les corresponden, está abierta a todos los Estados y se ratifica voluntariamente por parte de los mismos. España lo

³⁰ Para un excelente estudio de las bases jurídicas del principio de igualdad de género, tanto en la Constitución Española como en los diferentes Estatutos de Autonomía, consultar Granados (1999: 233- 253).

ratifica muy poco después de la creación del IM, en diciembre de 1983 y en 1995 eran más de 150 los países que la habían ratificado. Basándose en el artículo 17 de dicha Convención se crea el Comité para la Eliminación de la Discriminación de la Mujer, que es el que se encarga de velar por la aplicación de la Convención en los Estados Participantes.

Esta atención a los derechos de la mujer, inicialmente impulsada por la Comisión de la Condición Social y Jurídica de la Mujer de Naciones Unidas (CSW)³¹, tuvo repercusión tanto en las agencias internacionales especializadas como la Organización Mundial de la Salud (OMS), la Organización Internacional del Trabajo (OIT), la UNESCO, etc. como en las propias Comunidades Europeas y el Consejo de Europa³².

El artículo 119 del Tratado de Roma (1957), constitutivo de la Comunidad Económica Europea, reconoce el principio de igualdad de entre mujeres y hombres, referido exclusivamente a la igualdad de remuneración entre los trabajadores masculinos y femeninos que realicen un trabajo igual. Sin embargo, la obligación formal de aplicar su contenido no se da hasta 1975, año en el que se aprueba una directiva (nº 75/117) para armonizar las legislaciones de los Estados miembros en esa materia. Además de la base jurídica de las políticas de igualdad para las mujeres en la Unión Europea (anteriormente CEE), hay que mencionar la aprobación, en 1972, del Primer Programa de Acción Social Comunitario, en el que se incluían medidas para garantizar la igualdad en el empleo, la formación profesional, mencionando la conciliación de las responsabilidades familiares con las profesionales. En 1982 se aprueba el Primer Programa Comunitario para la Igualdad de oportunidades (1982-1985), con el fin de contribuir a crear las condiciones necesarias para que la aplicación de la legislación comunitaria resultara más efectiva. A este primer programa le siguen consecutivamente el II (1986-1990), el III (1991-1995) y el IV Programa de Acción Comunitario para la Igualdad de Oportunidades (1996-2000).

Como se puede observar, el reconocimiento internacional de los derechos de la mujer a lo largo de la década de los años 70 y sus consiguientes convenios, estructuras y medidas son ligeramente anteriores al comienzo del período democrático en España (las primeras elecciones democráticas son en 1977). De hecho, el Decenio de la Mujer coincide con el período de transición española, aprobación de la Constitución y segunda legislatura democrática postconstitucional en la que gana el PSOE y acaba poco antes de la entrada de España en la Comunidad Europea (1986). Más que una influencia directa de toda esta maquinaria internacional en la entrada en la agenda de los poderes públicos españoles, constituye una fuente primordial de legitimación y fundamentación jurídica y teórica para la acción. En realidad, la Convención para la Eliminación de la Discriminación de la Mujer

³¹ Esta Comisión, creada el 21 de Junio de 1946 es el organismo intergubernamental que propone a la Asamblea General, a través del Consejo Económico y Social (ECOSOC), las acciones que se deben desarrollar en favor de las mujeres.

³² El Consejo de Europa, fundado en 1949, también contribuye, aunque en menor medida, al reconocimiento internacional de los derechos de las mujeres. Hay que mencionar el Convenio Europeo para la Protección de los Derechos Humanos, que entra en vigor en 1953 y la Carta Social Europea, de 1961 que, aunque no se refiere a la discriminación de forma específica, varios artículos tienen relación con los derechos de las mujeres

se ratifica una vez ya está constituido el IM y la entrada de España en la Comunidad Europea no llega hasta 1986, aunque las negociaciones habían comenzado mucho antes y ya existía una clara tendencia "europeista" en los poderes públicos. Efectivamente, "la existencia de ejemplos internacionales sirvió a las feministas socialistas no sólo como fuente de inspiración (...), sino también como poderoso argumento para ser esgrimido ante políticos provenientes de todos los partidos, incluido el PSOE, reticentes a la creación de organismos de igualdad, pero deseosos de que España se integrara en el grupo de los países democráticos y desarrollados económicamente" (Valiente, 1994:9).

Relacionado con esto, además de la influencia de los organismos internacionales y supranacionales, hay que mencionar también el ejemplo de otros Estados con organismos similares, especialmente el Ministerio de los Derechos de la Mujer francés que fue establecido en el año 1981.

En el **ámbito autonómico**, queda pendiente la realización de estudios de caso para cada Comunidad Autónoma, en los que se analice con cierta profundidad el peso relativo de los factores en la creación de cada uno de los organismos y su posterior desarrollo en políticas de igualdad autonómicas³³.

De forma general, creemos que tiene especial importancia para la entrada en la agenda del tema de la igualdad, el modelo central del IM y, también, el propio desarrollo del Estado de las autonomías. Respecto al primer factor, hay que tener en cuenta que, además de que el IM tenía entre sus objetivos la promoción de organismos análogos en las comunidades autónomas, una buena parte de las comunidades autónomas estaban en manos del mismo partido que el Gobierno de la Nación, y ya hemos visto como el PSOE había sido protagonista principal en el nivel central. Respecto al propio **desarrollo del Estado de las autonomías** hay que tener en cuenta que no sólo las competencias relacionadas con la "promoción de la mujer" fueron asumidas por todas las comunidades autónomas desde un principio, sino que el hecho de que determinados temas entraran en las agendas de los Gobiernos Autonómicos era una de las vías para la propia institucionalización e identificación de las comunidades autónomas. Quizá tuvieran una relativa menor importancia en aquellos primeros momentos, tanto las influencias internacionales (que como veremos más adelante, van teniendo una importancia creciente en el espectro autonómico, especialmente a partir de 1995 y la IV Conferencia de Pekín), como el peso del movimiento de mujeres en la entrada en agenda de los poderes públicos de los temas de igualdad -en el que, probablemente, exista una variabilidad grande en las diferentes comunidades autónomas.

2.4. El marco institucional de las políticas de igualdad: la creación de organismos específicos de igualdad en los niveles central y autonómico

³³ El único estudio de caso conocido hasta ahora es el realizado por Elena Granados (1999). En él, la autora atribuye especial importancia en la creación del Instituto Andaluz de la Mujer al mimetismo respecto al IM en el nivel nacional.

Como ya se ha mencionado anteriormente, el desarrollo de organismos específicos de igualdad de género en el nivel autonómico en España comienza al final de los años 80, esto es, aproximadamente unos cinco años después de la creación del IM central, en 1983. Como ya ha sido señalado (Valiente, 1996, Martínez, 1997) el IM representó un modelo para la institucionalización autonómica de organismos de igualdad. También se ha mencionado que este desarrollo posterior en el nivel autonómico, respecto al central, tiene que ver con el propio ritmo y desarrollo del Estado de las autonomías en el territorio español. Las comunidades autónomas son creadas entre finales de 1979 (País Vasco y Cataluña) y principios de 1983 (en la que se aprueban los últimos cuatro Estatutos - Extremadura, Baleares, Madrid y Castilla y León-). Generalmente, en todos los Estatutos se mencionan derechos generales referidos a la igualdad de los ciudadanos y se establecen competencias relacionadas con la “promoción de la mujer”. Las comunidades autónomas institucionalizaron estas competencias a través de diferentes estructuras organizativas. Como ya se ha planteado en la introducción de este capítulo, estas “femocracias” autonómicas, junto con el nivel central, han sido y están siendo los actores principales de las políticas públicas de igualdad de género en España. En definitiva, se puede considerar que el marco institucional de estas políticas públicas de igualdad son las diferentes estructuras gubernamentales u organismos específicos que se encargan de impulsar dichas políticas.

En este apartado se analiza este marco institucional tomando en cuenta dos factores organizativos de dichas estructuras gubernamentales: uno, el tipo de organismo y el otro, su posición en el organigrama del respectivo gobierno -central o autonómico-, es decir, a qué departamento gubernamental está adscrito. Estos dos factores nos parecen claves de la capacidad institucional de dichas estructuras. En primer lugar, el tipo de organismo determina el grado de autonomía y en muchos casos fortaleza de la estructura. En segundo lugar, la posición en el gobierno es determinante para tener una mayor autoridad como impulsor de una política horizontal y para que las políticas de igualdad de género no estén especialmente sesgadas hacia una el área que las alberga (bienestar social, empleo, etc.). También se discuten los cambios que han acaecido respecto a estos dos factores en cada una de las comunidades autónomas

En el nivel central, el IM es un organismo autónomo de tipo administrativo, creado por ley, y que estuvo adscrito en sus inicios al Ministerio de Cultura. En 1988 se crea el Ministerio de Asuntos Sociales y una mujer del PSOE de fuerte tradición feminista - Matilde Fernández- asume su cartera y una serie de competencias y organismos entre los que estaba el IM. La posición relativa, en cierto modo, mejoraba entre otras cosas, porque la hasta entonces directora del IM -Carlota Bustelo- es nombrada subsecretaria del nuevo Ministerio. A Matilde Fernández, le sucede otra conocida feminista en la cartera ministerial, Cristina Alberdi. Desde 1996, fecha en la que el Partido Popular gana las elecciones generales, el IM pasa a estar adscrito al Ministerio de Trabajo y Asuntos Sociales, asignando a su cartera a hombres de muy poca conocida trayectoria o sensibilidad feminista (Javier Arenas, Manuel Pimentel y Juan Carlos Aparicio). Es indudable que este organismo es menos "feminista", pero además, al fundirse el Ministerio de Asuntos

Sociales con el de Trabajo, las cuestiones más "sociales" (y que ocupaban el corazón del anterior Ministerio de Asuntos Sociales) adquieren una relativamente menos importante posición en un Ministerio en el que los temas "estrella" son las pensiones, la seguridad social y los planes de empleo. Por otro lado, pone en una buena posición los temas de empleo de los planes de igualdad, aunque se siga identificando los temas de mujer con el área de asistencia social.

Respecto a las comunidades autónomas, se señala a continuación el marco institucional - es decir los organismos de igualdad, su configuración y adscripción- para las políticas de igualdad de género, teniendo en cuenta, por un lado, los datos del año 1998 y, por otro, los actuales³⁴. La información se resume en las tablas 3.3.a y 3.3.b que se representan a continuación. En estas tablas se presentan los siguientes datos:

- Se hace un cuadro de doble entrada con dos factores:

√ Horizontalmente, se señalan las áreas gubernamentales a las que están adscritos los organismos de igualdad. Las cuatro posibilidades existentes son: presidencia, bienestar social, cultura y familia, mujer y juventud.

√ Verticalmente, se ponen los diferentes tipos de organismos posibles: organismos autónomos, direcciones generales de la mujer, secretarías regionales de la mujer, comisiones interdepartamentales y otras direcciones generales no específicas.

- En cada una de las casillas se pone el nombre de la Comunidad Autónoma correspondiente que tiene un organismo de igualdad que corresponde a los dos factores señalados (por ejemplo, Andalucía tienen un organismo de igualdad que es autónomo y está adscrito a Presidencia)

- Se señalan, antes del nombre, el año de creación del organismo y, en su caso, se señala también el momento de su última modificación.

- En paréntesis aparecen los partidos políticos que formaban los diferentes gobiernos autonómicos en el año 1998 (Tabla 3.3.a), y que los forman en la actualidad (Tabla 3.3. b). Si el gobierno es de coalición se señalan los dos partidos con un / entre los mismos.

- Si el área gubernamental a la que está adscrita el organismo tiene más competencias que la señalada en las cuatro categorías iniciales se señala dentro del recuadro con un +. Esto ocurre sólo con las áreas de Bienestar Social, en la que en Canarias comparte competencias con Empleo, en Navarra con Deporte y juventud, en Madrid, Castilla y León y La Rioja con Sanidad.

³⁴ La razón de presentar los dos momentos estriba en que el marco institucional del año 1998 es con el que se realizan la mayoría de las evaluaciones que se estudian en este trabajo y es el que, por lo tanto, se analiza. Sin embargo, en las elecciones autonómicas de 1999 se producen algunos cambios de signo político en algunas comunidades autónomas y algunos cambios significativos de estructuras. Hemos querido reflejar estos cambios en la tabla 3.3. b.

- En las casillas de las otras direcciones generales no específicas, se señalan los nombres de las Direcciones Generales en las que se sitúan las competencias de igualdad de género.

**Tabla 3.3.a Marco institucional de las políticas de igualdad autonómicas.
Año 1998**

Area gubernamental / Tipo de organismo	Presidencia	Bienestar Social	Cultura	Familia, Mujer y Juventud
Organismos Autónomos	88 Andalucía (PSOE/PA) 93 Aragón (PP/PAR) 89 Cataluña (CiU) 88 País Vasco (PNV-EA)	94 Canarias (CC/PP) + Empleo 91-95 Navarra (UPN) + Deporte y Juventud		91 Galicia (PP)
Direcciones Generales de la Mujer	86-97 Cantabria (PP/PRC)	89 Castilla- La Mancha (PSOE) 89 Madrid (PP) + Sanidad 88-97 Comun. Valenciana (PP/UV)	88-91 Extremadura (PSOE)	
Secretarías Regionales de la Mujer		94 Castilla y León (PP) + Sanidad	88-95 Asturias (PP)	
Comisión Interdeparta mental	90 Baleares (PP)			
Otras Direc. Generales no específicas	91-97 Murcia (PP) D.Gnal. Juventud, Mujer y Familia	88 La Rioja (PP) + Sanidad y Consumo D. Gnal. Bienestar Social		

El primer factor a tener en cuenta es el **tipo de organismo**. Como se puede comprobar en la **tabla 3.3**:

a) En siete comunidades autónomas (Andalucía, Aragón, Canarias, Cataluña, Galicia, Navarra y País Vasco) los organismos o agencias de igualdad son **organismos autónomos**, al igual que el IM en el nivel central. Esta circunstancia marca unas diferencias con las diez comunidades autónomas restantes, a saber:

- Son creadas por ley. El resto de los organismos son unidades administrativas que están normalmente reguladas por decretos.
- Tienen personalidad jurídica propia y una autonomía funcional y administrativa mucho mayor. Tienen su propio presupuesto, independientemente de la Consejería o Departamento al que estén adscritos.
- Cuentan con un “Consejo de Administración” en el que existe la posibilidad de que estén representadas otras fuerzas sociales.

b) Los diez organismos de igualdad autonómicos restantes son unidades administrativas con algunas diferencias entre sí, a saber:

1. Una “Comisión Interdepartamental” en Baleares³⁵
2. Cinco “Direcciones Generales de la Mujer” (Cantabria, Castilla-La Mancha, Madrid, Comunidad Valenciana y Extremadura)
3. Dos “Secretarías Regionales de la Mujer” en Asturias³⁶ y Castilla-León, y que vienen a ser muy similares a las direcciones generales.

³⁵ Existe también una Comisión Interdepartamental en Cataluña, que fue creada previamente al *Institut Català de la Dona*, a través del decreto 25/1987 de 29 de Enero, denominada “Comisión Interdepartamental de Promoción de la Mujer” y que en el año 1993 cambia de nombre por el de “Comisión Interdepartamental para la Igualdad de Oportunidades de las Mujeres” (Decreto 209/1993 de 31 de Agosto). Dicha Comisión es la que elabora el I Plan catalán que tiene, entre sus medidas, la creación del Instituto Catalán de la Mujer que es creado por la ley 11/1989 de 10 de Julio. Esta comisión sigue subsistiendo en paralelo con el Instituto, la preside el Presidente de la Generalidad que nombra una vicepresidenta que se hay significado en la labor a favor de las mujeres y en ella están representados todos los departamentos de la Generalidad. Se concibe como un órgano de coordinación e impulso de las acciones que promueven la no discriminación entre hombres y mujeres y sus funciones son las de coordinar y fomentar las acciones y proyectos de los diferentes departamentos y proponer a la Junta de Gobierno del Instituto Catalán de la Mujer las medidas normativas necesarias para fomentar la igualdad de oportunidades y la no discriminación entre hombres y mujeres (Decreto 209/1993 de 31 de Agosto).

³⁶ Con los cambios producidos en el Gobierno del Principado de Asturias en las elecciones de junio de 1999 (en las que accede el PSOE al poder), se crea el Instituto Asturiano de la Mujer en agosto de ese mismo año. Sin embargo, este Instituto no se crea como sus homólogos autonómicos, como un organismo autónomo, sino como un órgano desconcentrado, dependiente de la Consejería de Presidencia, a través del Decreto 61/1999 de 10 de Agosto (B.O.P.A. 12 de agosto de 1999).

4. Dos Direcciones Generales no específicas de temas de mujer. Estas direcciones son las responsables de las políticas y planes de igualdad y de las relaciones con el IM y otras comunidades autónomas

- En La Rioja hay una unidad administrativa pequeña (“Centro Asesor de la Mujer”), dentro de la Dirección General de Bienestar Social. No ha sufrido cambios en su estructura desde su creación (1988). No contar con una unidad administrativa superior específica de mujer puede deberse, entre otras cosas, a la intención de mantener una administración pública autonómica pequeña, acorde con el tamaño de esta Comunidad Autónoma que tiene solamente 250.000 habitantes.
- En el caso de Murcia, también Comunidad Autónoma Uniprovincial pero con una extensión y una población mucho mayor que La Rioja, la Dirección General con competencias en mujer es también responsable de juventud y familia. Sin embargo, y como se analizará más adelante al ver los cambios que se han producido en Murcia sí llegó a haber una dirección general específica. Su configuración actual probablemente se deba más a la concepción de los problemas públicos relacionados con la mujer por parte del Partido Popular que es el que gobierna en Murcia desde 1995.

El segundo factor a analizar es la **Consejería o Departamento al que están adscritos** los organismos de igualdad y su situación en el organigrama de gobierno. Esta categorización también se puede observar en la **tabla 3.3.a**. Siete de los organismos dependen actualmente de los departamentos de **Presidencia** (Andalucía, Aragón, Baleares, Cantabria, Cataluña, Murcia y País Vasco)³⁷. Esto implica una mejor posición para conseguir que otros departamentos formulen y ejecuten las medidas correspondientes contempladas en los planes de igualdad (intentando así, introducir la perspectiva de género a través de todas las políticas), ya que Presidencia es un área gubernamental entendida, en cierto modo, como transversal y, por lo tanto, más “neutral” y con mayor autoridad con respecto a las restantes áreas –también suele ser un área más cercana al poder-. Este tipo de relación y de dependencia parece tener más sentido para políticas públicas que se pretendan globales e integrales (es decir, abordando todas las posibles cuestiones que afectan a las mujeres) y, por supuesto, horizontales (existiendo un compromiso por parte de todos los poderes públicos y no sólo de los relacionados específicamente con los temas de mujer o igualdad).

Entre las restantes, otros siete organismos dependen de los Departamentos o Consejerías de **Bienestar Social** (Canarias, Castilla-La Mancha, Castilla y León, Madrid, Navarra, La Rioja y Comunidad Valenciana). En el caso de Canarias, la Consejería incluye también la denominación y competencias de “Empleo” y en otras tres, Madrid, Castilla-

³⁷ Este número aumenta a ocho en 1999, con la creación del Instituto Asturiano de la mujer, dependiente de la Consejería de Presidencia.

León y La Rioja las competencias de bienestar social comparten Consejería con las competencias de Sanidad. En Galicia, acorde con la ideología conservadora del Partido Popular en el gobierno desde hace varias legislaturas, existe una Consejería específica llamada “Familia, Mujer y Juventud” de la que depende la “Agencia para la Promoción de la Igualdad entre el Hombre y la Mujer”. Los restantes dos organismos (Asturias y Extremadura) dependen de las Consejerías de **Cultura**.

Si se combinan estos dos factores mencionados, se puede tener una idea de cuán favorable o ventajoso es el marco institucional de cada una de las políticas públicas de igualdad de oportunidades autonómicas. Como se puede ver en la tabla 3.3.a, en el año 1998 había cuatro institutos (Andalucía, Aragón, Cataluña y el País Vasco) que cuentan con el marco institucional más favorable.

Las femócratas son conscientes de estos factores y la evolución y los cambios que se han producido a lo largo de estos años han ido, por lo general, en esta dirección. Bien creando un organismo autónomo (como es el caso de Navarra³⁸), creando unidades administrativas superiores o específicas (como en Cantabria, Castilla y León, Extremadura y Asturias)³⁹, o pasando el organismo de otro departamento o consejería al área de presidencia (como en Cantabria y Murcia).

Sólo hay tres excepciones significativas; una de ellas ilustra precisamente esta tendencia: el Instituto Andaluz de la Mujer que en sus comienzos fue adscrito a la Consejería de Presidencia, fue pasado en 1990 a la Consejería de Asuntos Sociales para volver a Presidencia en 1994. La evaluación que hacen sus responsables de este período en Asuntos Sociales es negativo, es decir, se consideran y perciben en mejor posición estando en Presidencia. Las otras dos excepciones tienen que ver con los cambios de tipo político que se dieron en las elecciones legislativas de 1995 en algunas comunidades autónomas al ganar el Partido Popular, entre ellas Madrid y Murcia. En 1995, la Dirección General de la Mujer de Madrid pasa de la Consejería de Presidencia a la de Sanidad y Servicios Sociales, única consejería madrileña ocupada por una mujer (Rosa Posada). Murcia cambia su anterior Dirección General de la Mujer en una nueva y significativamente diferente “Dirección General de Política Social y Familia”, también en 1995; en 1997 se reconvierte en Dirección General de Juventud, Mujer y Familia y se adscribe a la Consejería de Presidencia.

También en la tabla 3.3.a se reflejan los partidos políticos que formaban gobierno en cada Comunidad Autónoma en el año 1998, ya que, como se ha explicado antes, el marco

³⁸ El Instituto Navarro de la Mujer se crea en 1997. Antes existía una Subdirección general de la Mujer que dependía de la Dirección General de Bienestar Social.

³⁹ Cantabria crea en 1997 la Dirección general de la Mujer, dependiente de Presidencia. Antes, existía la Agencia de Promoción de la Mujer que dependía de la Dirección general de Bienestar Social. La Secretaría Regional de la Mujer de Castilla y León pasa de ser una pequeña unidad administrativa dependiente de la Dirección General de Acción Social a depender directamente del Consejero de Bienestar Social en 1994. En Extremadura se pasa de una unidad administrativa pequeña dependiente de la Dirección General de Acción Social (Asesoría Ejecutiva de la Mujer) a una Dirección General de la Mujer en 1991.

institucional correspondiente a aquel año es bajo el que se realizan las experiencias de evaluación analizadas. Así, se puede apreciar la importancia de algunos partidos nacionalistas (en su mayoría de tendencia conservadora) en el mapa político autonómico actual, gobernando en mayoría en las comunidades autónomas de Cataluña y Navarra, y en coalición en otras seis (Andalucía, Aragón, Canarias, Cantabria, País Vasco y Comunidad Valenciana). Con las elecciones autonómicas de 1999, el PSOE “recupera” tres comunidades autónomas (Aragón, Asturias y Baleares), y siguen siendo importantes los pactos con los partidos. Estos cambios se reflejan en la tabla 3.3.b.

**Tabla 3.3.b Marco institucional de las políticas de igualdad autonómicas.
Año 2001**

Area gubernamental / Tipo de organismo	Presidencia	Bienestar Social	Cultura	Familia, Mujer y Juventud
Organismos Autónomos	88 Andalucía (PSOE) 93 Aragón (PSOE/PAR) 90-2000 Baleares (PSOE/PSM/UM) 89 Cataluña (CiU) 88 País Vasco (PNV-EA)	94 Canarias (CC/PP) + Empleo 91-95 Navarra (UPN) + Deporte y Juventud		91 Galicia (1) (PP)
Otro tipo de organismos	88-99 Asturias (PSOE) Instituto Asturiano de la Mujer (órgano desconcentrado) (2)			
Direcciones Generales de la Mujer	86-97 Cantabria (PP/PRC)	89 Castilla- La Mancha (PSOE) 89 Madrid (3) (PP) 88-97 C. Valenciana (PP) 94-2000 Castilla y León (4) (PP) + Sanidad	88-91 Extremadura (PSOE)	
Otros organismos no específicos	91-97 Murcia (4) (PP) Secretaría Sectorial de Mujer y de la Juventud	88 La Rioja (PP) + Sanidad y Consumo D. Gnal. Bienestar Social		

(1) La Consejería pasa a denominarse “Familia, Promoción de Empleo, Mujer y Juventud”

(2) El órgano desconcentrado es una figura de la Administración Institucional autonómica que se establece por decreto

(3) En 1999 la Consejería pasa a ser de Servicios Sociales exclusivamente

(4) La Dirección general se denomina “de la Mujer e Igualdad de Oportunidades”

(5) En Murcia pasa de ser una Dirección General a una Secretaría Sectorial

Como se puede observar en la tabla 3.3.b los cambios que se han producido a partir de las elecciones autonómicas de 1999 y del 2000 (Galicia) son los siguientes:

- Cambio de signo político en los gobiernos autonómicos de Aragón, Asturias y Baleares, que pasan del PP al PSOE, en coalición con el PAR en Asturias y con el PSM y UM en Baleares. En la Comunidad Valenciana, el gobierno pasa a ser exclusivamente del PP y no en coalición con UV como en la legislatura anterior.
- En Baleares se crea por ley un organismo autónomo -el Instituto Balear de la Mujer-. En Asturias también se crea el Instituto Asturiano de la Mujer que pasa, además, a estar adscrito a la Consejería de Presidencia, aunque no se crea como organismo autónomo sino con la figura de “órgano desconcentrado”, creado por decreto y no por ley.
- Se dan algunos cambios de estructura en las Consejerías que afectan indirectamente al marco institucional de los organismos de igualdad: en Galicia la Consejería asume también las competencias de promoción de empleo y en la Comunidad de Madrid se vuelven a separar las competencias de Sanidad y Servicios Sociales, creándose dos Consejerías distintas. La Dirección General de la Mujer se adscribe a la de Servicios Sociales.
- En Castilla y León se pasa de una “Secretaría Regional” a una Dirección General de la mujer e Igualdad de Oportunidades. En Murcia, sin embargo, se pasa de una Dirección General no específica (de Juventud, Mujer y Familia) a una “Secretaría Sectorial” tampoco específica (de Mujer y de la Juventud).

En definitiva, se engrosan las filas de los organismos adscritos a presidencia y la de los organismos autónomos y otras estructuras algo más fuertes que las anteriores.

Finalmente, hay un curioso análisis a realizar. Si tenemos en cuenta el sexo del Consejero o Consejera del área al que está adscrito el correspondiente organismo de igualdad, vemos que, en 1998 (año en el que hemos fijado el marco institucional existente en el momento en el que se realizan las evaluaciones analizadas), casi una de cada tres personas responsables de las Consejerías de las que dependían los organismos de igualdad eran mujeres (5 mujeres consejeras de un total de 17). Este es un alto porcentaje si se considera que sólo el 13% del total de consejeros autonómicos son mujeres. Este hecho tiene probablemente que ver con la tendencia de asignar las competencias de mujer a departamentos que tienen mayor probabilidad de estar dirigidos por mujeres por ser considerados más "femeninos" (como Bienestar Social o Cultura). Curiosamente este panorama varía ligeramente en la estructura actual, siendo esa proporción de 3 mujeres consejeras de un total de 17. A este descenso probablemente se le pueda seguir dando la misma explicación, ya que existe una mayor tendencia a adscribir los organismos de igualdad al área de presidencia, área considerada tradicionalmente “masculina”.

2.5. Una revisión de los actores que contribuyen a la formulación de las políticas públicas de igualdad de género

El estudio de los actores⁴⁰ que contribuyen a la formulación, implementación y evaluación de las políticas públicas es una característica importante del análisis de dichas políticas. Los actores estudiados e identificados por los analistas como implicados en las políticas públicas de igualdad de género incluyen las femócratas y feministas de estado, los movimientos feministas y los grupos de mujeres, los partidos políticos, los sindicatos, los medios de comunicación de masas y el "feminismo académico". Estos actores han sido analizados en el caso español por Valiente (1996) y Martínez (1997), quienes centran su análisis en las políticas públicas de igualdad de género en general (Martínez), o en el nivel central (Valiente). En este apartado, se realiza una revisión de dichos análisis enfatizando el nivel autonómico.

Ya se ha apuntado anteriormente la importancia del **feminismo de Estado** en la aparición, formulación e impulso de las políticas de igualdad de género en España. Se han establecido diferentes organismos de igualdad en el nivel autonómico que han conseguido que los diferentes gobiernos autonómicos lancen planes de igualdad. Por lo tanto, este feminismo institucional también ha jugado un importante papel en las políticas autonómicas. Aunque se han de señalar las diferencias entre las distintas comunidades autónomas -diferentes marcos institucionales, diferentes ritmos y niveles de consolidación- podemos decir en general que las femocracias autonómicas han conseguido sensibilizar a la sociedad española sobre los problemas de la igualdad, introducir dichos problemas en las agendas de sus gobiernos, generar un marco institucional para las políticas públicas de igualdad y hacerse visibles como instituciones públicas.

Como actores importantes del feminismo de Estado en el nivel autonómico hay que considerar el nivel local. Así, las diputaciones provinciales y los ayuntamientos no sólo se consideran algunas veces como destinatarios de las políticas de igualdad de género en el nivel autonómico, sino que también para el cumplimiento de muchas de las acciones propuestas por los planes de igualdad cuentan con ellos como agentes activos. En algunos de los organismos de igualdad existe algún tipo de estructura en la que existe una representación de este nivel local. Por ejemplo en el Instituto Vasco de la Mujer (*Emakunde*), este nivel local está presente en su Consejo de Administración.

Celia Valiente (1996), en su estudio sobre los factores que contribuyeron a la creación del IM en el nivel central y al comienzo de las políticas de igualdad de género, ha señalado la influencia relativa del **movimiento feminista y los grupos de mujeres**. También ha destacado que las relaciones entre las feministas y las femócratas, durante ese

⁴⁰ Al igual que hace Eva Martínez (97:), utilizamos el término "actores" en masculino ya que el término en femenino no se utiliza, aunque resulta evidente que sería más lógico, en este caso, hablar de "actrices".

período, no han sido especialmente fluidas. Sin embargo, como señala Martínez, sí ha habido un impacto del movimiento feminista en el feminismo de Estado español, aunque este impacto ha podido variar dependiendo de los diferentes momentos. Con todo, es cierto que las organizaciones feministas españolas no son ni han sido tan fuertes y potentes como en otras democracias occidentales y, además, no han estado muy dispuestas a crear redes entre ellas (Martínez, 1997: 226-227). Si nos centramos en el ámbito autonómico, vemos que, a grandes rasgos, existen diferencias entre las comunidades autónomas. La presencia de los movimientos de mujeres en las políticas públicas de igualdad es mayor en algunos lugares, como por ejemplo, Cataluña, Canarias o el País Vasco, que en otros. Para estudiar a fondo la presencia o ausencia de los movimientos de mujeres en las políticas públicas de igualdad de género españolas, hacen falta estudios de caso en cada Comunidad Autónoma que analicen la flexibilidad y la disposición de las estructuras femocráticas y las femócratas para la participación de los movimientos de mujeres, la fortaleza de cada movimiento y el marco de relaciones entre femócratas y feministas en la formulación, ejecución y evaluación de las políticas de igualdad. Otra de las cuestiones que quedan por estudiar es el papel que están cumpliendo algunos grupos de mujeres no feministas (como asociaciones de viudas, de amas de casa, algunas asociaciones de mujeres rurales) en las políticas de igualdad de género en España en este momento, ya que los gobiernos conservadores tienden a estar especialmente dispuestos a recabar su participación en las políticas de igualdad.

Otros actores importantes suelen ser en las políticas públicas los **partidos políticos**. Ya se ha señalado la importancia del Partido Socialista Obrero Español (PSOE) en la creación del IM y el comienzo de las políticas públicas de igualdad de género en el nivel central. El protagonismo del PSOE en las políticas de igualdad fue también alto al comienzo de las políticas públicas de igualdad de género autonómicas. En primer lugar, por la propia influencia del IM y, en segundo lugar, porque era el partido que gobernaba la mayoría de las comunidades autónomas en aquel momento (10 de 17 en 1987). Con todo, a medida que ha ido pasando el tiempo, el PSOE ha ido perdiendo ese protagonismo en las políticas de igualdad de género. Esto se debe, principalmente a su posición en el mapa político autonómico, siendo especialmente débil desde que en las elecciones autonómicas de 1995 pierde siete Comunidades y en 1996 pierde las elecciones generales. Además, las políticas públicas de igualdad de género autonómicas se han ido desarrollando por su cuenta, con una influencia decreciente de la política a nivel central.

La llegada al poder del Partido Popular en el nivel central y en la mayoría de las comunidades autónomas lógicamente ha afectado a las políticas de igualdad de género ya que la concepción del papel de la mujer y cómo deben actuar los poderes públicos al respecto cambia en cierta manera con una ideología conservadora. Sin embargo, el freno que se le ha puesto a las políticas de igualdad por parte de los gobiernos más conservadores ha sido menor, al menos al nivel formal, de lo que se esperaba o era previsible en su momento. Celia Valiente escribía que el futuro a largo plazo del IM era incierto, si ganaba las elecciones el partido conservador, no sólo porque éste no estaba en ningún modo identificado con el feminismo de Estado sino, sobre todo, porque el IM era una institución muy ligada a e identificada con el PSOE (Valiente, 1995: 234; 1996: 197). Nuestra hipótesis de este mantenimiento de las instituciones femocráticas y de la estructura formal

para las políticas de igualdad de género -esto es, los planes de igualdad- por parte de los gobiernos conservadores es la influencia tan importante que ha tenido la IV Conferencia Mundial de la Mujer en Pekín en Septiembre de 1995 y su Plataforma de Acción, así como el IV Programa de Acción Comunitario para la Igualdad de Oportunidades para Mujeres y Hombres (1996-2000).

En el análisis de los partidos políticos como actores de las políticas públicas autonómicas, hay que tener en cuenta, como ya se ha mencionado en el apartado sobre el marco institucional, a los partidos nacionalistas. Como se vio en la tabla 3.3.a hay diferentes partidos nacionalistas en ocho de las 17 comunidades autónomas, bien sólo (Cataluña, Navarra) o en coalición, siendo el partido mayoritario de dicha coalición (Canarias, País Vasco) o el minoritario (Andalucía, Aragón, Cantabria, Valencia). Incluso en el nivel nacional, el PP está gobernando desde 1996 en minoría con el apoyo de CiU, PNV y Coalición Canaria, cuestión que pone a estos partidos en una posición estratégica para influir en las políticas públicas, tanto en el nivel central como en el autonómico. En lo que respecta a las políticas de igualdad autonómicas, también es importante señalar que en Cataluña y el País Vasco los gobiernos autonómicos llevan siendo nacionalistas desde hace mucho tiempo y estas dos comunidades autónomas, junto con Andalucía, cuentan con los Institutos de la Mujer autonómicos más antiguos y estables. Se puede considerar que son líderes de las femocracias y políticas de igualdad autonómicas.

En los **sindicatos** también se han movilizado mujeres feministas que han tenido cierta influencia en las políticas de igualdad de género, especialmente en las políticas relacionadas con el empleo. Efectivamente, en España se han impulsado en los últimos años bastantes medidas para la promoción del empleo femenino y, en esto, han tenido un papel importante las secretarías de la mujer de los principales sindicatos. En los últimos años, también están teniendo un papel importante como *agenda setters* y como impulsoras de medidas en los temas de acoso sexual en el ámbito laboral.

También se ha reconocido la importancia del llamado **feminismo académico** en las políticas de igualdad de género, especialmente en el estudio de la situación de la mujer y sus necesidades y en señalar algunos temas importantes para su entrada en las agendas políticas. En el caso de España, la relativa importancia creciente de este feminismo académico podría ser más consecuencia de las primeras políticas de igualdad de género y su apoyo y fomento que factor desencadenante de las mismas. El papel que ha jugado el IM a nivel nacional ha sido importante no sólo apoyando la investigación (apoyando los seminarios universitarios de mujer y a feministas académicas individuales, además de creando un programa sectorial específico dentro del Plan Nacional de Investigación), sino también sintetizando, publicando y divulgando la investigación en temas de mujer. Al igual que los planes de igualdad nacionales, todos los planes de igualdad autonómicos contienen acciones referidas al apoyo de la investigación en temas de mujer. Sin embargo, está claro que este feminismo académico todavía no es demasiado potente en España, sobre todo si lo comparamos con otros países europeos o norteamericanos. Esta carencia es especialmente grave en lo que se refiere al estudio de las políticas públicas, tanto de las propias políticas de igualdad de género como en el estudio del resto de las políticas públicas con perspectiva

de género⁴¹. Probablemente, tanto el apoyo continuado como la necesidad de evaluar sistemáticamente los planes y las políticas emprendidas en los próximos años fortalezca el papel de este actor y haga aparecer nuevos equipos de investigación en este área.

⁴¹ Para tener una visión general del estado de la cuestión en los estudios de mujeres y política, consultar Uriarte, 1997 y Valiente, 1998b.

3. LOS PLANES DE IGUALDAD COMO INSTRUMENTO BÁSICO PARA LA ARTICULACIÓN DE LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO

3.1. Qué significan los planes de igualdad y qué papel cumplen en las políticas de igualdad de género

Como ya se ha señalado anteriormente, en España, las estructuras gubernamentales relacionadas con la mujer u organismos de igualdad que conforman el “feminismo de Estado”, tanto en el nivel central como en el autonómico, se configuran más como impulsoras que como ejecutoras directas de las políticas de igualdad. Es decir, lo que pretenden es persuadir a las otras unidades de sus respectivos gobiernos que asuman una serie de objetivos para promover la igualdad de género en cada una de las áreas de las que son responsables. Esta es la razón por la que el IM, después de sus primeros años de existencia, adoptó como principal instrumento de las políticas de igualdad la elaboración y aprobación de planes de igualdad de oportunidades para las mujeres. Todas las comunidades autónomas han seguido el mismo camino. El I Plan nacional se aprueba en el año 1988 y los primeros autonómicos en 1989 (Asturias, Cataluña, Madrid y Comunidad Valenciana). Esto significa que en España se lleva ya diez años articulando las políticas de igualdad de género a través de este instrumento que son los planes de igualdad. Por lo tanto, a parece como especialmente interesante realizar un análisis de dichos planes en los niveles central y autonómico.

Así, muchas de las investigaciones más recientes tienen en cuenta el análisis de dichos planes (Gil Ruiz, 1996; Martínez, 1997; Valiente, 1997; Sensat y Varella, 1998), ya que éstos son el principal instrumento con el que se articulan las políticas de igualdad, desde los organismos específicos que promueven la igualdad de género.

Estos planes son un conjunto de objetivos y medidas tomadas y aprobadas por un gobierno y que deben llevar a cabo los diferentes departamentos gubernamentales, en un período concreto de tiempo (de dos a cinco años). Los planes también suelen implicar, al menos formalmente, otros niveles gubernamentales y organizaciones no gubernamentales.

La concepción de planes de carácter integral como son los planes de igualdad está íntimamente ligado a los conceptos de "perspectiva de género" y "*mainstreaming*" o transversalidad de dicha perspectiva de género, que son conceptos clave de las políticas de igualdad de género a nivel internacional (tanto la Plataforma de Acción de la IV Conferencia Mundial de la Mujeres en Pekín, 1995 como el IV Programa de Acción Comunitaria, 1996-2000, establecen sendas directrices y recomendaciones respecto a

dichos conceptos). Conseguir que todas las políticas públicas sean formuladas, ejecutadas y evaluadas con perspectiva de género es la meta. El "*mainstreaming*" o transversalidad es la estrategia para conseguir dicha perspectiva de género.

La **perspectiva de género** implica poner en cuestión la idea implícita de que las políticas públicas sean neutrales respecto al género, ya que éstas reproducen las estructuras sociales en cuya base se encuentra la desigualdad, entre otras razones, porque muchas políticas públicas afectan diferencialmente a los colectivos masculino y femenino; por lo tanto, se trata de introducir la variable género desde la propia definición del problema público que se intenta resolver y a la hora de formular, ejecutar y evaluar las acciones públicas ¿La definición del problema contempla alguna variación respecto al género? ¿Se está formulando, ejecutando y evaluando la política teniendo en cuenta dichas diferencias? Un buen ejemplo para ilustrar este concepto puede ser el de las políticas de prevención de drogodependencias dirigidas al colectivo juvenil. Para abordar este tema y formular acciones públicas al respecto se estudian los patrones de consumo de drogas y los estilos de vida y actitudes ante la salud. Al estudiar los datos más recientes se ve que, junto con la variable de edad, la variable sexo es la que explica las mayores diferencias respecto a estos patrones de conducta y actitudes hacia la salud. Mientras la variable de edad sí es más atendida a la hora de plantear las acciones públicas a llevar a cabo, la del sexo no se considera de igual manera. Es decir, los programas de prevención de drogodependencias en el ámbito escolar están (teóricamente) dirigidos a chicos y chicas, sin hacer diferencias. ¿A qué destinatario tenían en la cabeza los formuladores de dichos programas? ¿A qué perfil de destinatario -siendo el de las chicas y los chicos tan distintos en este terreno- se dirigen? ¿Existen diferencias en cuanto al impacto en unos y otras? Nos atrevemos a establecer una hipótesis tentativa sobre un posible sesgo hacia el perfil masculino como destinatario tipo de la mayoría de las acciones.

El concepto de "*mainstreaming*", de difícil traducción al castellano⁴² es la estrategia que se propone para conseguir esta perspectiva de género. No se trata de hacer un "ghetto" de las políticas de igualdad, realizando exclusivamente acciones específicas para las mujeres, sino que se trata de impregnar de perspectiva de género el resto de las políticas que abordan todo tipo de problemas públicos que, lógicamente, también afectan a las mujeres y, en muchos casos, de forma diferencial respecto a los hombres. De alguna manera, además de la idea de "transversalidad" tiene que ver también con el concepto de "normalización", con la idea de "normalizar" la perspectiva de género en la formulación de todas las políticas públicas, con la idea de integrar dicha perspectiva de género en la "corriente principal".

⁴² Existen muchas críticas a la traducción de *mainstreaming* como "transversalidad" porque se entiende que es algo más; se cree que por "transversalidad" se puede interpretar la simple contabilidad de las mujeres que se han beneficiado de una determinada actuación o política. Quizá esto podría ser subsanado traduciéndolo como "transversalidad de la perspectiva de género", aunque parece que es un término algo complejo. En cualquier caso, la utilización del término en inglés se ha generalizado en la comunidad de política (*policy network*) de la igualdad de género. Este no es un fenómeno exclusivo de España, sino que se da internacionalmente.

Ambos conceptos (perspectiva de género y *mainstreaming*) están intrínsecamente relacionados entre sí, de forma que el término más utilizado en la actualidad en la comunidad internacional es el de “*gender mainstreaming*” que podría ser traducido como “transversalidad de género”. Este término implica tanto el objetivo (la perspectiva de género) como la estrategia para conseguirlo (la transversalidad y “normalización” en todas las políticas públicas). El concepto de “*gender mainstreaming*” apareció por primera vez en los textos internacionales después de la III Conferencia Mundial de la Mujer en Nairobi y se integra plenamente en la Plataforma de Acción de Pekín como concepto clave. Dicha Plataforma hace un llamamiento constante a promover la política del *gender mainstreaming* o de la transversalidad de género, señalando repetidamente que “los Gobiernos y otros actores deben promover una política activa y visible de transversalidad de la perspectiva de género en todas las políticas y programas, de forma que, antes de que las decisiones sean tomadas, se haga un análisis de los efectos en las mujeres y los hombres, respectivamente”⁴³. Como la Plataforma, sin embargo, no ofrece ninguna orientación concreta sobre cómo desarrollar y ejecutar esta política de la transversalidad de género, se están haciendo esfuerzos tanto en el nivel nacional en los diferentes países⁴⁴ como en el nivel internacional por su definición y difusión. En el nivel internacional, destaca el esfuerzo realizado por el Consejo de Europa, a través de su Grupo de Especialistas en *Mainstreaming*. Este grupo establece una definición de *gender mainstreaming* apuntando a su meta, a su proceso, a sus objetos y a sus sujetos activos. Así, la transversalidad de género (*gender mainstreaming*) es la (re)organización, mejora, desarrollo y evaluación de los procesos de políticas, de forma que la perspectiva de la igualdad de género se incorpora en todas las políticas y a todos los niveles y en todas las fases, por parte de los actores normalmente implicados en la formulación de políticas (*policy-making*)” (Consejo de Europa, 1998: 19).

En cualquier caso, ésta es una estrategia común e internacionalmente aceptada, al menos en el nivel teórico. Stetson y Mazur (1995) concluyen en su estudio sobre feminismo de Estado comparado que todas las agencias o estructuras gubernamentales que son clasificadas, según su análisis, en la categoría más alta de feminismo de Estado, cuentan con este enfoque centralizado (a través de un organismo específico) pero horizontal y trans-sectorial (que intenta llegar a todos los sectores y políticas) a la hora de promover la igualdad de género en las políticas “normales” o “corrientes” (*mainstream policy*).

Por lo tanto, los **planes de igualdad** son un instrumento que tiene como filosofía subyacente la perspectiva de género en los diferentes sectores o políticas, ya que, como se ha comentado anteriormente, estos planes son un conjunto de objetivos y medidas que se deben llevar a cabo por los diferentes departamentos gubernamentales y no sólo por el organismo de igualdad correspondiente.

⁴³ Véanse los párrafos 79, 105, 123, 141, 164, 189, 202, 229, 238, 252, 273 de la Plataforma de Acción de la IV Conferencia Mundial de la Mujer. Pekín, 1995.

⁴⁴ En España existe algún proyecto como el “Hera 2001. Visibilizando el Mainstreaming” que desarrolla la Fundación Mujeres en el marco del IV Programa Comunitario de Igualdad de Oportunidades. Para más información consúltese la página web: www.fundacionmujeres.es/hera2001.

Esta concepción de planes promueve, o al menos facilita las siguientes características:

Conciencia de los poderes públicos sobre el problema de la igualdad. Los poderes públicos deben pensar sobre el problema y sobre cómo resolverlo. Este factor refuerza, además, la fundamentación de las políticas de igualdad de género.

Compromiso Público: los planes se aprueban normalmente en Consejo de Ministros o de Gobierno. En algunas comunidades autónomas también se debe informar o presentar el plan en el parlamento autonómico correspondiente. Estas circunstancias implican un compromiso formal importante respecto a la igualdad de género.

Globalidad: los planes se diseñan para ser globales y comprensivos, es decir, para tratar el problema de la igualdad desde diferentes ámbitos y sectores (empleo, educación, legislación, salud, etc.) y desde diferentes perspectivas. Este factor hace que se vea el desarrollo de la mujer desde una perspectiva más integral y realista y que se vaya superando poco a poco la visión tradicional y limitada de relacionar casi exclusivamente la promoción de la mujer y la igualdad de género con los temas de bienestar social.

Visión estratégica: Un buen plan significa que se deben establecer unas líneas estratégicas y relacionarlas con períodos de tiempo concretos y determinados. Se tienen también que prever con antelación los recursos -tanto económicos como humanos e institucionales- y se deben adquirir compromisos respecto a los mismos. Esta es una gran ayuda para situar a los poderes públicos sobre dónde quieren llegar y cómo y cuándo piensan hacerlo.

Horizontalidad: La globalidad de los planes de igualdad requiere de la participación activa de diferentes departamentos y sectores gubernamentales. Este es el primer paso y una condición necesaria para promover el "mainstreaming" de la perspectiva de género hacia todas las políticas públicas.

Coordinación interdepartamental: Con un plan horizontal, se estimula, de alguna manera, a los diferentes departamentos a que trabajen juntos o, al menos, que sepan y conozcan lo que están haciendo otros.

Participación: La mayoría de los últimos planes de igualdad en el nivel central y autonómico han contado con un "proceso de elaboración" anterior a su aprobación. Estos procesos normalmente incluyen algún tipo de participación de grupos de mujeres, aunque en algunos casos dicha participación pudiera ser, sobre todo, una forma de legitimar al organismo de igualdad y al plan. En general, se da menos participación en el proceso de ejecución de los planes y prácticamente ninguna en los procesos de evaluación, si éstos existen. Sin embargo, la estructura de plan estimula la participación de diferentes agentes y, a veces, ofrece un marco institucional para la misma.

Seguimiento y evaluación: La utilización de planes como instrumento principal de las políticas lleva a pensar en términos de hacer seguimiento y evaluar dichos planes, especialmente cuando esos planes cubren un período específico de tiempo y se supone que

deben ser seguidos por otros planes. Las femócratas han comenzado a pensar en la evaluación de los planes de igualdad como algo necesario y deseable. Más aún, esto puede ser importante para la creación de una "cultura evaluativa" en las administraciones públicas, dado además, el carácter horizontal de dichos planes.

Por supuesto, todos estos factores están presentes en los planes de igualdad de forma potencial. Es decir, aunque la estructura de los planes los permiten, facilitan y estimulan, también estos pueden quedarse en compromisos exclusivamente formales o en declaraciones generales de buenas intenciones.

3.2. La evolución de los planes de igualdad en España

Ya se ha señalado que el primer plan de igualdad en el nivel nacional fue aprobado en 1988 (I Plan de Igualdad de Oportunidades para las Mujeres 1988-90) y sólo un año más tarde se aprueban planes en cuatro comunidades autónomas (Asturias, Cataluña, Madrid y Comunidad Valenciana). En 1990 se aprueban otros dos planes en otras dos comunidades autónomas (Andalucía y Castilla-La Mancha). En 1991 eran ya once comunidades autónomas las que ya habían lanzado y aprobado sus respectivos primeros planes. En los últimos años, la aprobación de planes se ha generalizado a todas las comunidades autónomas (Galicia en 1999, Murcia en 1993, Aragón y Castilla y León en 1994, Canarias en 1995 y la última, Navarra, con su primer plan aprobado en 1998. En la **tabla 3.4** se muestra un cronograma de todos los planes de igualdad centrales y autonómicos.

Como se puede observar en dicha tabla 3.4, cuando finaliza el **período de aplicación** de un plan, no siempre se ha continuado de forma inmediata con el siguiente. Normalmente, el proceso de elaboración y posterior aprobación es largo y, en algunos casos, se señala que dicho retraso es debido, precisamente, a la necesidad de evaluar los resultados del plan, los cuales no suelen estar listos nada más acabarlo. Salvo en los casos de Cantabria, Extremadura y Comunidad Valenciana, en donde este período "interplanes" ha resultado mucho más largo por diferentes razones de inestabilidad organizativa, las "lagunas" de tiempo suelen durar un máximo de uno o dos años. La tendencia parece dirigirse, además, hacia la reducción de estos períodos.

TABLA 3.4. EVOLUCION DE LOS PLANES CENTRALES Y AUTONOMICOS DE IGUALDAD DE GENERO EN ESPAÑA

	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03
CENTRALES	I Plan 88-90			II Plan 93-95				III Plan 97-2000								
Andalucía	I Plan 90-92				II Plan 95-97											
Aragón					I Plan 94-96			II Plan 97-2000								
Asturias	I Plan 89-91			II Plan 93-95			III Plan 1996-2000									
Baleares	I Plan 91-93				II Plan 1996-99											
Canarias					I 95-96		II Plan 97-2000									
Cantabria	I Plan 91-93				II Plan 98-2001											
Castilla y León					I Plan 94-96			II Plan 97-2000								
Castilla-La Mancha	I Plan 1990-1994			II Plan 1995-1999			III Plan 1999 - 2003									
Cataluña	I Plan 1989-1992			II Plan 94-96			III 98-2000									
Extremadura	I Plan 1991-?															
Galicia				I Plan 92-94		II Plan 95-97		III Plan 98-2001								
Madrid	I Plan 89-91			II Plan 93-95			III Plan 97-2000									
Murcia					I Plan 93-95			II Plan 97-2000								
Navarra									I 98-2000							
La Rioja				I Plan 1991 - 1995			II Plan 1996-2000									
C. Valenciana	I Plan 89-91							II Plan 97-2000								
País Vasco				I Plan 1991-1994			II Plan 1995-98		III Plan 1999-2002							

La duración de los planes es variable y no parece deberse a ninguna razón concreta, aunque se puede observar una tendencia hacia planes ligeramente más largos. Esto puede deberse a que algunos de los primeros planes eran de “tanteo”, de manera que tendían a ser más cortos en sus inicios. Como se puede observar en la **tabla 3.5** el total de 38 planes centrales y autonómicos aprobados en el año 1998⁴⁵, sólo ha habido uno de dos años (I canario), 21 (55%) de tres años, 10 (26%) de 4 años y 6 (16%) de 5. Es decir la mayoría de ellos han sido de tres años y la media de duración se sitúa en tres años y medio. Curiosamente, ningún plan rebasa la barrera psicológica del año 2000.

TABLA 3.5. DURACION DE LOS PLANES DE IGUALDAD

DURACION	nº planes	%
Dos años	1	3%
Tres años	21	55%
Cuatro años	10	26%
Cinco años	6	16%
TOTAL	38	100%

Duración media= 3,5

En lo que se refiere al **formato** de los planes, se pueden establecer dos tendencias. Por un lado, los planes tienden a ser crecientemente **más específicos y “realistas”**; es decir, los primeros planes eran más generales y tendían a tener un mayor número de acciones que, por término general, se han ido reduciendo como demuestra la **tabla 3.6**.

⁴⁵ A finales del año 1998 se aprueba el III plan gallego (1998-2001) y en el año 1999 el III castellano-mancheño (1999-2003). Estos dos planes no están incluidos en el total de 38 analizados.

Tabla 3.6
NUMERO DE AREAS, OBJETIVOS Y ACCIONES
PLANES DE IGUALDAD CENTRALES Y AUTONOMICOS

	I PLANES	II PLANES	III PLANES
<i>PLANES NACIONALES</i>	<i>6-30-122</i> <i>(88-90)</i>	<i>10-46-170</i> <i>(93-95)</i>	<i>10-25-192</i> <i>(97-2000)</i>
ANDALUCIA	7-28- 200 (90-92)	4-21- 66 (95-97)	
ARAGON	8-34- 239 (94-96)	9-25- 154 (97-2000)	
ASTURIAS	6-26- 87 (89-91)	6-16- 62 (93-95)	(ND) (97-2000)
BALEARES	ND (91-93)	7-7- 68 (96-99)	
CANARIAS	7(26)-51- 168 (95-96)	7(34)-71- 225 (97-2000)	
CANTABRIA	7-22- 172 (91-93)	10-15- 99 (98-2000)	
CASTILLA-LA MANCHA	ND (90-94)	5-14- 65 (95-99)	
CASTILLA y LEON	8(19)-42- 227 (94-96)	5-7- 64 (97-2000)	
CATALUÑA	6- 34 (89-92)	7(26)-56- 157 (94-96)	8-28- 98 (98-2000)
EXTREMADURA	7-7- 126 (91-sin plazo)	(en elaboración)	
GALICIA	5- 56 (92-94)	5(8)-16- 84 (95-97)	
MADRID	7-26- 160 (89-91)	3-13- 83 (93-95)	4-14- 127 (97-2000)
MURCIA	6-29- 142 (93-95)	6-28- 82 (97-2000)	
NAVARRA	10-23- 76 (98-2000)		
PAIS VASCO	7(34)-86- 320 (91-94)	9(32)-92- 397 (96-2000)	
LA RIOJA	5-28- 158 (91-95)	7-45- 110 (96-99)	
C.VALENCIANA	7-32- 125 (89-91)	11-69- 177 (97-2000)	

Medias
Autonómicas

$\Sigma = 2290/15$
Media act.= 153

$\Sigma = 1893/15$
Media act.= 126

$\Sigma = 225/2$
Media act.= 113

Media de actuaciones de los planes autonómicos en 1998 (en itálica) $\Sigma = 1700/15 = 126$
 (Quitando el caso extremo del País Vasco= 107)

Esta tendencia hacia planes más específicos tiene que ver, en cierto modo, con las necesidades de seguimiento y evaluación, ya que es muy difícil, si no imposible, seguir y evaluar objetivos y acciones que estén formulados de forma vaga y general.

Por otro lado, existe una segunda tendencia a formular **mayores compromisos explícitos** de los poderes públicos. Estos pueden consistir en señalar los Ministerios, Consejerías o Departamentos que son responsables de cada una de las acciones, en un compromiso formal de seguir y evaluar el plan o por una intención formal de dar a conocer el Plan al correspondiente Parlamento autonómico. Un ejemplo ilustrativo es cómo el I Plan de Igualdad de Oportunidades para las Mujeres (1988-90) en el nivel central, no fue aprobado por el Consejo de Ministros, sino que éste "tomó conocimiento" de dicho Plan, es decir, "anunció públicamente que sabía de su existencia, pero que no lo consideraba una tarea de Gobierno" (Valiente, 1994:25). Esto no sólo ya no ocurre con el II Plan central (1993-95) que sí es aprobado formalmente por el Consejo de Ministros, ni tampoco con ninguno de los planes autonómicos que son todos aprobados por acuerdo de los respectivos Consejos de Gobierno.

Además de las tendencias evolutivas respecto al período de aplicación y duración de los planes y respecto al formato de los mismos, los planes de igualdad han evolucionado respecto a las **influencias externas**, teniendo una importancia creciente la influencia internacional, expresada actualmente a través de la Plataforma de Acción de Pekín y el IV Programa Europeo. Esta tendencia se aprecia especialmente si se analiza el conjunto de los planes autonómicos de igualdad. En estos últimos planes se pueden encontrar, además, dos tendencias opuestas: una creciente influencia internacional y una decreciente influencia nacional. Prácticamente todos los planes de igualdad autonómicos actuales mencionan en su presentación, y al enmarcar y fundamentar los planes, tanto la IV Conferencia de Pekín y su Plataforma de Acción y el IV Programa de Acción Europeo. Sólo hay tres planes que no lo hacen y son los que fueron aprobados antes de Septiembre de 1995, fecha en la que se celebra la IV Conferencia Mundial: los II Planes de Andalucía, aprobado en enero de 1995, el de Castilla-La Mancha aprobado en febrero y el de y Galicia aprobado en marzo del mismo año. El II plan andaluz, sin embargo, hace una referencia general al impulso tanto de las Naciones Unidas como de la Unión Europea en lo que respecta a la promoción de la mujer. Sin embargo, no todos los planes autonómicos actuales mencionan los planes nacionales o el IM. Más allá, sólo se pueden encontrar referencias generales a los planes nacionales en tres de los planes autonómicos (Andalucía, Aragón y La Rioja).

3.3. La estructura y contenidos de los planes de igualdad en España

En el análisis de los planes de igualdad de género en España, tanto en el nivel central como en el autonómico, conviene detenerse en su estructura y sus contenidos a grandes rasgos, aunque queda pendiente un análisis más pormenorizado de cada uno de los planes y que excede el objeto de este trabajo de investigación.

Los **contenidos** de los planes de igualdad son relativamente homogéneos en lo que se refiere a las “grandes áreas” que organizan las diferentes acciones. Como ya se ha comentado anteriormente, los primeros planes autonómicos estuvieron especialmente influenciados por los planes nacionales. Posteriormente, la influencia nacional pierde fuerza y sirven como modelos los programas y la legislación europea e internacional. Los planes no son, sin embargo, tan similares entre sí respecto al número de acciones, la concreción de dichas acciones o la importancia relativa de cada conjunto de acciones.

En general, las “grandes áreas” de los planes de igualdad son: **legislación** (adaptación, información y publicidad), **educación** (programas de co-educación, formación del profesorado, contenido de los currícula, estímulo y vigilancia de libros de texto no discriminatorios, ...), **salud** (planificación familiar, programas de prevención de cáncer y de menopausia, humanización del embarazo y parto, ...), **servicios sociales** (programas para mujeres en situaciones de exclusión social, madres solteras, prostitución, ...), **empleo** (estudios del mercado laboral, formación ocupacional, asesoramiento y apoyo a emprendedoras, ...), **participación** (apoyo a grupos de mujeres, ...), y **cooperación** (algunos se refieren a la cooperación nacional especialmente con administraciones locales y otros a la cooperación internacional⁴⁶). Existen otras áreas que se estructuran normalmente dentro de otras como son **investigación** (apoyo a seminarios universitarios de mujer, becas y ayudas, ...), **cultura** (apoyo a mujeres artistas, actividades culturales, ...), **imagen pública de la mujer** (promover una imagen más ajustada a la realidad, control de publicidad, ...). Finalmente, hay algunos temas “de moda” que han aparecido o han cobrado importancia específica en los últimos planes como son **medioambiente, urbanismo, mujer rural o violencia** (este último siempre ha estado presente pero en los últimos planes suele aparecer como un área diferenciada).

Respecto al **tipo de acciones**, tanto las catalanas como las vascas en sus respectivos institutos, han elaborado diferentes categorizaciones de acciones. Además de los grupos destinatarios y las ya mencionadas áreas, han desarrollado otras interesantes categorías para clasificar las acciones.

⁴⁶ De los quine planes autonómicos actuales, dos se refieren a la cooperación internacional (Aragón y Castilla-La Mancha), cuatro a la cooperación nacional y la relacionada con la Unión Europea (Castilla-León, Madrid, Murcia y La Rioja), otros cuatro tanto a cooperación nacional como internacional (Canarias, Navarra, País Vasco y Comunidad Valenciana) y las cinco restantes mencionan el término "cooperación" relacionado con la promoción y apoyo a grupos de mujeres (Andalucía, Baleares, Cantabria, Cataluña y Galicia).

En *Emakunde* (Instituto Vasco de la Mujer) se han establecido siete categorías según tipos de acción, según sean de 1. sensibilización, información, divulgación; 2. orientación, asesoramiento, apoyo; 3. formación; 4. realización de estudios, investigaciones, evaluación; 5. adaptación y ampliación de recursos; 6. coordinación, colaboración, intercambio y 7. intervención sobre las normas jurídicas y sociales (estas últimas acciones tratan de modificar las situaciones de base que generan discriminación). De acuerdo con la memoria de *Emakunde* de 1996, el 65% de las acciones realizadas pertenecían a las categorías 1, 3 y 5, y sólo el 5% de la última categoría. En la memoria de 1997, el porcentaje para la suma de las categorías 1, 3 y 5 seguía siendo prácticamente idéntico, subiendo algo el de la última categoría (8, 4%).

Las catalanas han desarrollado tres categorías basadas en el “tipo de política” que subyace a cada acción, y que clasifican de la siguiente manera: a) promoción (servicios dirigidos a la mujer), b) igualdad (acciones dirigidas a establecer las condiciones necesarias para que hombres y mujeres tengan las mismas oportunidades) y c) paridad (dirigidas a modificar las relaciones sociales que hagan que hombres y mujeres compartan responsabilidades en el ámbito público y el privado). En el análisis realizado en el Instituto Catalán de la Mujer, la mayoría de las acciones son de promoción y de igualdad; por lo tanto, las acciones de paridad deben ser reforzadas.

Ambos análisis sugieren que las acciones dirigidas a una transformación social verdadera tienen todavía una importancia relativa en los planes. Como se verá más adelante en el análisis de las evaluaciones, este interesante tipo de análisis se ha hecho, exclusivamente, en el País Vasco y Cataluña, precisamente en dos de las comunidades autónomas donde tanto los organismos como las políticas de igualdad están relativamente más consolidados.

3.4. Fortalezas y debilidades de los planes como principales instrumentos de las políticas de igualdad de oportunidades

En el punto anterior de este capítulo se han mencionado algunas características de los planes de igualdad que pueden ser consideradas como fortalezas potenciales de dicho instrumento y que pueden ser transformadas en oportunidades para su utilización efectiva como instrumento al servicio de la política. Como ya se ha visto, los planes de igualdad facilitan una mayor conciencia por parte de los poderes públicos, una visión global y estratégica, mayores cotas de compromiso público, la horizontalidad de la política, una mayor coordinación interdepartamental, oportunidades estructuradas para la participación y una mayor preocupación por el seguimiento y la evaluación de las acciones emprendidas. Otros dos factores no intrínsecos a los propios planes pueden ser añadidos al listado de fortalezas. El primero de ellos es que este tipo de instrumento está apoyado

internacionalmente, tanto por la Plataforma de Acción de la IV Conferencia Mundial de la Mujer de Pekín (Naciones Unidas) como por el IV Programa de Acción Comunitario de la Unión Europea. El segundo factor es que los planes de igualdad de género, como instrumento, representan un marco de referencia para ser emulado por otros poderes públicos (especialmente por los gobiernos locales).

Sin embargo, los planes de igualdad también tienen una serie de debilidades, también potenciales, y que pueden convertirse en amenazas de este instrumento de la política.

En primer lugar, no sólo las políticas públicas de igualdad de género en España son todavía muy pequeñas si las comparamos (tanto en términos presupuestarios como en términos de atención de los poderes públicos) con otras políticas públicas, sino que su principal instrumento, los planes de igualdad, dependen en cierto modo de la **persuasión**. Es decir, los agentes promotores de dichos planes que son los organismos de igualdad, tienen como objetivo impulsar, más que ejecutar directamente las políticas de igualdad de género. De manera que las femócratas a la cabeza de dichos organismos de igualdad deben convencer o persuadir a otros departamentos y poderes públicos para que asuman y ejecuten unos objetivos que inicialmente no son suyos. Esta tarea no es fácil y es el gran caballo de batalla, tanto a niveles políticos como técnicos de las femócratas y los organismos de igualdad. Es más, los organismos de igualdad no tienen poder para sancionar el incumplimiento de los objetivos y acciones propuestas en un plan. Esto significa que, en definitiva, los planes de igualdad necesitan ser reforzados con estrategias de persuasión y convencimiento por parte de las femócratas y sobre las que no se ha teorizado en la literatura sobre políticas públicas.

En este caso, es interesante hacer un breve análisis sobre las estrategias utilizadas por las femócratas para persuadir al resto de las instancias gubernamentales de que incorporen la perspectiva de género y hagan suyos los objetivos propuestos. Entre las estrategias que mencionan las responsables de los organismos de igualdad entrevistadas para esta investigación, están las de “ir con los datos por delante”, “tener mucha paciencia para explicar y repetir las cosas muchas veces” y “no enfrentarse nunca directamente”⁴⁷. Efectivamente, además de la obtención de datos que avalen la necesidad de actuación para conseguir una mayor igualdad, las otras estrategias mencionadas dependen más de la posición de liderazgo de la responsable y de su capacidad personal de persuasión y persistencia, y esto, lógicamente, se convierte en un punto especialmente débil si se dan cambios en las responsables. Aunque excede claramente los objetivos de este trabajo, quedaría pendiente investigar a fondo sobre las estrategias empleadas. Esto permitiría, además de evaluar su efectividad, establecer y “protocolizar” las estrategias más adecuadas para la introducción de la transversalidad de género en todas las políticas y programas.

Otra posible debilidad de los planes de igualdad es que pueden llevar al **mimetismo**. En los primeros pasos de elaboración de los planes, una de las primeras tareas de las técnicas de los organismos de igualdad es mirar y tener en cuenta otros planes de igualdad

⁴⁷ En el capítulo ocho se analiza pormenorizadamente el discurso extraído de las entrevistas realizadas. Para conocer las entrevistas realizadas, véase el anexo correspondiente.

(sobre todo del mismo nivel gubernamental). Los planes se construyen, por lo tanto, con una primera base de lo ya formulado en planes similares; aunque se hagan adaptaciones posteriores y se tengan en cuenta las características propias de la población diana, esa primera base es importante. Probablemente esta sea una de las razones por las que los planes se parezcan bastante entre sí, al menos desde el punto de vista formal. Esto, que no tiene por qué ser una debilidad en sí misma, puede sin embargo llevar a reproducir rutinas y, sobre todo, tender hacia una falta de innovación. Más allá, los planes estereotipados pueden hacer más difícil la reacción ante nuevas contingencias no contempladas en planes formulados de una forma rígida.

La tercera debilidad tiene relación con la **participación de actores** en los planes de igualdad de género. Aunque en los procesos de elaboración de los planes, algunos organismos de igualdad han tratado de implicar a diferentes grupos de interés, incluidos grupos de mujeres (entre otras cosas, para conseguir legitimación de dicho plan), los planes de igualdad de género en España están claramente orientados hacia los poderes públicos y escorados hacia lo gubernamental. En otras palabras, aunque el modelo teórico esté basado en la participación de la sociedad civil, el modelo en la práctica no lo está. Y unas políticas que están avaladas y apoyadas activamente por los movimientos no gubernamentales están siempre más fortalecidas. Esta tendencia hacia lo gubernamental es, hasta cierto punto inevitable, ya que los organismos de igualdad han tenido que dedicarle especiales esfuerzos a una nueva estrategia de horizontalidad que ha requerido y está requiriendo muchas energías en dicho nivel gubernamental. Por ahora, las femócratas parecen tener una tendencia a estar especialmente concentradas en obtener la participación y colaboración de sus propios colegas en los diferentes departamentos públicos, y no tanto en la participación de otros grupos no gubernamentales.

Por último, como ya se ha mencionado antes, otro peligro de los planes es quedarse en el terreno simbólico. Las políticas de igualdad de género tienen el peligro de ser políticas simbólicas, debido a la incorporación de la retórica sobre la igualdad de género en los discursos políticos (Meehan y Devenhuijsen, 1991) y debido a que lo “políticamente correcto” incluye el respeto por dicha igualdad. Aún reconociendo la importancia de las políticas simbólicas (Valiente, 1998) y una cierta evolución de unas políticas de igualdad de género simbólicas iniciales a ciertas políticas sustantivas a partir de los años 60 (Martínez, 1997: 212), existe sin embargo una tendencia hacia lo simbólico y lo retórico en el discurso de la igualdad de género. Los planes de igualdad pueden quedarse en un **plano meramente formal** y, precisamente, servir como "cortina de humo" y como instrumento para legitimar políticas públicas de puro carácter simbólico. No es raro que se justifique como parte de las actuaciones de un plan de igualdad, por ejemplo, la contabilización de las mujeres beneficiarias de unos cursos de búsqueda de empleo de otro departamento, ministerio o consejería, cuando ni esas actividades formativas se han planificado con perspectiva de género, ni el análisis de la realidad previo se ha hecho teniendo en cuenta realmente a las mujeres y cuando dichas actuaciones se hubieran hecho en cualquier caso, existiendo o no el Plan de Igualdad. Esta es una de las razones por las que la evaluación de las políticas y los planes de igualdad, una evaluación consciente y desde diferentes puntos de vista, resulta crucial en estos momentos en los que se cumple una década de utilización

de los planes de igualdad como instrumento principal de las políticas públicas de igualdad de género.

CAPÍTULO 4

LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO EN ESPAÑA. MARCO PARA EL ANÁLISIS Y LA VALORACIÓN

INTRODUCCIÓN

En este cuarto capítulo se ofrece una primera visión general de las experiencias de evaluación en torno a las políticas de igualdad de género en España en los niveles de gobierno autonómico y central. También se fundamenta y se expone el marco para su análisis y valoración, la metaevaluación, que permitirá la exposición estructurada de los datos en los capítulos siguientes (capítulos cinco, seis y siete).

En primer lugar, se fundamenta la metaevaluación como el enfoque metodológico en el que se basa el análisis y la valoración de los procesos de evaluación que se realiza en este trabajo y que se expone en esta parte III de la tesis doctoral. En la literatura al uso no se distingue claramente el concepto de metaevaluación del de síntesis de evaluaciones y el de meta-análisis; en este apartado se realiza una distinción de dichos términos, además de señalar las diferentes funciones atribuibles a la metaevaluación, destacando las de descripción y análisis de procesos de evaluación y la de control de calidad de los mismos. Al ser una evaluación de evaluaciones, se pueden también distinguir diferentes tipos de metaevaluación y se deben definir, asimismo, criterios de metaevaluación. Una vez discutidos estos extremos, se establecen los parámetros concretos bajo los que se realizó la metaevaluación presentada en esta tesis doctoral. Finalmente, en el primer apartado de este capítulo se señala y desarrolla la necesidad y conveniencia de analizar y valorar los procesos de evaluación de las políticas públicas de igualdad de género en España en los niveles autonómico y central.

En segundo lugar, se ofrece un panorama general de las evaluaciones realizadas en torno a las políticas de igualdad de género en España en los niveles autonómico y central. Ateniéndose al contexto temporal, todas las experiencias de evaluación existentes en torno a los planes de igualdad de género se clasifican en una primera, una segunda y una tercera generación. El objeto de la metaevaluación de este trabajo es la segunda generación, que queda delimitada en este apartado del capítulo como el conjunto de evaluaciones realizadas entre los años 1995 y 1999. Las primeras experiencias de evaluación fueron pocas, limitadas y están ya lejanas en el tiempo. Por el contrario, la tercera generación de evaluaciones está, en su mayoría, sin desarrollar y concluir en el momento de elaboración de este trabajo, por lo que se descarta necesariamente como objeto de esta tesis doctoral. Sin embargo, en el panorama general que se presenta en este capítulo, se ha querido exponer de forma somera, lo que representó la primera generación de evaluaciones y lo que significan los proyectos de evaluación de la tercera. De esta forma, la segunda generación de evaluaciones –

consistente en diez experiencias de ámbito autonómico y una de ámbito central- queda adecuadamente enmarcada en el contexto histórico en el que se produce.

En tercer y último lugar, se expone el marco para el análisis y valoración de la segunda generación de evaluaciones. Así, se exponen y fundamentan los criterios utilizados en esta metaevaluación. Estos criterios, que son definidos en el tercer apartado de este capítulo, se dividen en tres grandes apartados:

1) El planteamiento de las evaluaciones, teniendo en cuenta el contexto, la claridad de los propósitos y objetivos de las evaluaciones, la existencia de estructuras institucionales y de recursos para la evaluación, además de la utilización de diferentes tipos de evaluación.

2) Los elementos clave de la evaluación, siendo éstos los agentes y participantes en la misma, los momentos y el calendario, los criterios de evaluación establecidos y la metodología o procedimientos utilizados.

3) La utilización de los procesos de evaluación, teniendo en cuenta la pertinencia y utilidad de la información producida, la comunicación y difusión de los resultados y los propios procesos, así como el impacto y la utilización de los mismos en las políticas y los organismos de igualdad de género.

1. FUNDAMENTACIÓN

1.1. La metaevaluación

El concepto de metaevaluación y su diferenciación con el concepto de síntesis de evaluaciones

La metaevaluación, entendida ésta como evaluación de evaluaciones, es la base en la que se sustenta el trabajo empírico de esta tesis doctoral. A continuación, se discute su definición, y se diferencia de otros conceptos como son la síntesis de evaluaciones y el meta-análisis. Uno de los primeros autores que desarrolla el concepto de metaevaluación es Michael Scriven, cuyo término aparece definido en su “Thesaurus de Evaluación” como

“la evaluación de evaluaciones -indirectamente, la evaluación de los evaluadores- representa una obligación tanto ética como científica cuando el bienestar de otros está en juego” (Scriven, 1991: 228).

El *Joint Committee on Standards for Educational Evaluation*¹ (a partir de ahora *Joint Committee*) desarrolla la metaevaluación como una de las normas de precisión que deben cumplir las evaluaciones. La norma A12 Metaevaluación dice:

“La evaluación en sí misma debe ser formativa y sumativamente evaluada según éstas y otras normas pertinentes, de manera que su realización se oriente de forma apropiada y, en cuanto se termine, los agentes críticos (stakeholders) puedan examinar de cerca sus fortalezas y debilidades” (Joint Committee, 1994: 185).

Para Patton (1997: 143) la metaevaluación es una “evaluación de la evaluación basada en una serie de normas y principios profesionales”.

Sin embargo, para algunos autores, el concepto de metaevaluación puede incluir la síntesis de evaluaciones, e incluso ser la acepción principal, como es el caso de la clasificación realizada por la *Evaluation Research Society* (Evaluation Research Society

¹ Como se señaló en el capítulo 1, el *Joint Committee on Standards for Educational Evaluation* comenzó a reunirse al inicio de los años ochenta publicando la primera edición de sus normas en el año 1981. Dicho Comité está integrado por más de doscientos profesionales en el campo de evaluación quienes se reúnen periódicamente para revisar las normas. Aunque este Comité se inicia con el ámbito educativo de fondo se convierte rápidamente en referencia en las restantes disciplinas que se ocupan de la evaluación de programas y políticas (Bustelo, 1998).

Standards Committee, 1982), y en la que uno de los 6 tipos de evaluación descritos es la metaevaluación. En esta clasificación, a la metaevaluación se le otorga el significado de síntesis de evaluaciones, es decir, una forma de evaluación que reanaliza los resultados de una o varias evaluaciones para determinar qué es lo que se ha aprendido sobre una política o programa (Chelimsky, 1989). Para Ballart, la metaevaluación “consiste en el análisis de los resultados de varias evaluaciones”, aunque también habla de la “revisión de tipo más amplio de la investigación evaluativa desarrollada en un determinado período” (Ballart, 1992: 81-82). En la primera acepción -análisis de los resultados arrojados por distintas evaluaciones-, más estricta, el autor es consciente de que su desarrollo tiene grandes complicaciones debido a la heterogeneidad que normalmente se encuentra en los estudios de evaluación a sintetizar. Para Vedung (1997), la metaevaluación significa la evaluación de metaprocesos de evaluación y retroinformación, aunque también considera metaevaluación al resumen de los hallazgos de diferentes estudios de evaluación. Weiss (1998: 236) comenta que, cuando los estudios analizados son evaluaciones, el meta-análisis se llama a veces metaevaluación. Sin embargo, reconoce que dicho término, metaevaluación, se debe reservar para otro concepto diferente: la evaluación de evaluaciones, tal y como se ha definido en este trabajo.

En este trabajo se propone la distinción del concepto de metaevaluación y los de síntesis de evaluaciones y meta-análisis (como un tipo específico de síntesis de evaluaciones o investigaciones) ya que, en nuestra opinión, son procesos de naturaleza muy distinta.

La **síntesis de resultados** sirve para aunar, resumir e incluso potenciar los resultados de varios estudios de evaluación de programas o políticas similares. En realidad es un procedimiento para revisar estudios similares anteriores. El énfasis aquí se pone en hacer comparables los resultados de varios estudios: el foco está en dichos resultados, en los datos que han arrojado los estudios de evaluación. Se interesará en la metodología seguida, pero simplemente como camino por el que se ha llegado a dichos resultados. La función última no será juzgar los procesos de evaluación sino tener una base empírica para juzgar los resultados generales de los programas o políticas evaluadas.

Un tipo especial de síntesis es el **meta-análisis** que consiste específicamente en un procedimiento estadístico que permite la comparación entre resultados de diferentes estudios; es decir, es un procedimiento particular para sintetizar estudios cuantitativos² que tengan un mismo objeto de estudio, en el que se calcula un parámetro especial (llamado el “tamaño del efecto”) para cada estudio. Para Weiss, el meta-análisis, aplicado a la evaluación, puede ser definido como “el análisis sistemático de los resultados de un conjunto de evaluaciones de programas similares para producir una estimación del efecto general de los programas” (Weiss, 1998: 333). La técnica del meta-análisis no se ha utilizado sólo en el campo de la evaluación de programas, sino para sintetizar diferentes investigaciones en un mismo campo. En España se han realizado estudios de meta-análisis en el campo de la psicología, sobre efectos de

² No existe una metodología o procedimiento especial para “metanalizar” o sintetizar estudios de tipo cualitativo. Dado que una gran cantidad de estudios de evaluación combina metodologías cualitativas y cuantitativas, la utilización concreta del metanálisis en el campo concreto de la evaluación de programas es, de alguna forma, limitada.

diferentes terapias y de estudios que relacionaban memoria y estado de ánimo (Navarro, 1990; Matt, Vázquez y Campbell, 1992). Es más, los estudios de meta-análisis, por sus propias características, podrían ser clasificados más en el campo de la investigación que en el de la evaluación de programas propiamente dicha. El meta-análisis en España en el campo de la evaluación de políticas públicas parece estar lejos de ser posible, debido a la falta de estudios sistemáticos que pudieran ser meta-analizados en la mayoría de los sectores de las políticas públicas.

Sin embargo, la **metaevaluación** consiste en hacer evaluación de los procesos de evaluación. El foco aquí se encuentra en los procesos de evaluación, no en los resultados arrojados por los mismos. Por ejemplo, los procesos de metaevaluación se interesarán más en los resultados de la evaluación no en sí mismos, sino en si son buenos (es decir si se ha recopilado la información y datos suficientes, y si las conclusiones interpretaciones y valoraciones están bien fundamentadas), si han resultado útiles o si han sido realmente utilizados para la mejora de las políticas, el control de responsabilidades y/o la iluminación de acciones futuras. Es decir, el foco de las metaevaluaciones es en el cómo se han realizado las evaluaciones y en la valoración de dichos procesos, y no en los resultados que han arrojado.

De esta manera, en este trabajo se distinguen estos tres conceptos. La metaevaluación, que se define como evaluación de evaluaciones; la síntesis de evaluaciones, que son análisis comparados de los resultados de varias evaluaciones y el meta-análisis, que se reservaría para las síntesis de evaluaciones que utilizan específicamente la técnica cuantitativa de dicho tipo de estudios.

En los últimos tiempos se ha insistido mucho en la necesidad de aprovechar los estudios ya realizados a través de la síntesis de evaluaciones y el meta-análisis (Chelimsky, 1997; Cook, 1997). Esto se ha propuesto, sobre todo, desde los países en los que la función y la cultura de evaluación están muy desarrolladas (especialmente Estados Unidos y Canadá y, en menor medida, Australia y algunos países europeos como Gran Bretaña y Suecia). Para los países, como España, en los que la función de evaluación está todavía por desarrollar, parece aún lejana la posibilidad de sintetizar estudios de evaluación. Primero tiene que darse el hecho de que existan suficientes evaluaciones, es decir, las energías deben concentrarse en que se realicen estudios de evaluación que, posteriormente, podrían ser metaevaluados.

Mientras el meta-análisis y la síntesis de evaluaciones están más ligadas al concepto de revisión del estado de la cuestión y a la necesidad de un mayor aprovechamiento del saber acumulado, es decir, están más vinculadas a finalidades de la investigación, la metaevaluación se vincula a las finalidades de asegurar y evaluar la calidad de los procesos de evaluación y a la comprensión de la función de evaluación en el proceso de las políticas públicas.

El significado y las funciones de la metaevaluación

La metaevaluación tiene un primer significado como **descripción, análisis y valoración de los procesos de evaluación**. En este sentido, se inscribiría en el componente descriptivo del análisis de políticas públicas, en el “*knowledge of*” acuñado por Lasswell, quien establece la distinción entre dicho conocimiento del proceso de la política (*knowledge of*) y el conocimiento en o para el proceso de la política (*knowledge in*). En el primero (conocimiento del o *análisis de la política*) se pretende conocer el proceso de decisión de la política así como de hecho sucede. Si tomamos como referencia las políticas públicas en su ciclo completo, se trataría de describir por qué determinados problemas son calificados como públicos, entran en agenda, se definen, cómo y por qué se eligen ciertos cursos de acción y ejecución y se evalúan a través de ciertos criterios y metodologías de evaluación. La parte que se propone como descripción de los procesos de evaluación de las políticas públicas correspondería a uno de los significados y funciones de la metaevaluación. El autor que se acerca más a esta concepción, en el campo de la evaluación de programas y políticas públicas, es Vedung, quien considera que una posibilidad de la metaevaluación es la de auditar la función de evaluación de un determinado organismo gubernamental (Vedung, 1997). Aunque este componente descriptivo se ha desarrollado en el campo del análisis de políticas públicas en general, no ha ocurrido lo mismo en el aspecto o fase concreta de la evaluación de dichas políticas. Es decir, no existen muchos estudios sobre los procesos de evaluación de las políticas públicas, en los cuáles se estudie el desarrollo de la función de evaluación en torno a una determinada política y que describan, analicen y valoren las evaluaciones que se llevan a cabo.

El segundo significado de la metaevaluación es su función como **control de calidad de las evaluaciones**. Para Schwandt y Halpern (1988: 14), la metaevaluación se inscribiría en una serie de aproximaciones sistemáticas para el seguimiento, fomento y comprobación de la calidad de un servicio profesional como es, en este caso, la evaluación. El interés por la metaevaluación surge, en cierto modo, cuando se da una preocupación por la calidad de las evaluaciones y la necesidad de control de las mismas. En realidad tiene que ver con lo que Scriven define como la característica autoreferenciada o reflexiva que debe tener la evaluación, es decir, con la idea de “la evaluación empieza en casa” y con la de “¿quién evalúa al evaluador?” (Scriven, 1993: 51-52).

Como ya se ha comentado anteriormente, en el campo de la evaluación de programas el concepto de metaevaluación se ha desarrollado, sobre todo, en este segundo sentido. Se trata de un concepto principalmente técnico de la evaluación, del control técnico de las mismas. Incluso teniendo en cuenta que el concepto de calidad general de una evaluación incluye normalmente una serie de criterios no tan estrictamente técnicos (como son su utilización), el énfasis se pone en la utilización de los procesos de metaevaluación como una forma de control de calidad de las evaluaciones y no tanto como una forma de estudio de dichos procesos. Para Patton, la metaevaluación se pregunta sobre si se hizo bien la evaluación, si ha merecido la pena y se basó en normas y principios profesionales (Patton, 1997: 193). De hecho, en esta función de metaevaluación como control de calidad de las evaluaciones, existe un cierto

sesgo hacia metaevaluar una única evaluación. Es decir, se metaevalúa cada estudio de evaluación por separado; por lo tanto, cada evaluación, individualmente, debe ser a su vez evaluada. De esta manera, la idea de metaevaluar varios procesos de evaluación simultáneamente, y que esto ofrezca una base para la comparación (es decir, la idea de síntesis de evaluaciones, pero no de sus resultados sino de los procesos de evaluación) no se deja entrever en la literatura que ha desarrollado el concepto de metaevaluación como control de calidad de las evaluaciones. En realidad, este desarrollo del concepto de metaevaluación en su acepción más aplicada (para controlar la calidad de los procesos y no tanto para estudiar dichos procesos) resulta acorde con la propia definición de evaluación, diferenciándose de otros procesos de análisis e investigación: las evaluaciones se dan siempre en escenarios de acción, en los que el programa o política a evaluar es lo prioritario. Por lo tanto, la mejora y conocimiento concreto de dichos programas y el control de responsabilidades en torno a los mismos será siempre prioritario sobre la construcción de conocimiento -aunque éste verse sobre los propios procesos de evaluación-.

La metaevaluación es, por lo tanto, algo añadido a los procesos de evaluación, aunque para muchos autores algo absolutamente necesario. Patton señala la necesidad de metaevaluación aunque, “desde una perspectiva práctica, sin embargo, no todo esfuerzo de evaluación merece los recursos necesarios para una metaevaluación completa” (Patton, 1997: 352). Es decir, al igual que la evaluación, la metaevaluación debe tener su sentido y su medida. Sin embargo Scriven, quien es uno de los primeros autores que acuña y desarrolla el concepto de metaevaluación, es más contundente, y una de las 31 tesis de su obra “Lecciones duramente aprendidas en evaluación de programas” es la que dice que “los diseños de evaluación sin provisión de evaluación de la evaluación son poco claros en concepto” (Scriven, 1993: 51). Para Scriven, la evaluación debe ser reflexiva y, por lo tanto, debe ser evaluada; es decir, los evaluadores deben desarrollar los dispositivos necesarios para, a su vez, ser evaluados y esto deben hacerlo desde un principio, desde el mismo diseño de la evaluación. En realidad, en el concepto de metaevaluación de Scriven hay un especial énfasis en la idea de ser una forma de **control del sesgo del evaluador** (Shadish et al, 1991: 81-82).

Esta concepción de la metaevaluación como una forma de **control de calidad** no sólo pretende dicho control sino que también puede ser una forma de asegurar la **credibilidad** de una evaluación, especialmente si se dan procesos de metaevaluación externa. Las auditorías de evaluación y las metaevaluaciones aseguran la credibilidad de la evaluación para sus usuarios, “del mismo modo en el que las auditorías financieras independientes aseguran la credibilidad de los informes de beneficios a los inversores” (Patton, 1997: 352). Precisamente, este paralelismo entre el concepto de auditoría y el de metaevaluación se trata ampliamente en la obra de Schwandt y Halpern *Linking Auditing and Metaevaluation* del año 1988. Esta necesidad de credibilidad de la evaluación anteriormente mencionada, es mayor cuanto más politizado sea el contexto en el que se producen los procesos de evaluación (Chelimsky, 1987), por lo que “a mayor politización del contexto en el que se realiza una evaluación y a mayor visibilidad de la evaluación, mayor importancia de una valoración independiente sobre la calidad de la evaluación para su credibilidad” (Patton, 1997: 352). Este concepto de la metaevaluación como forma de aumentar la credibilidad de las evaluaciones está de alguna manera relacionada con la idea de Scriven de control del sesgo del evaluador. De hecho, reconocer que el evaluador puede tener un cierto sesgo, pero que hay un forma

de controlarlo a través de mecanismos de metaevaluación, significa también asegurar la credibilidad de los estudios de evaluación.

Efectivamente, con esta idea de control de calidad de fondo, Schwandt y Halpern (1988) desarrollan la metaevaluación como metáfora de la **auditoría** fiscal. Así, la metaevaluación tiene que ver con la acreditación y la comprobación de la calidad. Los autores creen que el desarrollo de este tipo de metáfora fuerza, además, a la autodisciplina de los evaluadores. Es decir, el hecho de que existan metaevaluaciones y se auditen los procesos de evaluación fomenta, en cierto modo, la calidad de las mismas aunque, por supuesto, no la crea o genera. Dichos autores señalan que en los últimos años los profesionales y académicos de la evaluación se han preocupado crecientemente de la calidad de los planes, los procedimientos y los hallazgos de las evaluaciones. Esta preocupación se refleja en diseñar evaluaciones sensibles a las necesidades de los críticos, en incorporar a los planes de evaluación procedimientos que faciliten el uso de los resultados, en establecer normas que sirvan como guías y baremos para la práctica profesional de la evaluación, y en la comprobación de la calidad de las evaluaciones a través de la metaevaluación (Schwandt y Halpern, 1988: 17).

En cualquier caso, la metaevaluación se convierte en una norma del *Joint Committee* a partir de su segunda edición (*Joint Committee*, 1994). Esto significa que en la revisión de las normas que se deben cumplir a la hora de hacer evaluaciones, la metaevaluación aparece como una obligación de todo proceso evaluativo. Su reconocimiento teórico, sin embargo, no se corresponde totalmente con la práctica evaluativa o, al menos, existen pocas referencias bibliográficas sobre metaevaluaciones realizadas.

Tipos de metaevaluación

La metaevaluación es, en realidad, una evaluación que tiene por objeto de estudio, en lugar de un programa o política pública, la evaluación de un programa o política. Esto quiere decir que puede ser tratada y, por lo tanto, clasificada, como cualquier otra evaluación.

Atendiendo, pues, a las cuatro clasificaciones básicas ya mencionadas en el capítulo 2 de este trabajo, la primera distinción que podría realizarse es según el papel que cumple la (meta)evaluación. En este caso la evaluación puede ser **formativa**, de manera que la metaevaluación se realiza sobre la marcha y con la intención de mejorar el diseño y ejecución de la evaluación, o **sumativa** haciendo recapitulación de cómo se ha realizado un proceso de evaluación para ofrecer una valoración global sobre el mismo. De hecho, la norma de “metaevaluación” elaborada por el *Joint Committee* menciona explícitamente que las evaluaciones deben ser formativa y sumativamente metaevaluadas, lo que ofrece un reconocimiento de ambos tipos de metaevaluación. Scriven en su *Thesaurus* también señala que la metaevaluación se puede hacer formativa, sumativamente o de ambas maneras; el autor hace especial hincapié en la

metaevaluación sumativa ya que ofrece al cliente información sobre la competencia técnica del evaluador primario; incluso si no es pedida por el cliente, Scriven cree que el evaluador primario debe considerar la metaevaluación sumativa como otro servicio más al cliente y que siempre mejora su evaluación (Scriven, 1991: 230).

En relación con esta primera clasificación, se puede también establecer otra clasificación básica, la que se hace respecto al momento en el que se realiza la (meta)evaluación. De esta manera se pueden realizar metaevaluaciones **ex-ante** (antes de que el proceso de evaluación se ejecute) o **post-facto** (durante el proceso de evaluación o una vez finalizado). Schwandt y Halpern (1988: 8) distinguen, precisamente entre *provaluation* (si la metaevaluación se lleva a cabo antes de que se realice la evaluación), *retrovaluation* (si se hace una vez que ya se ha realizado) y, por último, el seguimiento o monitorización del proceso de evaluación (durante dicho proceso).

La tercera clasificación tiene que ver con el agente que realiza la evaluación (en este caso metaevaluación). Así, pueden existir las metaevaluaciones **internas**, que son las que se llevan a cabo por parte de los mismos evaluadores; es decir, quienes realizan la evaluación y también la metaevalúan, o las metaevaluaciones **externas** que son las que se realizan por alguien ajeno al proceso de evaluación metaevaluado. En realidad estos dos tipos de metaevaluación son bastante diferentes ya que la interna significa que alguien interno al equipo de evaluación lleva el control de calidad de la misma (existen varias listas de comprobación que se proponen para este control y que se verán en el punto siguiente). Sin embargo, la metaevaluación externa significa un reconocimiento y, en cierto modo, la institucionalización de esta actividad, ya que implica recursos adicionales a los procesos de evaluación regulares. De todas formas, pueden existir diversas combinaciones y Patton (1997) pone precisamente como ejemplo la utilización de la metaevaluación como forma de combinar evaluación interna y externa: se trataría de hacer una evaluación interna y tener a alguien que, externamente, revisa y comprueba el proceso de evaluación. Esto resultaría en una evaluación que, aún siendo interna, goza de la credibilidad de las evaluaciones externas.

Por último, se aplica la clasificación según el contenido o fase del proceso de evaluación que se metaevalúa. Según esta clasificación, se puede hablar de (meta) evaluación del **diseño**, del **proceso** y de los **resultados**. En este caso consiste en la evaluación del diseño, el proceso y los resultados de la evaluación y no del programa evaluado.

Así la metaevaluación de **diseño** se ocuparía de estudiar y valorar el diseño del estudio y del proceso de evaluación, es decir, si está bien fundamentado, se ha planteado y responde a las preguntas de PARA QUE (qué pretende la evaluación) y PARA QUIENES (quiénes son los destinatarios) de la evaluación. Aquí se incluiría el estudio y análisis del contexto evaluativo y si el estudio de evaluación ha respondido a dicho contexto; es decir, si se han contemplado las diferentes variables de la situación evaluativa y se han tenido en cuenta en el diseño de la evaluación. También si es realista y tiene medida respecto a lo que significa el programa que evalúa y los resultados que se esperan obtener (su utilidad potencial y su posible contribución a la mejora del programa, el control de responsabilidades y la iluminación de acciones futuras); si es realista respecto al calendario planteado y los costes el estudio, si está prevista la

participación de los agentes críticos y procedimientos para fomentar la utilidad potencial del estudio, si las técnicas escogidas son suficientes y adecuadas para la recopilación del análisis, si los instrumentos de recopilación y análisis y los sistemas de información son válidos y si se cuenta con la documentación necesaria, si se ha previsto un análisis adecuado y con suficiente tiempo y si se ha diseñado un plan de informes y difusión de la información.

La metaevaluación de **proceso** analiza el proceso de puesta en práctica de la evaluación. Este tipo de metaevaluación tiene que ver con si se ejecuta la evaluación como se pretendía, qué dificultades se encuentran sobre la marcha, fortalezas y debilidades del proceso, el ajuste al calendario previsto, si se consigue la participación de los agentes críticos y si se llega a los informantes y si es fiable la información que se recopila en el proceso. También tiene que ver con el hecho de si se aplica la metodología y las técnicas previstas de forma adecuada y si se utilizan los recursos para el estudio de forma eficiente. Finalmente, con este tipo de evaluación cabe también preguntarse sobre el equipo de evaluadores y sobre si éste funciona de forma adecuada y ética.

Por último una metaevaluación de **resultados** debe preguntarse por los *outputs* (productos) y *outcomes* (efectos) de la evaluación o evaluaciones meta-evaluadas. En este caso, el producto por excelencia de un proceso de evaluación son los resultados e información (datos, interpretaciones, valoraciones y recomendaciones), que típicamente se reflejan en informes de evaluación, pero que también pueden expresarse en otro tipo de soporte (sesiones orales, presentaciones, etc.) . Por lo tanto, una metaevaluación de *outputs* analizará la calidad de los resultados e información que produce y cómo éstos se reflejan en los informes u otros soportes, si el informe final se entregó a tiempo y si se difundió a todas las partes responsables de la evaluación y si incluye una descripción del evaluando, su contexto, el propósito de la evaluación, los procedimientos empleados, los resultados, etc. La metaevaluación de *outcomes*, sin embargo, tiene que ver con los efectos que la evaluación consigue: aprendizaje sobre el programa e intervención evaluado, modificaciones en dicho programa como consecuencia de la evaluación, control de responsabilidades y aprendizaje organizativo en general, así como contribución a la comunidad de programa o política evaluada. En definitiva, los efectos de una evaluación tienen que ver con su utilidad y utilización.

En la tabla 4.1 se expone un resumen de las cuatro tipologías básicas de evaluación aplicadas a los diferentes tipos de metaevaluación.

Tabla 4.1. TIPOS DE META EVALUACION
Clasificación según las cuatro tipologías básicas

SEGUN EL PAPEL DE LA META EVALUACION

Formativa: Se hace sobre la marcha con la intención de mejorar el diseño y la ejecución del proceso de evaluación.

Sumativa: Se hace una recapitulación del proceso de evaluación para ofrecer una valoración global y final del mismo.

SEGUN EL MOMENTO DE LA META EVALUACION

Ex-ante (*provaluation*): Se realiza antes de que el proceso de evaluación se ejecute.

Ex-postfacto (*retrovaluation*): Se realiza durante o una vez el proceso de evaluación ya se ha ejecutado.

SEGUN EL AGENTE META EVALUADOR

Interna: Se llevan a cabo por parte de los mismos evaluadores, siendo en realidad una (auto)metaevaluación. Los propios evaluadores se metaevalúan, tratándose de un control interno.

Externa: Se realizan por parte de alguien ajeno al proceso de evaluación. Se trata de un control externo.

SEGUN EL CONTENIDO O FASE DEL PROCESO DE EVALUACION

Metaevaluación de Diseño: Analiza y valora el diseño y conceptualización del estudio de evaluación.

Metaevaluación de Proceso: Analiza y valora el proceso de puesta en práctica de la evaluación, es decir, cómo se ejecuta proceso.

Metaevaluación de Resultados:

De *outputs* o productos: Analiza y valora la calidad de los resultados y la información arrojada por la evaluación y su reflejo en el informe.

De *outcomes* o efectos: Analiza y valora los efectos de la evaluación, incluyendo su utilización para la toma de decisiones y el aprendizaje organizativo.

Criterios para la metaevaluación

Schwandt y Halpern (1988) proponen la metaevaluación como metáfora de la auditoría fiscal y como forma de fomentar la calidad de la evaluación. Estos autores señalan dos objetos generales de dicha metaevaluación: analizar y valorar, por un lado, si la metodología utilizada es la adecuada y, por otro, si los resultados a los que se llega son buenos. Es decir, para los autores los criterios de calidad de una evaluación y que serían, por lo tanto, criterios para la metaevaluación serían metodología adecuada y buenos resultados. Avanzando en el análisis, resulta interesante preguntarse entonces sobre el significado de cada uno de estos extremos.

La utilización de una **metodología adecuada** tiene que ver con los procedimientos empleados para conocer el objeto de estudio. En primer lugar, debe existir una cierta coherencia entre los niveles epistemológico, metodológico y tecnológico y, sobre todo, debe responder ante un determinado contexto y resultar adecuada para el estudio del objeto a evaluar (programa o política), es decir la metodología escogida debe permitir e ir dirigida a conseguir una evaluación que sea relevante al contexto y programa a evaluar. Una cuestión clave es que se hayan establecido previamente los criterios de evaluación, habiéndose descrito la base lógica sobre al que se harán los juicios sistemáticos posteriores. También debe existir una coherencia técnica entre el diseño para el establecimiento de criterios de evaluación, el diseño para la recopilación y análisis de la información y las técnicas concretas empleadas para dicha recopilación y análisis. Y además esto debe hacerse con una adecuada relación coste-beneficio (es decir, la evaluación debe tener su medida y debe merecer la pena el esfuerzo y recursos empleados en relación a la información y resultados que se obtienen) y teniendo en cuenta el tiempo y el calendario, factor clave en los procesos evaluativos. Las técnicas deben ser válidas y permitir una recopilación fiable de información. También, la metodología empleada debe permitir la participación de los diferentes agentes críticos en el programa, especialmente si se sigue un enfoque participativo.

También se puede analizar en qué consisten unos **buenos resultados** en los procesos de evaluación. Esencialmente en evaluación, y tomando como referencia los cuatro niveles del análisis valorativo establecidos por Patton, se trataría de conseguir una coherencia lógica y una clara separación de esos cuatro niveles: hallazgos o resultados (datos, información), interpretaciones, valoraciones y recomendaciones. Las interpretaciones o explicaciones deben realizarse basándose en unos datos fiables y recopilados sistemáticamente. Dichas interpretaciones son diferentes a las valoraciones y les sirven precisamente como base. Por último, las recomendaciones deben basarse en las valoraciones que, a su vez, se basarán en las interpretaciones que, a su vez, estarán basadas en los hallazgos. Profundizando algo más, ¿qué criterios se pueden establecer para saber si unos **hallazgos** (los datos, la información, la evidencia que se recopila) son buenos? Schwandt y Halpern³ establecen una serie de atributos deseables:

³ Schwandt y Halpern exponen una tabla sobre los atributos que debe tener la evidencia (información, datos) que se recoge en una auditoría. Estos atributos pueden servir como referencia general sobre características deseables o criterios de bondad para la información o datos que se recopilen en cualquier estudio de evaluación. (Schwandt y Halpern, 1988: 34, tabla 1.3, adaptada de Edds, J.A. 1980. *Management auditing: Concepts and practices*. Dubuque, IA: Kendall Hunt).

Relevancia: la información y datos recopilados deben estar directamente relacionados con los objetivos de la evaluación.

Fiabilidad: que la información y los datos sean fiables significa que sean verdaderos y confiables.

Suficiencia: tiene que haber suficiente información para que se puedan establecer conclusiones dentro de unos límites razonables de riesgo.

Representatividad: La información y los datos deben ser razonablemente representativos de la totalidad de la información bajo examen.

Temporalidad: se deben sopesar los factores de costes y tiempo y la necesidad de entregar una evaluación válida en un tiempo determinado.

Si la información y los datos son buenos (relevantes, fiables, suficientes, representativos y están conseguidos a tiempo), las **interpretaciones** y explicaciones de esos datos serán también adecuadas, si el análisis es correcto y suficiente. Los criterios y normas para el análisis están normalmente asociados a cada una de las técnicas empleadas. Las **valoraciones** deben hacerse de acuerdo con los criterios de evaluación previamente establecidos, definidos y consensuados con los diferentes agentes críticos, habiéndose establecido la base lógica sobre la que se van a emitir los juicios. En los criterios de evaluación deben también reflejarse las normas absolutas (que hacen referencia a lo que se considera bueno en términos absolutos) y las normas relativas (que hacen referencia a lo que es bueno en comparación con otras realidades similares concretas) que guiarán la emisión de juicios. Normalmente las valoraciones que se realizan en los procesos de evaluación se hacen teniendo en cuenta una mezcla de estos dos tipos de normas; sin embargo, conviene clarificarlo, exponerlo previamente y hacer el proceso visible por parte de los evaluadores.

Unas buenas **recomendaciones** además de basarse lógicamente en los tres niveles anteriores, deben tener muy presentes la realidad del evaluando y su contexto. Deben ser realistas y claras y el equipo evaluador tiene que permitirse el suficiente tiempo para establecerlas. También es deseable que las recomendaciones vayan acompañadas de una estimación de los costes que supondrían su cumplimiento, ya que está demostrado que las recomendaciones que van acompañadas de una valoración de los costes de su implementación tienen mayores probabilidades de ser utilizadas (Chelimsky, 1998).

Sin embargo, desde aquí postulamos la idea de que la metaevaluación puede también atender a otros criterios que vayan más allá de una metodología adecuada y unos buenos resultados. En general, la metaevaluación puede utilizarse para comprender los procesos de evaluación, ver qué papel han cumplido dichos procesos en la política pública evaluada y analizar y valorar cómo se ejecuta la función de evaluación en un determinado sector, política, organismo o institución. En definitiva, se trataría de entender los procesos políticos que dan lugar, intervienen en las evaluaciones y se generan en la recepción y utilización de las mismas. Mientras que en la concepción de la metaevaluación como forma de control de calidad de las evaluaciones se puede e incluso conviene concentrarse en metaevaluar un único estudio de evaluación, para esta segunda función de comprensión de los procesos de evaluación en el contexto de las

políticas públicas suele ser más conveniente la descripción, análisis y valoración de más de un proceso o estudio de evaluación.

A continuación, en la tabla 4.2 se hace un resumen de los criterios para la metaevaluación expuestos.

Tabla 4.2.
CRITERIOS PARA LA METAEVALUACION

Función de CONTROL DE CALIDAD

Utilización de una METODOLOGIA ADECUADA

- Coherencia entre los niveles epistemológico, metodológico y tecnológico
- Adecuación al contexto y al objeto a evaluar (relevancia del estudio)
- Establecimiento previo de los criterios de evaluación
- Coherencia técnica entre el diseño de criterios, la recopilación y el análisis de la información, así como en el enjuiciamiento posterior a la luz de los criterios
- Adecuada relación coste-beneficio del estudio de evaluación
- Adecuación al calendario disponible
- Validez y fiabilidad de las técnicas de recopilación y análisis de la información
- Participación de los agentes críticos en el proceso

Obtención de unos BUENOS RESULTADOS

- Coherencia entre hallazgos, interpretaciones, valoraciones y recomendaciones
- Buenos hallazgos, datos e información recopilada:
 - √ relevancia
 - √ fiabilidad
 - √ suficiencia
 - √ representatividad
 - √ temporalidad
- Buenas interpretaciones: análisis correcto y suficiente
- Buenas valoraciones: de acuerdo con los criterios de evaluación y normas relativas y absolutas, se deben extraer de la base empírica y de su interpretación
- Buenas recomendaciones: basadas en los hallazgos, interpretaciones y valoraciones previas, deben ser realistas y adecuadas al evaluando y a su contexto

Función de COMPRESION DE LOS PROCESOS DE EVALUACION (mejor comprensión y conocimiento de las políticas públicas)

- Papel que los procesos de evaluación han cumplido en la política pública
- Valoración de la función de evaluación: integración de dicha función en las políticas públicas, sectores, organismos e instituciones
- Utilidad de la evaluación para la mejora de las políticas y el control de responsabilidades

Otros autores mencionan criterios muy generales para la metaevaluación. Por ejemplo, para Vedung “el objetivo (de la metaevaluación) es averiguar la calidad, legibilidad, fidelidad a los hechos y otras propiedades” (Vedung, 1997: 41). Otros, como Michael Scriven, pionero en la definición del concepto, afina algo más y propone la aplicación de su famosa KEC (*Key evaluation checklist*) a los procesos de evaluación, insistiendo, al igual que en su “teoría” de evaluación (Shadish, Cook y Levinton, 1991), en la importancia a la explicitación previa de los criterios de valor utilizados por el metaevaluador para el enjuiciamiento de las evaluaciones. De esta manera se aplicaría la KEC (especificada a continuación en la tabla 4.3.) como instrumento de metaevaluación (Scriven, 1991: 230-231):

Tabla 4.3. Key Evaluation Checklist (KEC) de Michael Scriven

<p>1. DESCRIPCION. El diseño de la evaluación, el equipo evaluador, el calendario del estudio.. las fuentes de información utilizadas para la metaevaluación y la elección del metaevaluador.</p> <p>2. ANTECEDENTES Y CONTEXTO. ¿Quién ha demandado/pedido/se ha opuesto a esta evaluación? ¿Ha habido evaluaciones previas? ¿Por qué se hace otra? Se debe determinar los antecedentes de credibilidad de los enfoques de evaluación propuestos.</p> <p>3. CONSUMIDOR. Identificar los clientes y los destinatarios de la evaluación, de anteriores evaluaciones. ¿Hay destinatarios potenciales que podrían haber sido reales?</p> <p>4. RECURSOS. Recursos para la evaluación. Fuentes de las limitaciones en el diseño y, por lo tanto, en la evaluación del diseño y la ejecución.</p> <p>5. VALORES. Criterios de mérito para la evaluación, dadas las necesidades del cliente y las normas profesionales relevantes: temporalidad, relevancia, coste-efectividad, justicia, validez, cobertura de las necesidades apropiadas y relevantes del cliente, etc. Deben ser validados y no sólo señalados. ¿Existe alguna premisa evaluativa del evaluador evidente en la evaluación, o es lo suficientemente transparente al respecto?</p> <p>6. PROCESO. Se diseño y ejecutó bien la evaluación, por ejemplo, se hizo de manera que se facilitara su propia evaluación? Para comunicar y diseminar bien los resultados? ¿Respetó el evaluador la confidencialidad? ¿Fomentó la credibilidad? ¿Se han valorado los resultados de errores falsos negativos vs. falsos positivos? ¿Se mostró consideración hacia aquellos a los que ha afectado la evaluación? ¿Se dispuso para un equilibrio de poder?</p> <p>7. RESULTADOS. Incluye los informes y sus efectos buenos y malos, pero sólo hasta el extremo de lo previsible por el evaluador. También incluye los efectos del <i>proceso</i> de evaluación; ha producido productividad, una moral minada o debilitada, el uso de la información en decisiones?</p> <p>8. COSTES. ¿Cuánto costó la evaluación (especialmente costes no monetarios)? ¿Cuáles hubieran sido las alternativas al coste (incluida no hacerla)?</p> <p>9. COMPARACIONES. Los competidores críticos de la evaluación pueden incluir no molestar de ninguna manera, utilizar juicios validados, tirar una moneda al aire y utilizar otros modelos de evaluación.</p> <p>10. POSIBILIDADES DE GENERALIZACION. Otros usos para el diseño, en momentos futuros, para programas diferentes, en otros lugares, etc.</p> <p>11. RELEVANCIA. Una síntesis de todo lo anterior, para llegar a valorar el valor y el mérito de la evaluación.</p> <p>12. RECOMENDACIONES. Puede haberlas o no, pero pueden incluir: rechazar, repetir (para confirmar), aceptar, ejecutar.</p> <p>13. INFORME. Del metaevaluador.</p>
--

De esta manera, y siguiendo la lista de Scriven, se podrían ir estableciendo criterios de evaluación según cada una de las 13 dimensiones mencionadas: si se describieron el diseño de la evaluación, el equipo evaluador y el calendario, etc., si se contemplaron los antecedentes y el contexto, si se han identificado a los consumidores, si hubo los suficientes recursos, si fueron explicitados los criterios de valor, cómo se ha llevado a cabo el proceso, etc.

Otros criterios para la metaevaluación son las propias normas del *Joint Committe*. Estas son las normas más utilizadas como referencia para metaevaluar la calidad de las de evaluaciones, probablemente debido a su especificación, claridad, orden y estructuración. Como ya se ha mencionado anteriormente, en su segunda edición, publicada en el año 1994, aparece precisamente la metaevaluación como una norma específica, aunque ya en su primera edición se dejaba entrever la idea de utilizar el listado de normas como una guía para la metaevaluación, y se proponía su utilización como lista de comprobación. Estas normas aparecen en la tabla 4.4.⁴ que se expone a continuación. Cada una de las normas es, en realidad, un criterio o pregunta de la metaevaluación.

⁴ Las normas se presentan en cuatro bloques: (U) utilidad, (F) factibilidad/viabilidad, (C) corrección/ legitimidad y (P) precisión

Tabla 4.4: NORMAS PARA LA EVALUACION DE PROGRAMAS
The Joint Committee on Standards for Educational Evaluation. 2ª Edición. 1994⁵

NORMAS DE UTILIDAD
<p>Las normas de utilidad pretenden asegurar que una evaluación servirá a las necesidades de información de los usuarios pretendidos.</p> <p>U1 Identificación de los agentes críticos (<i>stakeholders</i>) Se deben identificar las personas implicadas o afectadas por la evaluación, de manera que sus necesidades puedan ser cubiertas.</p> <p>U2 Credibilidad del evaluador Las personas que dirijan una evaluación deben ser merecedoras de confianza y competentes para realizar una evaluación, de manera que los hallazgos de la evaluación consigan la máxima credibilidad y aceptación.</p> <p>U3 Selección y alcance de la información La información recopilada debe ser seleccionada ampliamente para contestar preguntas pertinentes sobre el programa y debe responder a las necesidades e intereses de clientes y otros agentes críticos.</p> <p>U4 Identificación de valores Las perspectivas, los procedimientos y la fundamentación que se utilicen al interpretar los resultados deben describirse con cuidado, de manera que los fundamentos de los juicios de valor sean claros.</p> <p>U5 Claridad del informe Los informes de evaluación deben describir claramente el programa que está siendo evaluado, incluyendo su contexto, y los propósitos, procedimientos y resultados de la evaluación, de manera que se ofrezca información esencial y fácilmente comprensible.</p> <p>U6 Oportunidad y difusión del informe Los hallazgos intermedios significativos y los informes de evaluación se deben difundir a los usuarios deseados, de manera que puedan ser utilizados de manera oportuna.</p> <p>U7 Impacto de la evaluación Las evaluaciones se deben planificar, dirigir y comunicar de forma que animen a los agentes críticos a su cumplimiento, de manera que aumente la probabilidad de que se use la evaluación.</p>
NORMAS DE FACTIBILIDAD/VIABILIDAD
<p>Las normas de factibilidad tienen el propósito de asegurar que una evaluación sea realista, prudente, diplomática y frugal.</p> <p>F1 Procedimientos prácticos Los procedimientos de la evaluación deben ser prácticos, de manera que las interrupciones sean mínimas mientras se obtiene la información necesaria.</p> <p>F2 Viabilidad política La evaluación se debe planificar y dirigir anticipando las diferentes posturas de diversos grupos de interés, de manera que se pueda obtener su colaboración y así se puedan advertir y contrarrestar posibles intentos de restringir las operaciones de evaluación o de sesgar o abusar de los resultados por parte de estos grupos.</p> <p>F3 Coste Efectividad La evaluación debe ser eficiente y producir una información de valor suficiente, de manera que se justifiquen los recursos empleados.</p>

⁵ Adaptación y traducción por parte de María Bustelo.

NORMAS DE CORRECCION /LEGITIMIDAD⁶

Son las normas dirigidas a asegurar que la evaluación se dirija legal y éticamente y con la debida atención al bienestar de aquellos implicados en la evaluación, así como a aquellos afectados por sus resultados.

C1 Orientación de Servicio Las evaluaciones se deben diseñar para asistir a las organizaciones a tratar y cubrir efectivamente las necesidades de todo el espectro de destinatarios.

C2 Acuerdos formales Las obligaciones de las partes formales de la evaluación (qué se hará, cómo, por quién y cuándo) se deben acordar por escrito, de forma que dichas partes estén obligadas a seguir todas las condiciones del acuerdo o, por el contrario, renegociarlo formalmente.

C3 Derechos de los sujetos Las evaluaciones se deben diseñar y dirigir de tal forma que se respeten y protejan los derechos y el bienestar de los sujetos.

C4 Interacciones humanas Los evaluadores deben respetar la dignidad y el valor humano en sus interacciones con otras personas asociadas a la evaluación, de manera que no se amenace o dañe a los participantes.

C5 Valoración completa y justa La evaluación debe ser completa y justa en su análisis y en su registro de fortalezas y debilidades del programa evaluado, de tal modo que las fortalezas sean reforzadas y las áreas problemáticas tratadas.

C6 Revelación de los resultados Las partes formales de la evaluación deben asegurar que el conjunto completo de resultados de la evaluación junto con sus pertinentes limitaciones son accesibles para las personas afectadas por la evaluación y para cualquier otra persona con derechos legales expresados a recibir los resultados.

C7 Conflicto de intereses Los conflictos de intereses deben manejarse abierta y honestamente de tal forma que no comprometan el procesos ni los resultados de la evaluación.

C8 Responsabilidad fiscal La distribución y los gastos de recursos asignados por el evaluador deben reflejar procedimientos solventes de información y responsabilidad y si no el evaluador debe ser prudente y éticamente responsable de manera que los gastos se justifiquen y sean apropiados.

⁶ El término original en inglés es "*propriety standards*". Aunque nos parece más adecuada la traducción con el término de "normas de corrección", mantenemos también el término de "legitimidad" que es el que aparece en la traducción al español de la 1ª edición de las normas (*Joint Committee*: 1988: 3).

NORMAS DE PRECISION

Estas normas tienen el propósito de asegurar que una evaluación revele y proporcione información técnicamente apropiada sobre las características que determinan el valor o mérito del programa evaluado.

P1 Documentación del programa El programa evaluado debe ser descrito y documentado de forma apropiada y clara, de tal manera que el programa se identifique claramente.

P2 Análisis del contexto El contexto en el cual el programa existe se debe examinar con suficiente detalle de manera que se puedan identificar sus posibles influencias sobre dicho programa.

P3 Descripción de propósitos y procedimientos Los propósitos y procedimientos de la evaluación deben ser seguidos y descritos con el suficiente detalle de manera que se puedan identificar y valorar.

P4 Fuentes de información defendibles Las fuentes de información que se utilicen en una evaluación de programa se deben describir con suficiente detalle de forma que se pueda estimar la adecuación de la información.

P5 Información válida Los procedimientos de recopilación de información se deben elegir y desarrollar y entonces poner en práctica de manera que se asegure que la interpretación a la que se ha llegado es válida para el uso que se le pretende dar.

P6 Información fiable Los procedimientos de recopilación de información se deben elegir y desarrollar y entonces poner en práctica de manera que se asegure que la información obtenida es suficientemente fiable para el uso que se le pretende dar.

P7 Información sistemática La información recopilada, procesada y transmitida en una evaluación debe ser revisada sistemáticamente y cualquier error encontrado debe ser corregido.

P8 Análisis de la información cuantitativa La información cuantitativa de una evaluación se debe analizar apropiada y sistemáticamente de manera que las preguntas de evaluación se contesten de forma efectiva.

P9 Análisis de la información cualitativa La información cualitativa de una evaluación se debe analizar apropiada y sistemáticamente de manera que las preguntas de evaluación se contesten de forma efectiva.

P10 Conclusiones justificadas Las conclusiones a las que se llega en una evaluación deben estar justificadas explícitamente de manera que los agentes críticos de la evaluación puedan ponderarlas.

P11 Informe imparcial Los procedimientos para informar deben asegurar que no haya distorsión causada por sesgos y sentimientos personales de cualquiera de las partes de la evaluación, de manera que los informes reflejen de forma justa los resultados de la evaluación.

P12 Metaevaluación La evaluación en sí misma se debe evaluar formativa y sumativamente a través de éstas y otras normas pertinentes, de manera que se guíe de forma apropiada su realización y, en cuanto se termine, los agentes críticos puedan examinar de cerca sus fortalezas y debilidades.

La metaevaluación de evaluaciones de políticas públicas de igualdad de género

En este apartado se describe el tipo de metaevaluación que se ha realizado en esta tesis doctoral. En general, se le ha dado prioridad a la metaevaluación como procedimiento para describir, analizar y valorar los procesos de evaluación que se han producido en relación a las políticas de igualdad de género en España, en los niveles autonómico y central. Es decir, no se plantea la metaevaluación como una forma de controlar la calidad de las evaluaciones realizadas, sino como una forma de estudio de las políticas públicas, en su fase concreta de evaluación. Más que la valoración de la calidad de las evaluaciones estudiadas, se realiza una valoración de su adecuación y oportunidad al contexto evaluativo y de la política, y de su contribución al cumplimiento de las funciones de la evaluación. Se pretende, además, dar una visión global sobre cómo se desarrolla la función de evaluación en el sector de las políticas de igualdad de género en España, y en el nivel autonómico, tomando también como contraste lo realizado en evaluación en este sector en el nivel central. Esto no obsta para que hayamos tenido también como referencia la perspectiva de la metaevaluación como control de calidad para elaborar los criterios concretos que hemos utilizado en nuestro estudio.

Por esta razón se ha optado por tener por objeto de la metaevaluación un conjunto de varias evaluaciones en lugar de centrarse en uno o dos procesos de evaluación. Como se ha comentado anteriormente, en este trabajo no se pretende juzgar los procesos concretos sino la función de evaluación en las políticas públicas de igualdad de género. Por lo tanto, se ha favorecido la descripción general comparativa entre varias evaluaciones por encima de la descripción pormenorizada y en profundidad de cada proceso de evaluación.

Respecto al tipo de metaevaluación realizada en este trabajo, se señala lo siguiente:

Es una metaevaluación *sumativa*, ya que hace recapitulación de los procesos de evaluación habidos y completados en un tiempo determinado. La intención es realizar un análisis global y sumativo de dichos procesos.

Es una metaevaluación *post-facto* ya que los procesos de evaluación que se meta-evalúan ya han ocurrido y se describen y valoran *a posteriori*.

Es una metaevaluación *externa* ya que se realiza por alguien ajeno tanto a los equipos de evaluación que realizan las diversas evaluaciones como a los organismos de igualdad que las encargan.

Finalmente, pretende ser una metaevaluación de *diseño*, de *proceso* y de *resultados*.

Respecto al *diseño* de los procesos de evaluación, se plantean cuestiones relacionadas con el planteamiento general de los mismos, teniendo en cuenta el contexto evaluativo en el que se producen, la claridad en los propósitos y funciones de la

evaluación, las estructuras institucionales que apoyan las evaluaciones, los tipos de evaluación realizadas y los recursos empleados en las mismas. También se estudian variables como la inclusión o no de diferentes agentes críticos en los diseños de evaluación, la adecuación del calendario diseñado y, por supuesto, los criterios e indicadores de evaluación utilizados, además de los procedimientos, metodología y técnicas empleadas en la recopilación, análisis y valoración de la información.

Del **proceso** de las evaluaciones, se tienen en cuenta, sobre todo, cuestiones relacionadas con los actores participantes en los mismos y los tiempos y calendarios de los estudios de evaluación. También se señalan las incidencias y dificultades significativas a lo largo del proceso de las evaluaciones.

Finalmente, en la metaevaluación de **resultados**, se pretende analizar tanto los productos -los informes y el formato en la comunicación y difusión de resultados- como los efectos -tanto en lo que se refiere a la utilidad y pertinencia de la información producida como a la utilización de dicha información evaluativa producida para modificar los planes y políticas, controlar responsabilidades o iluminar acciones futuras. También se hace especial hincapié en los efectos que han tenido los procesos de evaluación sobre el aprendizaje organizativo en los organismos de igualdad y sobre una mayor conciencia sobre la situación de los planes y las políticas de igualdad.

Sin embargo, las limitaciones en cuanto a la información disponible (por dificultad en su acceso en algunos casos y por lo reciente de los procesos, en otros) sobre dichos procesos de evaluación hacen que, en algunos casos, sólo se apunten algunas conclusiones de forma tentativa, especialmente en lo que se refiere a los efectos de los procesos de evaluación.

Igualmente se descarta la idea de realizar una **síntesis de evaluaciones**, en el sentido de sintetizar los resultados hallados a través de las evaluaciones meta-evaluadas, para poder llegar a conclusiones sobre las políticas de igualdad de género en España en los niveles autonómico y central. Esta idea se rechaza desde un principio cuando se detecta, en la primera recopilación de información tentativa que, en general, los procesos de evaluación de los planes de igualdad han sido pocos y muy heterogéneos entre sí, existiendo en realidad muy poca base para la comparación de resultados. Además, se ha preferido dar prioridad al estudio de los procesos de evaluación y no a los resultados que éstos producen; en definitiva, recalamos la idea de que el objeto de estudio de esta tesis no son las políticas de igualdad de género, sino los procesos de evaluación habidos en torno a las mismas.

1.2. El análisis de los procesos de evaluación de las políticas públicas de igualdad de género en España

Ya se ha comentado anteriormente que la fase de evaluación es la menos estudiada desde el análisis de políticas públicas. Especialmente esta fase ha sido muy poco tratada desde el componente descriptivo de dicho análisis, es decir en el estudio de los procesos de evaluación. Aún así, existen algunas excepciones en el análisis de la función de evaluación en las diferentes naciones (Rist, 1990; Ballart, 1993). En su componente prescriptivo, desde el análisis de políticas públicas, sí existen algunas publicaciones especializadas y algunos manuales que desarrollan el concepto de evaluación y de políticas públicas y, principalmente, sobre cómo aplicarlo en dichas políticas y, especialmente, en el ámbito gubernamental (algunos ejemplos son Ballart, 1992; Vedung, 1997; Monnier, 1997; Weiss, 1998).

Por otro lado, los poderes públicos necesitan evaluar para obtener signos explícitos que permitan ajustar sus acciones. Cuentan también con otras fuentes de información como sondeos y encuestas, investigaciones, servicios de información, etc., pero la evaluación puede ser una herramienta valiosa para facilitar el aprendizaje organizativo y la toma de decisiones para la mejora, el futuro y el control de responsabilidades de una determinada política. Sin embargo, aunque el concepto y la fase de evaluación de las políticas públicas se haya desarrollado de forma más o menos contundente y vertiginosa en los últimos años en la teoría, la práctica real de la evaluación está todavía muy lejos de lo que se propugna en los manuales. Esto ocurre de forma especial en algunos países, como España, en los que el desarrollo institucional de la evaluación ha sido muy bajo y la consiguiente cultura evaluativa prácticamente inexistente, sobre todo en las administraciones públicas. Hace falta, por lo tanto, que las instituciones públicas sean conscientes de su bajo grado de desarrollo teórico-práctico respecto a la evaluación de sus acciones e intervenciones. Es necesario, pues, que se comience a crear una cierta cultura de evaluación que puede generarse no sólo con la propia práctica evaluativa, sino también con el hecho de hacer de los procesos de evaluación objeto de estudio y con el aprendizaje extraído de los análisis realizados de las primeras experiencias de evaluación.

Si nos centramos en concreto en las políticas de igualdad de género, vemos que, en su análisis sobre el feminismo de Estado y las políticas de igualdad en España, Celia Valiente (1997) concluye que dicho feminismo de Estado ha sido eficaz en la definición de problemas y en la introducción de los mismos en la agenda política, pero no lo ha sido tanto ni en la fase de formulación y ejecución de dichas políticas ni, sobre todo, en la función de control de dicha ejecución. Aquí es donde la evaluación de dichos planes adquiere su importancia. Las “femócratas” o feministas de Estado⁷, conscientes de la importancia de dicha función de control y el papel que los organismos de igualdad pueden cumplir respecto a la misma, comienzan a dirigir e impulsar, de forma creciente, procesos de evaluación de los planes de igualdad. Por lo tanto, resulta interesante un

⁷ Término acuñado para mencionar a las responsables de los organismos específicos de igualdad (véase capítulo tres).

análisis de dichos procesos de evaluación, profundizando en la respuesta que se está dando por parte del feminismo de Estado para el cumplimiento de la función de control de las políticas públicas de igualdad.

Las evaluaciones que se llevan a cabo en torno a las políticas de igualdad de género son intentos de evaluación de los **planes** de igualdad, que como se ha visto en el capítulo tres de este trabajo, se generalizan en todas las comunidades autónomas en los primeros años de la década de los años noventa. Como también se ha visto anteriormente, los planes de igualdad son un conjunto de objetivos y medidas tomadas y aprobadas por un gobierno y que deben llevar a cabo los diferentes departamentos gubernamentales, en un período concreto de tiempo (de dos a cinco años). Los objetivos y medidas suelen incluir todos los aspectos o áreas que afectan a la mujer (educación, cultura, sanidad, legislación, medio ambiente, medio rural, etc.); por lo tanto, los planes de igualdad están diseñados para ser un paraguas común bajo el que se explicitan todas las acciones públicas referidas a la igualdad de género de un determinado gobierno. Con esto insistimos en el carácter integral e inclusivo de los planes. Los planes también suelen implicar, al menos formalmente, otros niveles gubernamentales y organizaciones no gubernamentales. Esto hace que los planes de igualdad sean la principal herramienta que articula las políticas de igualdad de género, y que sea rara la actuación que se realice en ese área por parte de los poderes públicos y que no esté contemplada en los planes. Por lo tanto, la evaluación de los planes es esencial para entender y evaluar las políticas de igualdad. Además, y como se ha visto en el capítulo tercero de este trabajo, en España y tanto en el nivel central como en el autonómico, los planes no sólo han sido aceptados como el instrumento principal sino que, en cierto sentido, como el único. De esta forma, el análisis de la idoneidad, eficacia y eficiencia de estos planes aparece como crucial no sólo para el estudio de la función de evaluación en un determinado sector, sino también para el estudio de las propias políticas de igualdad de género. Por lo tanto, el **objeto** de análisis y valoración en este trabajo se centra en la evaluación de los **planes**, como máxima y principal expresión de las políticas de igualdad de género.

Respecto a la delimitación **territorial**, este trabajo se centra en estudiar cómo se han evaluado los planes de igualdad en las diferentes comunidades autónomas y en el nivel central. El nivel central es el primero en desarrollarse en España, respecto a las políticas de igualdad de género, tal y como se entienden en la actualidad⁸, y el nivel autonómico se va desarrollando a lo largo de los últimos 15 años, a la luz de la experiencia central y, sobre todo, por el propio desarrollo institucional y organizativo del Estado de las Autonomías en España. El nivel autonómico, por lo tanto, no sólo es el nivel gubernamental que más ha crecido y se ha desarrollado en los últimos años en España, sino que también, y debido a las competencias que las comunidades autónomas han adquirido, aparece como un nivel clave en la formulación de las políticas sociales en general. Por lo tanto, parece necesario contar con este nivel autonómico en el análisis, incluidos los procesos de evaluación, de las políticas públicas de igualdad de género en España.

El **marco temporal** del análisis está determinado por la propia realidad de los planes de igualdad y sus evaluaciones. Aquí se consignan los primeros intentos de

⁸ Véase el capítulo 3. Se toma como punto de partida de las políticas de igualdad de género en España la creación del IM, a finales del año 1983.

evaluación, coincidentes con algunos primeros planes⁹, pero los procesos de evaluación que se estudian en profundidad y que se seleccionan para la metaevaluación son los que se llevan a cabo a partir del año 1995 y se completan antes de terminar el año 1999, y que denominaremos de segunda generación¹⁰. Salvo en los casos de Aragón, Canarias y Castilla y León, en donde se evalúan los primeros planes, pero que por criterios cronológicos son incluidos en esta segunda generación, el resto de las evaluaciones son de los segundos planes.

⁹ Los primeros intentos de evaluación se realizan del I Plan a nivel central (1988-1990), del I Plan Catalán (1989-1992), el I Plan Andaluz (1990-1992), y el I PAPME Vasco (1991-1994).

¹⁰ Se incluye también la evaluación del II PAPME vasco (1996-2000), ya que la evaluación de dicho plan se plantea desde un principio y de forma continua, con resultados anuales. Al ser, además, una de las experiencias de evaluación más afianzadas y que cuenta con el precedente de la evaluación del PAPME anterior, se cree especialmente interesante incluirlo en el análisis. También se incluye el caso de la evaluación del II Plan riojano (1996-1999), que se plantea como una evaluación interna desde el inicio, aunque la fase final de la evaluación no se concluye hasta el año 2000 (esta fase final que consiste en una evaluación externa no entra, por lo tanto, en nuestro análisis).

2. UN PANORAMA GENERAL: EXPERIENCIAS DE EVALUACIÓN EN POLÍTICAS DE IGUALDAD EN LOS NIVELES CENTRAL Y AUTONÓMICO

En España se han dado algunas experiencias de evaluación de planes de igualdad centrales y autonómicos en los últimos años. Estas experiencias son heterogéneas entre sí y limitadas, sobre todo, en cuanto a su alcance (se suelen evaluar sólo algunos aspectos del plan) y en cuanto al calendario (la mayoría de las evaluaciones se planificaron una vez el plan ya estaba diseñado y había comenzado). Sin embargo, todas estas experiencias tienen el valor de su propia existencia y de haber “roto el fuego” en el terreno de la evaluación y, en general, han cumplido un papel importante en las políticas de igualdad de género en España, tal y como se verá más adelante.

En primer lugar, se debe señalar que las experiencias de evaluación en torno a las políticas de igualdad de género se circunscriben a evaluaciones de los planes de igualdad y, en algunos casos, a la evaluación de algunas actuaciones o programas concretos¹¹. Es decir, en ningún caso se ha hecho ni se ha formulado en términos de evaluación de la política global.

Se pueden clasificar las experiencias de evaluación en tres fases o momentos: los primeros intentos de evaluación que se dan de forma muy limitada en cuanto a número y alcance, y que se producen en los primeros años de la década de los noventa; lo que se podría llamar la segunda generación de evaluaciones, que es el objeto de esta investigación, y que coincide en casi todos los casos con las evaluaciones de los segundos planes (en alguna Comunidad Autónoma, como Aragón, Canarias y Castilla y León, se trataría de los primeros planes) y que se dan a partir del año 1995; por último, estarían los proyectos de evaluación para los terceros (en algunos casos segundos o incluso primeros, como es el caso de Navarra) y que se supone deberían estar completados hacia el 2000-2001 en algunos casos y, en otros, más adelante. A continuación, se exponen algo más pormenorizadamente estas tres fases o momentos.

¹¹ Aquí consignamos de las que tenemos constancia: en Asturias, la evaluación de los talleres itinerantes socioculturales (1996), y los programas NOW (que en el caso de esta Comunidad Autónoma se realizan por parte del mismo que los gestiona). Más reciente es la evaluación del programa ELIGE en Andalucía. En general, los programas que cuentan con dinero europeo -como los NOW- suelen requerir la existencia de una evaluación, que en muchos casos se da como una mera formalidad para “cubrir el expediente”, pero que en otros sirve para generar una información útil e ir generando una cierta cultura evaluativa en torno a los organismos de igualdad.

2.1. Las primeras experiencias de evaluación

Existe una primera “generación” de evaluaciones que se da en la primera mitad de la década de los noventa. Estos primeros intentos, como ya se ha mencionado, son muy reducidos en número (sólo se dan en el nivel central y en tres comunidades autónomas) y, en general, limitados en cuanto a su alcance, aunque existen diferencias entre los mismos en esfuerzos y recursos empleados, en la recopilación sistemática de información y su valoración y en su eficacia evaluativa.

Se dan procesos de evaluación de los siguientes planes¹²:

- I Plan de Igualdad de Oportunidades de las Mujeres **1988-1990**, lanzado por el Instituto de la Mujer (IM a partir de ahora) en el nivel central
- I Plan de Actuación del Gobierno de la Generalidad de Cataluña para la Igualdad de Oportunidades para las Mujeres **1989-1992** (Instituto Catalán de la Mujer - *Institut Català de la Dona*)
- I Plan de Igualdad de Oportunidades de las Mujeres **1990-1992** (Instituto Andaluz de la Mujer)
- I Plan de Acción Positiva para las mujeres en la Comunidad Autónoma de Euskadi. **1991-1994** (Instituto Vasco de la Mujer- *Emakunde*).

Salvo en el caso del País Vasco, en el que la apuesta y los esfuerzos empleados en la evaluación son muchos mayores (por lo que los resultados resultan más completos y útiles), y en el de Andalucía, en donde se contrata a una entidad consultora externa para que colabore en la evaluación el último año, y que hace que la evaluación finalmente tenga mayor alcance y amplitud, estas primeras evaluaciones mencionadas son muy débiles y se limitan casi exclusivamente a una comprobación del cumplimiento de las acciones que se habían aprobado en los respectivos planes. Esta debilidad está más que justificada en estas primeras experiencias de evaluación, no sólo por la carencia de precedentes, competencia y cultura evaluativa, sino también porque la calidad técnica de

¹² Se puede mencionar también la evaluación realizada del primer Plan extremeño (1991), que se realiza y da lugar a un documento en 1994. Los dos técnicos que elaboran esta evaluación no se encontraban trabajando en la Dirección General de la Mujer en el periodo de ejecución del plan, y la realizan recopilando información interna en esa misma Dirección. La evaluación consiste en la realización de una ficha por cada objetivo de actuación del Plan en la que se contemplan tres tipos de información: 1) fecha de inicio y fin prevista y fecha de inicio y fin real, 2) grado de cumplimiento, con 4 categorías (acción finalizada, se concluirá en breve, iniciada recientemente y acción sin iniciarse) y 3) Acciones/programas. Finalmente se hace una valoración del grado de cumplimiento por cada área en forma de porcentaje. Aunque esta “evaluación” supone un esfuerzo de ordenación de la información por parte de la Dirección General, no puede ser considerado como un verdadero proceso de evaluación sistemática. Entre otras cosas, dicha evaluación no parece que llegue a ninguna conclusión, ya que en Extremadura no se ha elaborado, por el momento, un II Plan. En este momento la Dirección General de la Mujer de Extremadura es, además, bastante diferente a sus organismos homólogos autonómicos, ya que tiene un fuerte carácter de ejecución de acciones y actuaciones, más que de impulso de las mismas. Precisamente, ésta es la única Comunidad Autónoma en la que el primer Plan no ha tenido continuidad real en un segundo Plan.

los primeros planes era menor que la de los posteriores, sobre todo en el sentido de que contaban con objetivos y medidas muy generales o vagos y, en ocasiones, excesivos; en definitiva, los primeros planes eran difícilmente evaluables. Precisamente, estas primeras experiencias de evaluación contribuyen al aprendizaje y rodaje de las femócratas y, por lo tanto, a mejorar la calidad técnica de los planes subsiguientes.

Si se observa su procedencia, vemos que las primeras experiencias de evaluación se dan en los organismos de igualdad más antiguos y con una estructura institucional más fuerte (organismos autónomos)¹³. Esta circunstancia es coincidente con el hecho de que, además del propio IM, son tres comunidades autónomas que acceden a la autonomía por la llamada vía rápida o artículo 151 de la Constitución Española. En algunas comunidades autónomas en las que se aprueban los primeros planes también en fases tempranas, no se dan, sin embargo, experiencias de evaluación, o al menos no se tiene constancia ni documentación de las mismas (Asturias, Madrid¹⁴ y Comunidad Valenciana -cuyos primeros planes van de 1989 a 1991-, Castilla-La Mancha -cuyo I Plan es de 1990 a 1994- y Baleares y Cantabria -cuyos primeros planes son de 1991-1993). A continuación se ofrece alguna información específica sobre cada uno de los casos.

Evaluación del I Plan de Igualdad de Oportunidades de las Mujeres 1988-1990

La primera evaluación que se realiza es la que se hace del **I Plan central**. Esta evaluación se hace de forma interna y consiste exclusivamente en comprobar y consignar si las 120 medidas contempladas en el Plan se habían cumplido o no (es decir, si se habían realizado). En el propio documento de evaluación¹⁵ (una publicación del Plan, que añade una segunda parte sobre la evaluación del mismo) se menciona que lo que se hace es un “balance de cumplimiento del mismo” (pág. 105 del documento)¹⁶. La conclusión es que el I Plan tiene un grado de cumplimiento del 96-97% porque se consigna que se han ejecutado 116 medidas. Las cuatro restantes (tres referentes a reformas en el Código Civil y Código Penal y la otra un Convenio con RTVE para contribuir al cambio de actitudes a través de los medios), quedan pendientes una vez finalizado el período de vigencia del Plan.

¹³ Véase el punto 2.4. del capítulo 2.

¹⁴ En el caso de Madrid, existe una pequeña publicación llamada “evaluación” y en cuya introducción se dice textualmente: “Esta publicación recoge la Evaluación del Plan, detallando todas las actuaciones desarrolladas, en los 3 años de vigencia”. Efectivamente, el documento sólo recoge un listado ordenado de las actuaciones realizadas. No se recoge ninguna otra información, ni valoración ni tampoco ninguna referencia a un posible proceso de evaluación. En la Dirección General de la Mujer no existe ninguna constancia de la existencia de dicho proceso. (Dirección General de la Mujer. Comunidad de Madrid. “Plan de Igualdad para las Mujeres. 1989-1991. Evaluación” 9 pp. Sin ISBN ni fecha).

¹⁵ INSTITUTO DE LA MUJER (Ministerio de Asuntos Sociales). 1990. “Plan de Igualdad de Oportunidades de las Mujeres 1988-1990. Evaluación”. 155 pp.

¹⁶ Todas las citas mencionadas en este apartado se refieren al documento INSTITUTO DE LA MUJER (Ministerio de Asuntos Sociales). 1990. “Plan de Igualdad de Oportunidades de las Mujeres 1988-1990. Evaluación”. 155 pp.

En este documento sobre la evaluación del primer Plan central, se realiza una valoración general de lo que ha supuesto el Plan. Posteriormente, se trata cada una de las 6 áreas del plan (ordenamiento jurídico, educación y cultura, empleo y relaciones laborales, salud, cooperación internacional y asociacionismo). En cada área se hace una introducción sobre el significado e importancia de la misma y una valoración general sobre lo que se ha realizado. Luego se ve medida por medida, explicando las actuaciones realizadas y su significado, simplemente describiendo lo que se ha hecho, sin realizar juicios ni valoraciones sobre el cumplimiento de las mismas. Al final del documento se exponen dos listados: uno sobre los Departamentos Ministeriales implicados en la ejecución del Plan para la Igualdad de Oportunidades, en el cual se listan los Ministerios implicados y su forma de implicación concreta (comisiones mixtas, convenios, acuerdos y grupos de trabajo) y otro sobre las Campañas de Información y Sensibilización realizadas, en el que se detallan las publicaciones editadas y las campañas en los medios de comunicación, con sus correspondientes fechas.

Sin embargo, y a pesar de las limitaciones que presenta esta primera evaluación, en las valoraciones generales que sobre el Plan se realizan en este documento se insiste en la importancia del Plan como forma de articular todas las actuaciones del Gobierno en materia de igualdad, reflejando el espíritu de lo que luego se conocerá como *mainstreaming* promovido por la Plataforma de Acción de la IV Conferencia de las Naciones Unidas en Pekín:

“Se puede afirmar que la virtualidad del Plan ha residido en ser un instrumento para articular e impulsar políticas destinadas a corregir desigualdades entre los sexos, para desarrollar una política integral para las mujeres por un lado, y por otro para, desde el diseño de coordinación horizontal, en el que se basaba el Plan, poder incluir la solución a las demandas específicas de las mujeres en los objetivos de política general del gobierno, sean éstas de empleo, educación o salud, huyendo del riesgo de algunas estrategias para la igualdad de oportunidades: la excesiva sectorialización, la política de gueto” (p. 105).

Y más adelante se señala que:

“El conjunto de medidas del Plan significa, por tanto, más allá de la importancia concreta de cada una de las actuaciones, que por primera vez se ha puesto en marcha en España, una política integral y coordinada para avanzar en la igualdad de oportunidades entre hombres y mujeres, y por lo tanto, en la consecución de una sociedad más justa” (p. 109).

Es decir, en la valoración que se realiza del Plan se ofrece una visión que va más allá de la valoración del cumplimiento de cada una de las actuaciones contempladas en el mismo. También se realizan otras valoraciones generales. Por ejemplo, se resaltan dos logros importantes como que a lo largo del desarrollo del Plan

“se han ido creando mecanismos de coordinación con los diferentes Departamentos ministeriales, consolidándose así canales de colaboración institucional con los Ministerios Implicados” (p. 106).

En segundo lugar, la influencia y el efecto rebote en las comunidades autónomas:

“y aunque el Plan sólo implicaba a la Administración Central del Estado, se ha producido durante este período un fenómeno de sinergia de forma que muchas Comunidades Autónomas han constituido en sus gobiernos Organismos para impulsar políticas específicas dirigidas a la mujer; dichos Organismos han elaborado programas para la igualdad que, en la mayor parte de los casos, son adaptaciones del Plan del Gobierno Central a las condiciones y necesidades específicas de las mujeres en las respectivas Autonomías” (p. 106).

Efectivamente, en 1990 había ya siete comunidades autónomas (Andalucía, Asturias, Cataluña, Castilla-La Mancha, Madrid, Valencia y el País Vasco) que habían aprobado sus primeros Planes, todas ellas gobernadas por el PSOE -el mismo que gobernaba en el nivel central y había impulsado las políticas de igualdad de oportunidades- o por partidos de base autonómica.

También se aportan otros datos, como que en la encuesta realizada por el Instituto en 1990 para conocer las demandas de las mujeres españolas, se había constatado que el 65% de las mujeres encuestadas tenía conocimiento del Plan (esto se expresa de un modo más triunfalista “el grado de conocimiento que había alcanzado el Plan de Acción entre la población era de un 65%” p. 107). Otra información que se aporta en forma de logro (aunque no resaltándola especialmente como el establecimiento de los mecanismos de coordinación interministeriales y la influencia en las comunidades autónomas) es la constitución de “una plataforma de asociaciones de mujeres de toda España, que, con la denominación de Comisión Consultiva para el seguimiento del Plan para la Igualdad, ha designado representantes para el Consejo Rector del IM, al igual que los sindicatos mayoritarios CC.OO. y U.G.T., lo que ha permitido que las organizaciones no gubernamentales de mujeres puedan hacer oír su voz en el órgano desde el que se coordinan e impulsan las políticas para la igualdad” (p. 107). Sabemos, por el estudio realizado por Celia Valiente, que la experiencia de dicha Comisión Consultiva fue muy desesperanzadora¹⁷ y que la representación de los colectivos de mujeres en un Consejo Rector, que apenas se reunió durante aquellos años, fue prácticamente simbólica (Valiente, 1995, 1996 y 1997).

Finalmente, en el documento de evaluación de este primer plan central se deja entrever la idea de que la evaluación del I Plan se utilizará para la redacción del II, que será continuidad del I y del que se aportan ya una serie de líneas de acción, aunque el II Plan no será aprobado hasta 1993.

Esta evaluación del I Plan central ha tenido ciertas críticas. Para Valiente (1996: 185) las afirmaciones de la evaluación han de tomarse con extrema cautela por dos razones: una, porque los objetivos no eran claros y concretos y, por lo tanto, resultaba

¹⁷ Dicha Comisión Consultiva se estableció en 1988 y su propósito era, precisamente, evaluar el cumplimiento del I Plan y elaborar recomendaciones y propuestas al respecto que serían tenidas en cuenta por el IM. Este intento de acercamiento, sin embargo terminó siendo lo que Valiente califica como “una crucial oportunidad perdida”. Las feministas criticaron haber sido llamadas para hacer el seguimiento de un Plan en el que no habían participado en su formulación y lo interpretaron como un mero gesto simbólico de las femócratas. Numerosas participantes se negaron a participar desde un principio y entre las que lo hicieron enseguida surgieron desavenencias y luchas por la competencia entre asociaciones por ser las más representativas de los problemas y las opiniones del conjunto de las mujeres. Resultó prácticamente imposible la aprobación de las recomendaciones más básicas. (Valiente, 1996:188-190).

muy difícil llegar a conclusiones sobre su verdadero grado de consecución; la segunda razón es que la base de la información sobre la que se realizó la evaluación era la información proporcionada por los propios departamentos ministeriales que probablemente tuvieran una tendencia a enfatizar en exceso lo conseguido y a minimizar lo que no llegó a hacerse. Efectivamente, la evaluación tiene un cierto carácter triunfalista y, además, carece de una metodología explícita, aunque denota una cierta recopilación sistemática y una cierta reflexión sobre sus actuaciones. Además, hasta cierto punto, es comprensible que se diera cierto triunfalismo, ya que el sólo hecho de haber conseguido la creación del Instituto y la elaboración y aprobación del primer Plan era ya, en sí mismo, un éxito. Había medidas cuyo cumplimiento era evidente (como eran todas las relacionadas con los cambios legislativos) y otras, por el mismo hecho de plantearse y de llevarse a cabo por primera vez eran ya de por sí, un triunfo. En muy pocos años, y a pesar de haber empezado con una década de retraso respecto a otros países occidentales, se había establecido un organismo específico con una estructura y un presupuesto homólogos a sus equivalentes europeos (Valiente, 1995, 1996 y 1997) y en menos de cinco años de existencia, la política de igualdad de oportunidades se articulaba en un instrumento innovador que implicaba a otros sectores del Gobierno y que era reconocido internacionalmente como la mejor estrategia.

Evaluación del I Plan de Actuación del Gobierno de la Generalidad de Cataluña para la Igualdad de Oportunidades para las Mujeres 1989-1992

No se ha podido obtener ningún documento referente a la evaluación del **primer plan catalán**, a pesar de que las responsables del Instituto Catalán de la Mujer dicen que la hubo aunque con limitaciones. En cualquier caso, en la introducción que hace Joaquima Alemany, Vicepresidenta del *Institut Català de la Dona*, se hace referencia a que el II Plan ha sido elaborado tomando como base el III Plan de Acción Comunitario a medio plazo para la igualdad de oportunidades entre hombres y mujeres (1991-1995), lanzado por la Comisión Europea y también las conclusiones del II Congreso de la Mujer en Cataluña y “las demandas del tejido asociativo de mujeres de nuestro país” (Instituto Catalán de la Mujer, 1994: 13). Se habla de que este segundo plan es una continuación y ampliación del primero, pero en ningún momento se menciona la evaluación del I Plan como base o como reflexión para el II.

Evaluación del I Plan de Igualdad de Oportunidades de las Mujeres 1990-1992 (Instituto Andaluz de la Mujer)

El **primer plan andaluz (1990-1992)**, cuya vigencia estaba prevista hasta 1992, se prolonga hasta 1994, año en el que se presentan los resultados de su evaluación al Parlamento de Andalucía, que había creado una Comisión de Seguimiento y Vigilancia del Plan. Esta evaluación se supone que da pie a la elaboración del II Plan que va de 1995 a 1997. El primer Plan andaluz se evalúa por parte del propio IAM, aunque se cuenta con la ayuda de una consultora externa (Fondo de Formación de Sevilla) que

colabora con el IAM para realizar la evaluación durante el año 1993¹⁸. También como parte del proceso de evaluación, en el año 1992 se realizó un primer informe de evaluación sobre el grado de cumplimiento de cada una de las acciones, que se realizó internamente.

En el Informe de Evaluación del primer Plan Andaluz se dedican sólo 5 páginas de las 141 que tiene el informe a hablar de la evaluación, sus objetivos, metodología, aspectos que se han evaluado y explicación de la exposición de los resultados. Respecto a los **objetivos** se plantean tanto la mejora de la gestión y la eficacia de las acciones programadas como el rendimiento de cuentas tanto al Parlamento como a los órganos de control del Consejo de Gobierno. La **metodología** parte de “varias reuniones entre las responsables del Instituto Andaluz de la Mujer y equipo evaluador para definir los cometidos y alcance de la evaluación”. Los **instrumentos** mencionados para la recopilación de información son, textualmente, “cuestionarios para cada una de las actuaciones del Plan, entrevistas en profundidad a responsables políticos y técnicos del mismo, análisis de los documentos existentes en relación con cada una de las actuaciones, estudio de factores contextuales, etc.” (p.41). Respecto a los **aspectos** que se han evaluado se mencionan: a) El grado del cumplimiento del Plan, b) El impacto del Plan, c) Los programas y acciones que han tenido mayor incidencia, d) La relación entre los recursos implementados y los resultados conseguidos y e) Las necesidades y demandas en torno a la igualdad de oportunidades de las mujeres que se han derivado con ocasión del Plan, pero que no fueron previstas en el mismo. En un análisis de los aspectos pretendidos, se puede observar que las intenciones son bastante ambiciosas pero que no pueden cumplirse con el proceso de evaluación que se llevó a cabo. Si además, se analizan los resultados, se ve que la recopilación y análisis sistemático de la información se han hecho sobre todo para cubrir el primer aspecto mencionado (grado de cumplimiento de las acciones), consignando el costo económico de las acciones de cada área y teniendo en cuenta, en algunos casos, las opiniones de algunos agentes críticos, pero no de forma sistemática. Por ejemplo, evaluar “el impacto del Plan en las personas beneficiarias y destinatarias directas y en los cambios que se han producido y que puedan estar relacionados con el mismo” (aunque se suavice con la coletilla “en aquellos aspectos en que se podía contar con datos significativos” resulta casi imposible con los instrumentos de recopilación de información utilizados que se mencionan más arriba. Respecto a los programas y acciones que han tenido “mayor incidencia”, depende de cómo se defina este extremo, pero se podrían contemplar cuestiones relacionadas con el diseño y el proceso de puesta en práctica del Plan; sin embargo, por la lectura de resultados, creemos que se trata sobre todo de saber en qué áreas se han empleado más recursos. Finalmente, se menciona que la **presentación de resultados** se realiza por los objetivos del plan, contribuyendo así a una evaluación que carga las tintas en el grado de cumplimiento de cada una de las acciones. Este sesgo hacia el grado de cumplimiento y las dificultades de cubrir otros aspectos, pueden considerarse hasta cierto punto normales si tenemos en cuenta que las 200 medidas que planteaba el I

¹⁸ El estatus “interna” o “externa” de la evaluación de este primer plan no está del todo claro. Se habla siempre de “colaboración” de una consultora externa. En el Informe de Evaluación, que aparece con la autoría del propio Instituto se lee textualmente: “A lo largo del presente año 1993 se ha llevado a cabo, con la colaboración de una empresa consultora, una evaluación externa del Plan para la Igualdad, cuyos resultados se ofrecen en el presente informe” *Informe de Evaluación del Plan para la Igualdad de Oportunidades. 1990-1993* p. 40). En el informe no existe ninguna referencia concreta a dicha empresa consultora aunque se les denomina “el equipo evaluador”.

Plan eran “muy vagas, muy blandas y muy pobres en relación a los objetivos propuestos (...) baratas, indirectas y de resultados inciertos, debido precisamente a la necesidad de obtener la colaboración voluntaria de otros departamentos” (Granados, 1999: 802). Ante tal panorama, es comprensible que sólo el hecho de “seguir la pista” a cada una de las medidas y comprobar su ejecución y la colaboración departamental no sólo sea una ardua labor, sino que baste y sea útil para una primera evaluación de un primer plan.

Aun así, y como era previsible dada la presentación de resultados, esta evaluación del primer plan andaluz ha tenido algunas críticas. Elena Granados, en su estudio sobre el Instituto Andaluz de la Mujer, concluye que “a pesar de la colaboración de una consultora externa, los resultados del I Plan se presentaron de una forma extremadamente optimista, si tenemos en cuenta que muchas de las actuaciones previstas perseguían objetivos muy difíciles de lograr en un período de tiempo tan corto” (Granados, 1997: 30). Para la autora, los indicadores utilizados son los que con mayor probabilidad van a arrojar resultados positivos, poniendo el ejemplo que para evaluar los programas de fomento del empleo en lugar de utilizar el indicador del nº de mujeres que logran acceder a un puesto de trabajo, se utilizan los indicadores de nº de cursos realizados, nº de asistentes a los cursos, coste de los cursos, etc. Y señala que es curioso que sólo se mencione un fracaso en todo el Plan¹⁹. Ciertamente, los resultados del Plan resultan casi siempre positivos y si nos fijamos en la valoración global del Plan que se realiza en las páginas 44 y 45 del informe, previamente a la presentación de resultados por cada una de las áreas, se observa un aire triunfalista y excepcionalmente optimista²⁰. Por ejemplo, el primer resultado mencionado “Ha alcanzado una elevada notoriedad, y sus principales programas han sido conocidos por un alto porcentaje de la población” (p. 44), no sólo adolece de concreción (¿qué porcentaje, qué programas, qué población?), sino que resulta muy difícil saber a través de qué instrumento de recopilación sistemática de información han llegado a esa conclusión (no desde luego con ninguno de los utilizados en el proceso de evaluación). Efectivamente, no se encuentra el necesario tono (auto)crítico que debe tener cualquier evaluación.

Sin embargo, y al igual que se valoraba la evaluación del I Plan central, creemos comprensible que la evaluación de un primer plan tenga un cierto tono triunfalista, ya que a través de estos primeros planes se realizan por vez primera una serie de acciones y porque supone un avance cualitativo y sustancial sobre la situación anterior. Además, insistimos en lo positivo de estas primeras evaluaciones, por el simple hecho de plantearse. De hecho, el IAM hizo una apuesta fuerte por la recopilación sistemática de información y por la evaluación. La complejidad de la evaluación del primer plan,

¹⁹ El fracaso mencionado se refiere a unos asentamientos de cooperativas de mujeres, que habían previsto el IAM y el Instituto Andaluz de Reforma Agraria para explotaciones agrarias de Huelva y Sevilla, y que no llegaron a producirse (señalado en Granados, 1997: 30, en nota a pie de página).

²⁰ El listado es el siguiente: 1º Ha alcanzado una elevada notoriedad, y sus principales programas han sido conocidos por un alto porcentaje de la población. 2º Ha alcanzado un alto grado de cumplimiento en general, habiendo generado incluso actuaciones importantes no previstas inicialmente. 3º Ha supuesto un gran esfuerzo de colaboración interinstitucional, lo que constituye un modelo positivo de actuación para distintos programas sociales. 4º Ha movilizó importantes recursos en varios de los programas. 5º La relación costo económico-resultados es muy favorable a causa de la implicación voluntaria de numerosos agentes y el aprovechamiento de medios ya existentes. 6º Ha generado un mayor conocimiento de la situación social de la mujer en todas las áreas, lo que puede servir para planificar políticas ajustadas en el futuro y 7º El IAM ha resultado un buen instrumento de coordinación y dinamización del Plan.

que supuso además un enorme esfuerzo interno en el IAM, les hizo conscientes de la necesidad de concretar los objetivos y actuaciones y de ser algo más realistas. Así, la obsesión en el II Plan era, por un lado, hacer un II Plan más sencillo y realista que el I (con menos medidas y más concretas) y, en segundo lugar, realizar una evaluación totalmente externa del II Plan y que no supusiera tanto esfuerzo interno para el Instituto. En cualquier caso, el II Plan incorporó las recomendaciones que en la evaluación del I Plan se hacían, por lo que dicha evaluación cumplió una de sus funciones más importantes.

Evaluación del I Plan de Acción Positiva para las mujeres en la Comunidad Autónoma de Euskadi. 1991-1994

Finalmente, la evaluación del **primer Plan de Acción Positiva para las Mujeres de Euskadi (PAPME)** es el primer esfuerzo de evaluación más continuo y completo, que se define como una prioridad del Instituto Vasco de la Mujer (*Emakunde*) y en el que se emplean bastantes esfuerzos y recursos. Aunque la evaluación del I PAPME es el primer intento o experiencia en *Emakunde*, es, en realidad, el comienzo de una verdadera política de evaluación, planteada de forma consciente y rigurosa desde un principio. El planteamiento de una metodología de evaluación ya para el I PAPME fue una de las “preocupaciones prioritarias de *Emakunde* ya que atribuía gran importancia a la evaluación, pues sólo a partir de ella se podía tener líneas futuras de actuación eficaces para la consecución de una igualdad de oportunidades real, objetivo estratégico del PAPME”²¹. Además del rigor y de la dedicación de *Emakunde* a la función de evaluación, los procesos evaluativos están muy bien documentados en memorias e informes, en los que se reflejan no sólo los resultados de la evaluación, sino también el planteamiento y el proceso de evaluación, incluidas las dificultades encontradas a lo largo del mismo. Más allá, como se verá en el análisis de las segundas evaluaciones, existe una continuidad entre la evaluación del I y el II PAPME, habiéndose ido puliendo los procesos de evaluación. Aquí apuntamos pues, de forma general, lo que supuso la evaluación del I PAPME y se analizará en mayor profundidad los procesos de evaluación con la evaluación del II.

Para realizar la evaluación se recurrió al trabajo conjunto de *Emakunde* y de personas expertas externas de la Universidad Autónoma de Barcelona (Profesora Judith Astelarra) y de la Universidad del País Vasco (Profesora Pilar Pérez Fuentes). Por lo tanto, aunque se debe considerar como un proceso de evaluación esencialmente interno, contó con colaboración externa. El formato de esta colaboración fue a través de seminarios en los que se diseñó y planificó el proceso de evaluación.

A lo largo del I PAPME, se fueron elaborando memorias de actividad (*Emakunde* presenta al Parlamento Vasco una memoria anual de evaluación de la actuación de los poderes públicos) que se realizaron siguiendo una metodología de evaluación que pudiera ser utilizada posteriormente en la evaluación de la ejecución de todo el Plan. Y al final del período de vigencia del mismo se presentó un informe final de 138 páginas. En este informe se refleja, en un primer capítulo, la filosofía y los aspectos

²¹ “Informe de Evaluación del I PAPME” *Emakunde*, pág. 8.

metodológicos de la evaluación, y en un segundo capítulo, los resultados, organizados en los tres ejes fundamentales que cuenta la evaluación: 1) Actividades realizadas, 2) Colectivos destinatarios y 3) Participación institucional. Finalmente, se realiza un capítulo de conclusiones, en las que se realiza un balance de los resultados y se señalan unas consideraciones finales.

En el planteamiento metodológico que se hace de la evaluación del I PAPME, se renuncia explícitamente a hacer una evaluación del impacto de la política, tanto por la complejidad técnica debida a la multiplicidad de ámbitos, agentes, áreas temáticas, objetivos y formas de intervención, como por el impacto presupuestario que eso podía suponer. Por lo tanto, proponen una evaluación de la propia política pública, es decir de las ventajas y desventajas que tiene la estrategia seguida, de las características específicas de las acciones desarrolladas y de las instituciones encargadas de la ejecución de las políticas y acciones en cuanto a su capacidad de implementación. Esto significa que habían profundizado y existía una reflexión seria ya desde un principio sobre el papel, los fines y la función de la evaluación. También demuestran en el documento un empeño en distinguir entre la realización de una memoria de actividades (que es lo que consideran que se ha hecho en otras evaluaciones de Planes de Igualdad, cuando se analiza el grado de cumplimiento del plan, en el sentido de comprobar si las acciones propuestas habían sido realizadas o no) y una evaluación de las políticas y acciones de los planes.

Las responsables de *Emakunde* son también conscientes de la generalidad de la formulación del PAPME (ya que en él no se determina cuáles son los resultados que se espera conseguir), característica que comparten con el resto de los primeros planes de igualdad. Entre otras razones, por eso en esta primera evaluación se trataba no sólo de saber qué acciones habían sido desarrolladas y cuáles no, sino de tener otra información sobre el contenido y naturaleza de las actuaciones.

Uno de los trabajos más a fondo realizados en este proceso de evaluación por parte de *Emakunde* es la categorización de actividades que establecen. Construyen una tipología de actividades, tomando en cuenta dos criterios: 1) La forma de actuación de que se trata, es decir, qué se pretende y a quién se dirige, y 2) el contenido concreto de la actividad en cuanto al área social o la temática que aborda. El primer criterio permitió el establecimiento de nueve tipos de actividad y el segundo, los subtipos de actividades dentro de cada uno de estos nueve tipos generales. Los nueve tipos eran²²:

1. Cultura/ocio (con 4 subtipos)
2. Información/asesoramiento general (5 subtipos)
3. Sensibilización (8 subtipos)
4. Creación de conocimiento de género (3 subtipos)
5. Formación dirigida a los profesionales (3 subtipos)
6. Acción directa ante la discriminación (4 subtipos)
7. Fortalecimiento de asociaciones de mujeres (5 subtipos)

²² Para una mayor especificación véase *Informe de Evaluación del I PAPME*

8. Introducción temática de género en administración pública (7 subtipos)

9. Cooperación Internacional (2 subtipos)

A su vez, estos nueve tipos han sido reagrupados en tres niveles en función de su estrategia general de actuación:

- 1) Actuar sobre el conocimiento, informando, sensibilizando y formando a las personas.
- 2) Actuar directamente en contra de la discriminación y en favor de la igualdad real entre mujeres y hombres.
- 3) Crear organizaciones y estructuras para abordar las políticas antidiscriminatorias: instituciones públicas y organizaciones privadas.

Los instrumentos de recopilación de información utilizados fueron: 1) Un cuestionario sobre actividades anuales, dirigido a toda la Administración Pública Vasca. Esta información se codificaba en fichas y se trataba informáticamente para su análisis posterior. Los informes presentados anualmente durante la vigencia dle I PAPME se realizaron basándose en dicha información. Tuvieron dificultades en la recopilación de información (no se obtenían respuestas a muchas preguntas y se producían retrasos que hacían difícil su posterior tratamiento). A lo largo de la vigencia del I PAPME se va perfeccionando el sistema de codificación y tratamiento estadístico de la información, así como la forma de detección y corrección de posibles errores. La segunda herramienta fue 2) Una encuesta específica enviada a toda la Administración Vasca en 1993, para obtener información sobre la forma en la que se percibía su compromiso con las políticas de igualdad de oportunidades entre mujeres y hombres y sobre el grado en que éstas habían sido incorporadas al quehacer habitual de cada una de las instituciones, con el fin de conocer si se había creado, o no, una dinámica de acción propia en cada una de ellas capaz de desarrollar una acción sostenida en el futuro²³. También con este instrumento tuvieron dificultades, ya que tuvieron una menor respuesta por parte de los organismos que la referida al correspondiente cuestionario sobre realización de actividades. Sin embargo consideran que la información es útil y se incorpora al análisis de la participación institucional.

La presentación de resultados es exhaustiva y se organiza en los tres ejes básicos de la evaluación: actividades, destinatarios y participación institucional. El tratamiento de estos tres ejes fue sistemáticamente cruzado a fin de conocer su interrelación con las categorías analíticas creadas, de modo que se da un perfil del tipo de intervención pública en políticas de igualdad de género que se ha llevado a cabo y de las tendencias establecidas al respecto en el País Vasco. El total de actividades de las que se recibe información y que son analizadas asciende a 1237.

²³ Como se argumentará en el capítulo 5, y aunque en *Emakunde* se considere que esto no tiene que véase con una evaluación de impacto, esta información permitiría en parte, a nuestro parecer, una verdadera evaluación del impacto de las políticas de igualdad entre los géneros. Si el objetivo último de dichas políticas es la perspectiva de género en otras políticas públicas, la búsqueda de los efectos de las políticas de igualdad debe hacerse en las otras políticas públicas y su incorporación de la perspectiva de género, además de analizar el impacto de cada una de las acciones concretas.

Finalmente, queremos señalar no sólo el espíritu autocrítico que impregna los documentos que reflejan el proceso de evaluación, apuntando los avances pero también los límites y las dificultades de dicho proceso, sino la característica de proceso continuo del mismo, característica que le hace adquirir una mucha mayor potencia que los correspondientes procesos de evaluación en los niveles central y autonómico. Respecto a la utilidad y utilización, además, no sólo parece que el proceso de evaluación arrojó unos resultados útiles sino que también el propio proceso fue útil para generar y desarrollar la función de evaluación en las políticas vascas de igualdad de género.

2.2. La segunda generación de evaluaciones

Como ya se ha comentado al principio de este punto, esta segunda generación de evaluaciones, es el objeto pormenorizado de la metaevaluación que se realiza en este trabajo. A partir del año 1991 se generalizan los planes de igualdad en las comunidades autónomas (de seis planes que había aprobados en el año 1990, se pasa a once en el año 1991 y va aumentando gradualmente hasta el año 1995, en el que todas menos una de las comunidades autónomas habían ya aprobado planes). Esto provoca un “efecto rebote” de evaluaciones de dichos planes a partir del año 1995.

Los procesos de evaluación de planes de igualdad de género que aquí se analizan son los que han tenido lugar a partir de 1995 y hasta 1999, en los niveles autonómico y central. Los **criterios de selección**, además del criterio temporal, han sido amplios, recogiendo cualquier experiencia de evaluación de los planes, independientemente del tipo de evaluación realizado, su alcance o el tiempo y los recursos empleados. Se consignan la experiencia de evaluación del II Plan central y las experiencias de evaluación de 7 segundos y 3 primeros planes autonómicos²⁴, que son las evaluaciones siguientes:

- II Plan de Igualdad de Oportunidades de las Mujeres 1993-95. IM. Realizada en 1995-1997. Evaluación mixta. (evaluación interna con alguna colaboración externa).
- II Plan Andaluz para la Igualdad de las Mujeres 1995-97. “Las Mujeres en Andalucía: Estrategias para avanzar” *Instituto Andaluz de la Mujer*. Realizada en 1997-99. Evaluación externa.
- (I) Plan de Acción Positiva para las Mujeres en Aragón 1994-96. *Instituto Aragonés de la Mujer*. Realizada 1996-97. Evaluación interna.
- I Plan de Igualdad de Oportunidades 1995-96. *Instituto Canario de la Mujer*. Realizada en 1996. Evaluación externa.
- II Plan de Igualdad de Oportunidades: “Las Mujeres de Castilla-La Mancha hacia el siglo XXI” 1995-1999. *Dirección General de la Mujer*. Realizada en 1999. Evaluación externa.
- (I) Plan Integral de Igualdad de Oportunidades para la Mujer en Castilla y León 1994-96. *Secretaría Regional de la Mujer*. Realizada en 1996-97. Evaluación externa.

²⁴ No se incluye la evaluación realizada en Asturias del II Plan de Igualdad, ya que ésta es una evaluación descriptiva de constatación de las acciones única y exclusivamente a través del gasto presupuestario. La razón de la evaluación es, además, saber “dónde se estaba” por parte del nuevo equipo que llega a la Secretaría Regional de la Mujer, con el cambio de gobierno (y de partido en el gobierno) en 1995. Las razones para haber incluido, por ejemplo, la evaluación del I Plan aragonés y no haber incluido la del II Plan asturiano en esta segunda tanda y la del I Plan extremeño en la primera, son claras: las evaluaciones extremeña y asturiana son dos acciones puntuales que no son inicio de una mínima política de evaluación. En el caso aragonés, la evaluación del I Plan, en la que se emplean ya bastantes esfuerzos, sirve como acicate para plantearse la evaluación del II Plan que es, además, inmediatamente aprobado.

- II Plan de Actuación del Gobierno de la Generalidad de Cataluña para la Igualdad de Oportunidades de las Mujeres 1994-96. *Institut Català de la Dona*. Realizada en 1996-97. Evaluación mixta (evaluación interna con alguna colaboración externa).
- II Plan para la Igualdad de Oportunidades de las Mujeres de la Comunidad de Madrid 1993-95. *Dirección General de la Mujer*. Realizada en 1995-1996. Evaluación externa.
- II Plan para la Igualdad de Oportunidades de las Mujeres Gallegas 1995-97. *Servicio Galego de Promoción da Igualdade do Home e da Muller*. Realizada 1996-1998. Evaluación externa.
- (II) Plan Integral de la Mujer 1996-99. *Dirección General de Bienestar Social del Gobierno de La Rioja*. Realizada 1997-99. Evaluación interna.
- II Plan de Acción Positiva para las Mujeres en la Comunidad Autónoma de Euskadi (1995-99). *Emakunde*. Realizada anualmente de forma continua. Evaluación mixta (evaluación interna con alguna colaboración externa).

Como ya se ha comentado anteriormente, estas experiencias de evaluación son muy heterogéneas entre sí en cuanto a tipos de evaluación realizadas, recursos invertidos, tiempo de duración, alcance y eficacia evaluativa. En realidad, para la mayoría de las comunidades autónomas estas evaluaciones representan sus primeras experiencias al respecto (salvo en el caso de Andalucía, Cataluña y el País Vasco, además del IM en el nivel central, que ya tenían experiencia previa). En el listado previo se han consignado sólo dos características: año de realización y tipo de evaluación según el agente evaluador, es decir si son internas, externas o mixtas. La intención es dar un panorama general previo antes de comenzar el análisis que se reflejará en el capítulo siguiente; por lo tanto, se ha considerado que estas dos características pueden ser las más ilustrativas para aprehender una visión global anterior al análisis pormenorizado. Precisamente, la distinción según el agente evaluador (evaluaciones externas, mixtas o internas) es la clasificación que se sigue más adelante para la exposición de los datos.

Las comunidades autónomas que no “entran” en esta segunda generación son: Asturias, Baleares, Cantabria, Extremadura, Murcia, Navarra y Comunidad Valenciana. O bien no se ha hecho ningún tipo de evaluación o se han llevado a cabo análisis muy débiles y someros (Asturias, Baleares, Murcia), o no se cuenta en ese momento con un plan evaluable, por ya lejano (Cantabria, Extremadura, Comunidad Valenciana) o porque no se había aprobado todavía ninguno (Navarra).

2.3. La tercera generación: Los proyectos de evaluación de los terceros planes o equivalentes

En la mayoría de los terceros (segundos en algunos casos y primero, en el caso de Navarra) se hacen menciones explícitas al seguimiento y la evaluación de los mismos. Esto ya marca una diferencia con respecto a la generación anterior de planes, en los que, como se ha visto en el capítulo 3 de este trabajo, estas menciones explícitas eran mucho menos frecuentes. A continuación, exponemos, de forma breve, cuáles son dichas menciones explícitas y cómo se plantean las evaluaciones al inicio y aprobación de los planes.

El **III Plan central (1997-2000)** se plantea la evaluación en términos prácticamente idénticos a los de la evaluación del II. Consistirá en un análisis de la implementación que permitirá su seguimiento anual y que se realizará a través del diseño de una ficha de recogida de información para cada actuación. “En estas fichas se dejará constancia de las actividades llevadas a cabo para ejecutar las actuaciones, en qué fechas se han realizado y los motivos, si hubiera lugar a ello, por los que no se ha puesto en marcha la actuación en cuestión. Y la evaluación, propiamente dicha, tendrá “dos vertientes”: 1) Una evaluación cuantitativa que consistirá en estudiar la evolución de la población femenina desde 1997 hasta el 2000, y que se llevará a cabo con datos secundarios y a través de unos indicadores objetivos referidos a demografía, vida familiar, educación, trabajo, salud, participación política y social, marginación y malos tratos y servicios sociales. 2) Una evaluación cualitativa, cuyo objetivo será conocer la valoración que hagan de la ejecución del II PIOM representantes de la Administración Pública Central y colectivos de mujeres destinatarios de algunas actuaciones del Plan. Se hará a través de entrevistas en profundidad con los representantes y de grupos de discusión con los colectivos, una vez haya acabado el período de vigencia del PIOM. Quizá una cuestión importante es que, por vez primera, en el III Plan se hace una compromiso formal respecto a la evaluación, y que se formula del siguiente modo:

Antes del 31 de diciembre del 2001, el Ministerio de Trabajo y Asuntos Sociales presentará al Consejo de Ministros un informe final sobre el seguimiento y evaluación del Plan, en el que se expondrá el grado de cumplimiento alcanzado, las actuaciones puestas en práctica y las causas que hubieran motivado la no realización de algunas de las actuaciones previstas.

Curiosamente, si se analiza el contenido del compromiso, se comprueba que se refiere al seguimiento y análisis de la implementación, no al resto de la evaluación planteada, es decir, para cumplir con lo que se dice bastaría con el análisis de la implementación y una valoración sobre el grado de cumplimiento y ejecución del Plan.

Si se tiene en cuenta lo planteado para la evaluación del III Plan, se puede aventurar la idea que el IM no ha extraído muchas enseñanzas de su experiencia evaluativa. Lo normal en los procesos de evaluación es ir ajustándolos gradualmente a

la realidad concreta y cambiante de los programas que evalúan de forma que se van modificando para resultar más útiles y con mayor calidad técnica. Sobre todo cuando, en este caso, el proceso de evaluación del II Plan tiene algunos elementos criticables y manifiestamente mejorables.

En el **segundo Plan aragonés (1997-2000)**, también se plantea un sistema de evaluación. En primer lugar se presenta un sistema de recopilación de información que está incluido en un sistema de seguimiento a través de grupos de trabajo y reuniones con los diferentes Departamentos e instituciones. El Instituto Aragonés de la Mujer plantea la realización de una evaluación interna continua, basada en la información recopilada a través del sistema de seguimiento. El Instituto se compromete a: 1) Realizar un informe anual de seguimiento del Plan que incluirá: “un análisis de los resultados obtenidos respecto a los objetivos planteados”, fichas de cada acción realizada o iniciada, una relación de las acciones pendientes, indicando motivos de la falta de realización, análisis de necesidades detectadas que no hayan sido recogidas en el Plan, posibilidades de realización de actuaciones no previstas y posibles fórmulas de reorientación o reprogramación, si se precisan. 2) Realizar unos informes finales en los años 2000-2001 en los que se recopile y analice la información recogida en los informes anuales de los cuatro años de vigencia del Plan, incluyendo entre otros: grado de consecución de objetivos y cumplimiento de las acciones, tipos de actividades realizadas, organismos e instituciones implicados en su ejecución, ámbitos territoriales de realización, temporalización, colectivos destinatarios, recursos humanos y materiales, presupuestos destinados, tanto desde el IAM como desde otras instituciones, impacto social e indicadores específicos. En el planteamiento del contenido de estos informes se puede percibir una cierta influencia del tipo de evaluación realizada por *Emakunde*. Por último se contempla la posibilidad de una evaluación externa a iniciar en el año 2000, “si se ve factible y necesario” y que será una “evaluación complementaria para profundizar y conocer el grado de asunción y la repercusión de la inclusión de las perspectivas de género en las políticas de los Departamentos de la Diputación General de Aragón, instituciones y entidades; así como para ahondar en el conocimiento y valoración de ejecución del Plan” (II Plan aragonés. p. 64). Por último, se señala también en este apartado el compromiso de conocer el avance en la situación social de las mujeres en Aragón, “actualizando parcial y permanentemente los datos del estudio ‘Mujer aragonesa en cifras’ en la medida que el acceso a los datos lo haga posible” y elaborando de nuevo la investigación en el año 2000. Es importante resaltar como se separa este estudio de la situación social de las mujeres de la evaluación propiamente dicha (cuestión que todavía no se diferencia en la evaluación de otros planes), aunque aparece como parte del sistema de seguimiento y evaluación, ya que estudios de este tipo pueden ser una fuente crucial de información para las evaluaciones, además de resultar objetivo prioritario de los organismos específicos de igualdad y ser otra base más para la planificación y formulación de políticas futuras.

Uno de los cinco objetivos del **III Plan asturiano (1996-2000)**²⁵ es el de “establecer medidas de seguimiento y evaluación de las acciones emprendidas para alcanzar los objetivos propuestos” y uno de sus principios básicos es la “evaluación permanente de las actuaciones realizadas, con el objeto de medir el progreso y reflexionar sobre los procedimientos para corregir o reforzar los planteamientos” (pág. 52). La estructura responsable para ello es la Comisión de Coordinación y Seguimiento del Plan, cuya presidencia la ocuparía la Consejería de Cultura, la Vicepresidencia la Secretaria Regional de la Mujer y en la que están representadas once Direcciones Regionales del Gobierno del Principado de Asturias. También al final de documento aparece un poco irregular y poco estructurado listado de “indicadores de evaluación” (pp. 73-74), aunque se señala que no se pretende “exponer con detalle todos los indicadores susceptibles de ser considerados en la evaluación, que tendrán que concretarse en los programas que desarrollen cada medida” (pág. 73). Sin embargo, este “compromiso” explícito con la evaluación del III Plan, reflejado a través de un objetivo específico y unas menciones en el propio Plan no parece haberse llevado a la práctica²⁶.

En el **II Plan de las Islas Baleares (1996-99)** no existe una mención explícita a cómo se va a realizar la evaluación. Sabemos que no se ha realizado ningún tipo de evaluación de este segundo plan²⁷, aunque en él si aparecen dos cuestiones relacionadas entre las 12 propuestas de resolución al II Plan de Igualdad de Oportunidades para las mujeres aprobadas por el Pleno del Parlamento de las Islas Baleares el 29 de noviembre de 1996:

El Parlamento de las Islas Baleares insta al Gobierno a dar cuenta cada año en el si de la Comisión de Asuntos Institucionales y generales del Parlamento, sobre el grado de cumplimiento del II Plan de igualdad de oportunidades para las mujeres 1996-9. (II Plan balear: p. 23. subrayado mío).

El Parlamento de las Islas Baleares insta al Gobierno de la Comunidad Autónoma a presentar una memoria valorando los resultados del I Plan de la mujer. (II Plan balear: p. 23. subrayado mío).

²⁵ La publicación de este plan es del año 1998 y en ella no se hace referencia ninguna a la aprobación de este Plan en Consejo de Gobierno ni aparece ninguna fecha de acuerdo sobre el mismo. Simplemente se menciona que “en el seno de la Comisión de Seguimiento para la Igualdad de Oportunidades de las Mujeres, constituida en Enero de 1996, se ha desarrollado este III Plan...” (pag. 7).

²⁶ Al igual que en Baleares, en esta C.A. se ha dado un cambio de partido en el Gobierno en la última legislatura y también se ha creado una nueva estructura en 1999, el Instituto Asturiano de la Mujer, aunque en este caso no es un organismo autónomo, sino un “órgano desconcentrado”, creado por decreto y dependiente de la Consejería de Presidencia.

²⁷ La información que obtuvimos de la Comisión Interdepartamental en Febrero de 2000 es que no se había realizado ninguna evaluación del II Plan y se estaba a la espera de la creación del Instituto Balear de la Mujer. En ese momento ya existía un proyecto de ley aprobado por el Parlamento.

Los mecanismos para el seguimiento y la evaluación que aparecen en el **II Plan Canario (1997-2000)**, aunque se relacionan los criterios para el seguimiento de la implantación del Plan, estos son poco concretos. Los criterios no están muy especificados, los “instrumentos de evaluación” que se mencionan son diferentes estructuras (Consejo asesor del Instituto Canario, Comisión de Igualdad, comisiones asesoras de área y movimiento asociativo de mujeres), sin especificar realmente mecanismos, ni formato ni periodicidad. No se adquiere, además, ningún compromiso al respecto.

Se dice que el seguimiento de la implantación del Plan tiene “dos vertientes: orgánica y factual” 1) Desde la orgánica se debe evaluar:

- el grado de compromiso de los responsables públicos y la participación de la iniciativa privada y asociativa
- la funcionalidad de los acuerdos y convenios alcanzados y su plasmación en colaboraciones concretas
- la creación de redes de coordinación con los recursos ya existentes
- la sensibilización de los responsables públicos y la sociedad con las políticas de igualdad

2) Y en lo que respecta a las “realizaciones objetivas” se deberá evaluar en términos de:

- participación y alcance poblacional
- colaboración con otros organismos y entidades en acciones comunes`
- conocimiento y diagnóstico del medio, los problemas que afectan a las mujeres y los perfiles que corresponden a cada situación
- grado de integración en la vida comunitaria y de utilización de los servicios que pueda ofrecer el organismo
- utilización de la red de recursos por las mujeres y de la demanda de nuevas acciones o servicios

En el borrador del **II Plan Cántabro**, no aparece ninguna referencia a la evaluación.

En el **III Plan de Castilla-La Mancha (1999-2003)**, se especifica la coordinación, seguimiento y evaluación del plan de forma pormenorizada. En primer lugar se señala que “todo plan exige la existencia de un diseño evaluativo desde la puesta en marcha y hasta su finalización”. La evaluación que se propone combina dos modalidades: interna y externa que se incorporará en la fase final. Se señalan objetivos de la evaluación, características, qué evaluar, la estructura organizativa y contenido del proceso de evaluación.

Respecto a las características de la evaluación, se contemplan tres fases de la misma: inicial, proceso y final. La evaluación inicial se identifica con una evaluación del contexto y que se plantea está realizada con la evaluación del Plan anterior. La evaluación del proceso se identifica con el seguimiento y la final con una evaluación de

resultados. Es decir, no parece que se contemple una evaluación sumativa que tenga por objeto, además de los resultados, el propio diseño del plan y su proceso de puesta en práctica. Ni tampoco que se vayan evaluando resultados de forma periódica y continua.

Respecto a qué evaluar, distinguen entre resultados, que hacen referencia a la ejecución de cada actividad (más en el sentido de productos o *outputs*) y que son “fundamentalmente de tipo cuantitativo”, e impacto en la sociedad, y que se define como los cambios experimentados globalmente y por grupos destinatarios (más en el sentido de efectos o *outcomes*) y suponen obtener y analizar informaciones de carácter cualitativo. No encontramos mucho sentido a asignar un carácter cuantitativo a unos y cualitativo a otros, ya que, tanto uno como otro podrían (y deberían) ser contestados con los dos tipos de información. También se menciona el desarrollo, de tipo organizativo y que incluye la coordinación, la colaboración entre entidades, las dificultades encontradas, la eficacia y la eficiencia resultante. Y por último, la cobertura, que la entienden de forma amplia ya que se trata, además de si las actuaciones previstas han llegado a la población a la que se dirigían, también si se han cubierto las necesidades y demandas a las que se trataba de dar respuesta (este último extremo necesita de información que va mucho más allá de la propia cobertura).

El II Plan castellano manchego también propone una estructura organizativa que se encargue del diseño, puesta en marcha y desarrollo de la evaluación. Este es uno de los pocos planes que especifica este aspecto tan importante. “Dicha estructura organizativa ha de coordinar las actuaciones previstas y la recogida de información sobre sus resultados e implementación, y organizarse en grupos de trabajo o comisiones de seguimiento con responsabilidad en las materias propias de cada grupo”. Se realiza una tabla en la que se especifica cada una de las estructuras, con sus objetivos y el contenido de trabajo de cada uno de ellos. Así a la Comisión de Igualdad de Oportunidades le corresponde, entre otras cosas, el diseño del proyecto de evaluación, y se apoya en el Consejo Regional de la Mujer, cuya misión consiste en realizar el seguimiento y evaluación de las actuaciones del Plan (a la Comisión de Igualdad le corresponde el seguimiento y evaluación de forma global) y que se apoya en la Comisión Técnica de Seguimiento, que realiza un asesoramiento más técnico y que se apoya en los grupos de trabajo que son los que hacen una tarea más ejecutiva, recogiendo información y emitiendo informes.

En el contenido del proceso de evaluación se propone una tipología de 12 grandes tipos de actuación:

1. Creación de organismos y centros
2. Estudios/análisis
3. Elaboración de materiales
4. Ayudas económicas
5. Formación
6. Convenios
7. Campañas -difusión, sensibilización, información, prevención,..
8. Actuaciones
9. Redes
10. Programas
11. Infraestructura
12. Equipos

Sobre cada uno de estos tipos se fijan una serie de indicadores mínimos, aunque se explicita que la fijación de los mismos está en función del objeto de la evaluación. Para definirlos se proponen puntos de referencia como los colectivos destinatarios, el tipo de programa o actuación, la cooperación con otras entidades o asociaciones, las estrategias de desarrollo y el presupuesto.

En el **III Plan catalán (1998-2000)** se menciona un apartado sobre la evaluación. En él, se menciona un seguimiento de su ejecución o implementación y una evaluación de impacto. El **seguimiento** se realizará a través de memorias anuales que recojan el grado de ejecución de las acciones, estableciendo las acciones realizadas a término, las fechas de realización y el presupuesto de las mismas. Se señala también que la **evaluación del impacto** se realizará tanto de forma cuantitativa como cualitativa, reservándose la cuantitativa para conocer la “evolución positiva o negativa en todos los campos en los que incide el Plan” (pág. 28) y la cualitativa para “todo aquello que tiene que ver con las actitudes, los valores y los estereotipos”. A esta evaluación cualitativa se le da especial importancia ya que el “III Plan está dirigido especialmente a la sensibilización y al establecimiento de un nuevo contrato social entre hombres y mujeres” aunque también mencionan que “hay que ser conscientes que las medidas encaminadas a modificar actitudes, conductas, valores y el substrato cultural, tienen siempre resultados a medio o largo plazo” (pág. 29). Se señala que esta evaluación se realizara antes de acabar el año 2000, una vez haya finalizado el período de implementación y que la metodología que se empleará será básicamente de entrevistas en profundidad, encuestas de opinión y estudios comparativos.

Aunque en el **III Plan madrileño (1997-2000)** no se hace ninguna mención explícita a su evaluación, fuentes de la Dirección General de la Mujer de la Comunidad de Madrid señalan que la evaluación de este tercer plan se plantea de igual modo a la evaluación del II y que es objeto de la metaevaluación que se plantea en este trabajo. Curiosamente, aunque la evaluación del II Plan ha recibido muchas críticas (como veremos más adelante, también nosotros mantenemos la tesis de que no se trata de una evaluación del II Plan, sino de un estudio de la situación social de la mujer madrileña durante los años de vigencia del Plan), la Dirección General de la Mujer no parece haberse planteado grandes modificaciones al respecto.

En el **II Plan de la Región de Murcia (1997-2000)** se contempla un apartado específico sobre la evaluación del Plan (pp. 77-78). En él se explicita que habrá una evaluación anual y una a la finalización del mismo y se establecen algunas condiciones para un sistema de recogida de información que sirva “para conocer en cualquier momento: a) el estado de ejecución de cada una de las acciones, b) las actuaciones concretas desarrolladas y c) el modo concreto de ejecución. También se señala que la recogida de información se hará a partir de modelos específicos diseñado para cada una de las acciones, aunque en cualquier caso deberán recoger los aspectos de: identificación de área y acción, descripción de las actuaciones, tiempos empleados,

entidades participantes y sus funciones, presupuesto, colectivo al que va dirigida la acción y colectivo que realmente ha participado o se ha beneficiado de la misma.

También en el **I Plan de actuación navarro (1998-2000)** existe un apartado, por un lado, de aplicación y seguimiento del Plan (pp. 68-69), y de evaluación (p. 70) por otro. En el apartado de seguimiento y aplicación se especifica que se hará un sistema de recogida de información a través de una ficha, en la que se dejará constancia, para cada actuación, de las actividades ejecutadas, fecha de realización y los motivos por los que, en su caso, no se haya puesto en marcha la actuación en cuestión. Se señala que la aplicación, seguimiento y control del Plan se efectuará a través de una Comisión de Seguimiento, en el que estarán representadas los diferentes departamentos gubernamentales implicados en el Plan, además de una persona representante del Consejo Navarro de la Mujer. También se menciona que en el primer trimestre de cada año se realizará un informe sobre dicho seguimiento. En el apartado sobre evaluación, se dice que “el plan se evaluará atendiendo los siguientes puntos de vista”: 1) Evaluación de objetivos y acciones, analizando el grado de cumplimiento del Plan (se hará mediante el seguimiento antes explicado); 2) Evaluación de la participación institucional, para conocer el grado de participación de los diferentes organismos implicados en la ejecución de las acciones y 3) Opinión de la población destinataria, mencionando que al finalizar el plan se recogerán opiniones de las representantes de la población destinataria y de los agentes implicados en el mismo, mediante grupos de discusión y entrevistas. Finalmente, se señala que en el primer trimestre del año 2001 el INAM presentará al Gobierno un informe final sobre el seguimiento y evaluación del Plan.

Por último, en el **II Plan valenciano (1997-2000)** no existe ninguna mención explícita a su evaluación, aunque en Junio de 1997 se está pendiente de la firma de un convenio con la Universidad de Valencia para llevar a cabo un proyecto de evaluación del II Plan, convenio del que no hemos recibido información posterior²⁸.

Sin entrar en un análisis profundo de esta tercera generación de evaluaciones, ya que no es el objeto de este trabajo, sí se puede hacer una valoración global, y extraer unas conclusiones de carácter general, que se presentan a continuación.

Es evidente que una primera conclusión positiva es que en esta tercera generación de planes, la evaluación está mucho más presente de lo que estaba en las generaciones anteriores. Salvo en alguna excepción, prácticamente todos los planes hacen alguna mención explícita a su seguimiento y evaluación. Esta preocupación por la evaluación, además, acarrea también la realización de planes más específicos y realistas, es decir, más “evaluables”; en definitiva, sólo el hecho de pensar en términos de evaluación, hace no sólo que los objetivos y actuaciones sean más concretos y realistas, sino también que,

²⁸ Información extraída de una carta de 25 de Junio de 1997, firmada por la entonces Directora General de la Mujer de la Generalidad Valenciana (Ester Fonfría), en la que se señala que todavía no está firmado dicho convenio y que nos será enviado tan pronto esté debidamente formalizado.

en general, los planes sean más técnicamente ajustados en términos de presupuestos, asignación de responsabilidades y resultados esperados.

Sin embargo, también detectamos que muchos de los problemas o debilidades encontradas en la segunda generación de evaluaciones (y que se discuten con mayor profundidad en los capítulos siguientes de este trabajo), persisten aún en esta tercera generación. En general, destacaríamos tres:

1) El hecho de que exista un compromiso explícito de evaluación no asegura que de hecho se dé dicha evaluación. Este es el caso, por ejemplo, de Asturias y Baleares.

2) Cuando se plantean evaluaciones externas sumativas, se les da un claro carácter finalista. Es normal que estas se planteen temporalmente al final de los años de vigencia del plan a evaluar, pero el hecho que se espere al final para parte de la recopilación, así como para el análisis y la valoración, no debería implicar que el diseño de esa evaluación también se deje para el final, como de hecho ocurre en la mayoría de los casos. Este empezar a pensar en la evaluación al final hace, inevitablemente, que la potencia y las posibilidades de la evaluación sea mucho menor, incluso a efectos de tener pensado un esquema de recopilación de información a lo largo del período de vigencia del plan, lo cual permitiría una evaluación sumativa final de mucha mayor calidad. Esto puede ser debido, en parte, al hecho de considerar estas evaluaciones finales con carácter externo; sin embargo, esto no debería ser una razón, ya que se podría contratar un primer asesoramiento externo al principio del Plan para diseñar el estudio de evaluación que se realizara al final, incluso un pequeño asesoramiento en la propia recopilación de información, además de contratar el propio estudio al finalizar el Plan.

3) Se siguen reflejando problemas de fondo en la concepción de la evaluación, y que vienen dadas por la manera en la que se han concebido y realizado las primeras experiencias de evaluación. De la misma manera que se ha comentado la importancia de estas primeras experiencias por “romper el fuego” en evaluación y porque la mejor manera de avanzar es “aprender haciendo”, sin embargo, también es necesario señalar que estas primeras experiencias han creado un cierto mimetismo a la hora de plantearse los procesos de evaluación. Por ejemplo, se sigue definiendo la evaluación de los procesos (o de la implementación como se denomina en algunos casos) como la simple evaluación de los productos (es decir, la comprobación de que se hace lo que se pretende hacer), dejando de lado una verdadera evaluación de procesos en la que se estudie, por ejemplo, el funcionamiento de la transversalidad y la implicación de los diferentes responsables. Otro ejemplo sería cómo se sigue hablando de “evaluación cuantitativa” y de “evaluación cualitativa”, distinguiendo falsamente dos tipos de evaluación dependiendo de las técnicas de recopilación de información a utilizar; cuando, en todo caso, habría que distinguir sobre qué cosas se evalúan y no sobre las técnicas a evaluar y cuando precisamente la utilización “dialéctica” de diferentes técnicas es lo que permite enriquecer los procesos de evaluación. Como veremos más adelante, creemos que poner el acento en las técnicas de recopilación y análisis de la información significa que, en cierto modo, falla el paso previo de definir claramente qué se quiere evaluar y dar entonces y sólo entonces, una respuesta técnica adecuada para ello. Creemos, por lo tanto, que, a pesar de los indudables avances, algunas de estas deficiencias o debilidades siguen persistiendo en esta tercera generación de evaluaciones.

3. EL MARCO PARA LA METAEVALUACIÓN: LOS CRITERIOS PARA EL ANÁLISIS VALORATIVO DE LAS ONCE EXPERIENCIAS DE EVALUACIÓN

Ya se han comentado anteriormente los **criterios de selección** de los casos que incluimos en esta metaevaluación: principalmente se trata de un criterio temporal (evaluaciones realizadas entre 1995 y junio de 1999 y completadas antes de dicha fecha). También se utiliza un criterio técnico muy laxo sobre el concepto de evaluación, es decir, a no ser que haya sido una experiencia muy puntual y aislada en el tiempo (como es el caso de la evaluación del II Plan asturiano), se han incluido como objeto de la metaevaluación todas las experiencias que han sido autodenominadas como evaluación por parte de sus respectivos organismos responsables.

Esto significa que las experiencias de evaluación que estudiamos en esta segunda generación son muy heterogéneas entre sí y, esta circunstancia, va a hacer difícil la comparación en algunos casos. Aún así, al ser una metaevaluación de **once** procesos de evaluación, se da prioridad precisamente a la comparación entre los casos sobre el análisis en profundidad de cada uno de ellos.

Como en toda evaluación, se pretende especificar para esta metaevaluación, el propósito y objetivos de la misma, los criterios de (meta)evaluación y los instrumentos para la recopilación y análisis de la información que se han utilizado. El **propósito** de la metaevaluación es contribuir al estudio de la evaluación de las políticas públicas de igualdad de género y los **objetivos** de la misma son, por un lado, analizar la función de evaluación en torno a dichas políticas en los niveles autonómico y central en España y, por otro, proponer alternativas para la mejora de dicha función. El objeto de la metaevaluación son los procesos de evaluación en sí mismos, no los resultados que arrojan. Por lo tanto se plantean tres grandes **criterios de (meta)evaluación**: 1) ¿Cómo se han planteado las evaluaciones y cuáles han sido las estrategias evaluativas planteadas? Para esto analizamos 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 realizados y los recursos empleados en los procesos de evaluación. 2) ¿Cuál ha sido el papel de algunos elementos clave en la evaluación?, centrándose el estudio en los agentes y los participantes, los momentos y los tiempos, los criterios e indicadores, así como los procedimientos e instrumentos de recopilación y análisis de la información utilizados. Por último, 3) ¿Cuál ha sido la utilización de las evaluaciones?, analizando 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.

A continuación, se especifican los criterios utilizados en la metaevaluación. Estos se organizan en un esquema para el análisis y la valoración que, como ya se ha comentado, se divide en tres grandes partes o dimensiones: el planteamiento de las evaluaciones, los elementos clave y la utilización e impacto de las mismas.

3.1. El planteamiento de la evaluación y las estrategias planteadas

En el análisis del planteamiento de una evaluación, resulta crucial analizar el contexto evaluativo en el que se generan, ya que ésta es una dimensión que ayuda a entender las respuestas evaluativas. En segundo lugar, se analizan, como parte de dicho planteamiento, los propósitos de cada uno de los procesos de evaluación, su claridad y coherencia con respecto a cómo se plantean dichos procesos. Además, se estudian el tipo de estructuras a las que se asigna la función de evaluación. Por último, y en lo que se refiere a la estrategia general de los procesos de evaluación, se señalan los tipos de evaluación realizados y los recursos empleados en las mismas.

El contexto evaluativo

Existen dos factores clave para entender el contexto en el que se genera esta segunda generación de evaluaciones: la generalización de los planes de igualdad, reforzados entre otros factores, por el panorama internacional. Y, en segundo lugar, los cambios de gobierno en algunas de las comunidades autónomas.²⁹

En el análisis del contexto evaluativo se tienen en cuenta las siguientes variables: organismo responsable, tipo y antigüedad del mismo; departamento al que está inscrito dicho organismo; partido en el gobierno en el momento de la evaluación y cambios producidos en dicho período; si han existido cambios en las responsables de los organismos de igualdad que han hecho y/o encargado la evaluación. Por último se ha tenido en cuenta la historia respecto a los planes y respecto a la evaluación de los mismos (es decir, si existía experiencia evaluadora previa).

La claridad en los propósitos y funciones de las evaluaciones

Uno de los factores más importantes a analizar en cualquier evaluación es la claridad o ambigüedad del planteamiento de dicha evaluación a través de sus propósitos y de las funciones que se espera cumplan dicha evaluación. Normalmente estos propósitos suelen quedar explicitados o bien en pliegos de condiciones técnicas, si el estudio de evaluación sale a concurso público, o bien en los proyectos de evaluación si es por convenio o adjudicación directa; también suelen explicitarse, o al menos debería hacerse, en los propios informes de evaluación.

Las estructuras institucionales de la evaluación

²⁹ Curiosamente, una vez asumidas las políticas de igualdad de género por todos los gobiernos autonómicos (incluidos los de signo más conservador, a partir de 1995) como un terreno propio, el cambio de gobierno en algunas CC.AA. ha resultado, en cierto modo, un factor facilitador de la evaluación, ya que las nuevas responsables necesitaban saber y aprehender que es lo que se había hecho hasta entonces.

En general, la existencia de un cierto marco institucional facilita la evaluación. El marco institucional para la evaluación de planes de igualdad en España es todavía incipiente, aunque los procesos de evaluación de dichos planes están haciéndose crecientemente más explícitos. En este apartado se analiza cuál ha sido el marco institucional en la segunda generación de evaluaciones de planes de igualdad. Para este análisis del marco institucional se toman en cuenta los siguientes factores: por un lado, si se menciona y se hace un compromiso explícito de evaluación en el plan evaluado y, por otro, el tipo de estructuras creadas alrededor de los organismos o los planes de igualdad y que tienen alguna responsabilidad en el seguimiento y la evaluación de los planes. En tercer lugar, también se tiene en cuenta en este análisis la estructura funcional y técnica de los organismos específicos de igualdad, ya que sobre ellos recae el peso ejecutivo de los procesos de evaluación. Conocer las dimensiones de la estructura de recursos humanos de estos organismos en el momento en el que se realizan las evaluaciones analizadas, ofrece una visión más completa de las estructuras institucionales que soportan los procesos de evaluación, bien sean internos o externos.

Tipos de evaluaciones realizadas

Otro de los posibles análisis a realizar en cuanto al planteamiento de los procesos evaluativos es el tipo de evaluación realizada en cada caso. Aquí se van a utilizar las cuatro clasificaciones básicas desarrolladas en el capítulo 2 de este trabajo: según el papel que se quiere que cumpla la evaluación (evaluación formativa vs. sumativa), según el contenido o fase de la política que se evalúa (evaluación de diseño, proceso y/o resultados), según el momento en el que se evalúa (evaluación *ex-ante* o *post-facto*) y según el agente evaluador (evaluación interna o externa).

Este criterio de metaevaluación va a ser el que utilizemos en primer lugar, con el fin de clasificar las once experiencias metaevaluadas. Por lo tanto, la información recopilada y analizada sobre el mismo, aparece en este capítulo número cuatro, previamente a la clasificación de evaluaciones externas, mixtas e internas.

Recursos empleados en las evaluaciones

Otro factor importante para valorar el planteamiento que se hace de una evaluación, es sin duda, los recursos que en ella se invierten, ya sean económicos o humanos. En el caso de las evaluaciones externas, resulta fácil la ponderación de los gastos directos. Más difícil es la medición de los gastos indirectos, sobre todo en los procesos de evaluación interna. También en los procesos de evaluación externa es interesante medir los gastos indirectos, es decir, el esfuerzo invertido por el correspondiente organismo de igualdad en la coordinación con el equipo externo, más el tiempo interno empleado en aportar y recoger información, así como en recibir, revisar, ponderar, analizar y procesar los resultados de una evaluación.

La medición de estos gastos indirectos es, sin embargo, muy difícil hacerla de manera fiable si no se ha establecido previamente unos criterios y un sistema para la

recogida de información (tiempo empleado por cada trabajador/a en el proceso de evaluación) que permita consignar los datos de forma periódica y de una forma frecuente (como máximo semanalmente), ya que la estimación del tiempo dedicado a una tarea “a toro pasado” es muy subjetiva y depende de muchos factores. Si existiera personal técnico dedicado a tiempo completo y exclusivo a la evaluación es más fácil, pero éste no suele ser el caso, ya que la mayoría del personal técnico con responsabilidades en las tareas de evaluación en los organismos de igualdad se dedican, además, a otras tareas distintas de la evaluación. Estas son las razones por las que se ha tenido que renunciar, en este trabajo, a una medición, que resulte objetiva y fiable, de los gastos indirectos de las evaluaciones. Simplemente, y según cada caso, se harán algunas puntualizaciones de carácter cualitativo a dichos gastos indirectos.

Con el fin de realizar una ponderación de los recursos empleados en términos relativos, se ha tratado de obtener información no sólo de los recursos empleados en los procesos de evaluación, sino también de los recursos empleados en el plan evaluado así como los presupuestos del organismo de igualdad durante el período de vigencia del plan correspondiente. Esto ofrece una idea relativa del esfuerzo invertido en los procesos de evaluación de los planes de igualdad. En este caso tampoco ha sido posible conseguir la información en todos los casos, ya que en algunos de ellos, ni siquiera había un presupuesto general del plan evaluado. Sin embargo, en los casos en los que se han podido obtener dichos datos, los resultados son de lo más ilustrativos.

3.2. Los elementos clave de los procesos evaluativos

En este apartado se analizan cuatro elementos clave de los procesos evaluativos. En primer lugar, se estudian los agentes y participantes en la evaluación, distinguiendo a los responsables y agentes de la misma (ya sean internos o externos) y a los diferentes participantes en las diversas fases del proceso evaluativo. En segundo lugar, se hace un análisis de los momentos en los que se plantea y realiza la evaluación, así como su duración y calendario. En tercer lugar, se analizan los criterios de evaluación utilizados, así como los indicadores establecidos, sus fuentes y el proceso de establecimiento. Finalmente, se describen los procedimientos e instrumentos utilizados para la recopilación y análisis de la información.

Agentes y participantes en el proceso de evaluación

En primer lugar, se realiza un análisis de los equipos técnicos que han diseñado y ejecutado las evaluaciones. Cuando las evaluaciones son externas, el peso de la responsabilidad técnica de las mismas recae sobre equipos externos de evaluación. En otros casos, aunque son evaluaciones principalmente internas, en ellas intervienen también otros agentes externos.

Además de los equipos técnicos responsables de las evaluaciones, es conveniente hacer un análisis de quiénes han participado, de una forma u otra, en los procesos de evaluación. Por participación en los procesos de evaluación, entendemos la participación activa en alguna de sus fases (diseño del estudio y estrategia evaluativa, establecimiento de criterios, diseño de instrumentos, recopilación de información, análisis de información y valoración y establecimiento de recomendaciones). Es decir, que por el hecho de preguntar la opinión de determinados colectivos a través de una encuesta, una entrevista o un grupo de discusión, no significa participar en el proceso de evaluación, sino ofrecer información o, mejor, convertirse en información que será utilizada en dicho proceso sistemático de evaluación. Ni siquiera el hecho de que se pregunte a los agentes que recomendarían ellos/as hacer significa necesariamente la participación: que alguien participara realmente en la fase de valoración y establecimiento de recomendaciones de una evaluación implicaría que se le ofreciera toda la información sistematizada y analizada y se le pidiera no su opinión sino una valoración de dicha información y sugerencias a raíz de la misma.

Momentos de los procesos de evaluación

Como se ha visto en el capítulo uno de este trabajo, otro de los elementos clave de cualquier proceso evaluativo es el de los tiempos en los que se produce dicho proceso. Todas las cuestiones relacionadas con el ajuste temporal de los procesos de evaluación al contexto y a los tiempos del programa o política evaluada, son cruciales en la determinación de la capacidad de respuesta y utilidad de dichos procesos. Relacionado con este elemento temporal, se analizan cuatro aspectos en cada una de las evaluaciones estudiadas. En primer lugar, se relaciona el **momento** en el que se plantea la evaluación; para esto, se consigna en qué momento del período de vigencia del plan evaluado se plantea y comienza la evaluación del mismo. En segundo lugar, se analiza la **duración** de los procesos de evaluación desde su inicio hasta la entrega del informe final. En tercer lugar, se consigna cuánto **tiempo pasa entre la finalización del período de vigencia del plan y la entrega del informe final** y, por último, se analizan los **retrasos** sobre el calendario inicial de los procesos de evaluación.

Criterios e indicadores de evaluación: fuentes y proceso de establecimiento

El contenido de las evaluaciones, es decir, el QUE se evalúa, se concreta a través de los criterios e indicadores de evaluación que representan las variables estudiadas y la base lógica en la que se fundamentan los juicios de valor que se establecen posteriormente en el proceso de evaluación. Como se verá más adelante, en la práctica totalidad de las experiencias de evaluación analizadas, es muy débil la fundamentación de la concreta elección y definición de los criterios e indicadores utilizados. Aunque lo adecuado sería analizar cuestiones tales como cuál es el sentido de elegir los criterios, su explicación y definición y la justificación de por qué se eligen esos y no otros, este tipo de información no suelen aparecer en los informes de

evaluación, y nos ha resultado difícil conseguirla. En muchos casos, se obtiene la información sobre los criterios de evaluación, a través de las especificaciones técnicas de las técnicas de recopilación y análisis utilizadas o a través de los propios resultados de evaluación.

Procedimientos y metodología empleada respecto a instrumentos y técnicas de recopilación y análisis de la información

En este apartado, hacemos un análisis de los procedimientos utilizados para la recopilación y análisis de la información. De forma similar a lo que ocurre con los criterios de evaluación, los procedimientos utilizados no suelen estar profusamente fundamentados; es decir, difícilmente encontramos una justificación de por qué se elige dicho instrumento y no otro. Sí encontramos, en algunos casos, especificaciones detalladas sobre los instrumentos en sí: estas especificaciones, sin embargo, son siempre de carácter técnico y versan sobre las características intrínsecas de dichos instrumentos, no tanto sobre su valía o adecuación para la situación concreta en la que se aplica.

3.3. La utilización de los procesos evaluativos

La verdadera utilización de los procesos evaluativos requeriría de una investigación propia, de corte longitudinal, y que estudiara los factores de utilización a largo plazo, es decir, desde que se entrega el informe de evaluación y al menos durante los cinco años siguientes, algo que excede de la investigación aquí planteada. Como ya se comentaba en el capítulo 1 de esta tesis doctoral, el concepto de uso y utilización de la evaluación se ha ido ampliando en los últimos años: no siempre la utilización de los resultados es tan inmediata y directa -nada más entregar el informe de evaluación-. Por lo tanto, se puede concebir la utilización de la evaluación no sólo para una concreta toma de decisiones, sino también para un aprendizaje organizativo más difuso. Además de la utilización de los resultados de una evaluación, pueden también “utilizarse” los propios procesos de evaluación, haciendo que las organizaciones desarrollen su función de evaluación, una cultura evaluativa y un *know how* en evaluación, incluyendo la capacidad de respuesta ante diferentes situaciones evaluativas, la metodología y los procedimientos de evaluación, así como la propia utilización de resultados y el aprendizaje para situaciones futuras. El propio hecho de estar evaluando, haciendo a las organizaciones protagonistas y conscientes de su propio quehacer, parece tener efectos positivos en sí mismos.

A pesar de las limitaciones de abordar esta dimensión de la utilización en el marco de esta investigación, se introducen como aspectos a estudiar en la metaevaluación realizada, la pertinencia y utilidad de la información producida por las experiencias de evaluación, la comunicación de los resultados y los destinatarios de la misma y, finalmente, el propio impacto de los resultados y los procesos de evaluación. Es decir,

con la visión amplia sobre el uso y la utilización de la evaluación anteriormente mencionada, se pretende extraer la información existente -en algunos casos, muy limitada- en primer lugar, sobre el contenido de los estudios de evaluación y su pertinencia y utilidad, haciendo una valoración general por nuestra parte y recurriendo a las propias opiniones de las responsables políticas de los organismos de igualdad; en segundo lugar, sobre la forma en la que el estudio de evaluación y sus resultados se han comunicado y difundido posteriormente y, en tercero, sobre el impacto que han causado, especialmente si se han modificado cosas en el siguiente plan como consecuencia de la evaluación o se ha generado una cierta cultura evaluativa por la realización del estudio.

Pertinencia y utilidad de la información producida

En este apartado se realizan dos tipos de análisis: en primer lugar, se hace una valoración general sobre la **pertinencia de los contenidos del estudio** para el plan evaluado, para el organismo de igualdad protagonista y para la situación evaluativa concreta. Esta valoración sobre la pertinencia de la información (y que es el primer aspecto a tener en cuenta de la potencial utilidad de la evaluación) se hace teniendo en cuenta, sobre todo, dos factores: el alcance de la información (es decir, si se ha recopilado, tratado y analizado la información suficiente) y la globalidad de la misma (es decir, si permite hacer una valoración global del plan o sólo de aspectos concretos sobre el mismo). También se tienen en cuenta las opiniones vertidas por las responsables políticas entrevistadas sobre este aspecto.

En segundo lugar, se estudia el **grado de análisis y elaboración de conclusiones y la existencia y la calidad de recomendaciones en los informes de evaluación**, como resultado de la misma: en definitiva, se analiza el **grado de “traducción práctica” y aplicabilidad de los procesos de evaluación**, como condición facilitadora (y casi nos atreveríamos a decir que necesaria) para su utilidad. Ya se comentó en el capítulo 1 de esta tesis doctoral la importancia de las recomendaciones en los procesos de evaluación, no sólo como cuarto nivel de análisis (después de los resultados, sus interpretaciones y las valoraciones) sino también como vinculación formal entre la evaluación y su uso. Las recomendaciones, en una evaluación sistemática, nacen de los tres niveles anteriores y representan la verdadera orientación a la acción de los estudios de evaluación. Resulta evidente, pues, que un estudio de evaluación que resulta en una serie de recomendaciones cuidadosamente seleccionadas y rigurosamente extraídas del proceso sistemático de evaluación, tiene mayores probabilidades de ser una evaluación útil y utilizada.

Comunicación de los resultados y destinatarios de los mismos

En este apartado se pretende analizar el proceso de comunicación de la evaluación, como otro elemento clave para la utilidad y utilización de la misma. Uno de los primeros elementos a analizar es el **informe(s) de evaluación**, en qué consisten, lo que reflejan del proceso de estudio llevado a cabo y cuán completos son. Se parte de

la base de que un buen informe³⁰ de evaluación debe explicitar los propósitos del estudio, describir el contexto del plan evaluado y la situación evaluativa concreta, relatar la metodología de evaluación (en la que se incluyan, además de las especificaciones técnicas de los instrumentos y técnicas utilizadas para la recopilación y el análisis de la información, una explicación de la estrategia evaluativa general, el modo en el que se han establecido los criterios de evaluación, los propios criterios de evaluación, las estructuras y foros que se han utilizado para la recopilación y cómo se han establecido las valoraciones y recomendaciones).

Otros aspectos a tener en cuenta en este apartado son la **difusión** que se ha hecho de los informes de evaluación, quiénes han sido los **destinatarios** de dicha difusión y su **calendario**. Además de la calidad del producto final del proceso de evaluación -el informe- hay que valorar la difusión que se le da, como otra condición para su uso y utilidad. Es necesario saber si se han difundido dichos informes, a través de qué instrumentos (reuniones, jornadas, publicación) y quiénes han sido sus destinatarios principales. En algunos casos, el simple hecho del calendario de la evaluación puede dificultar enormemente la utilización de la evaluación, al menos su uso instrumental para la mejora del plan siguiente.

El impacto de los resultados y los procesos de evaluación en la política y en los organismos de igualdad

El último criterio a tener en cuenta para valorar las experiencias de evaluación analizadas es el impacto que han tenido dichas experiencias, tanto en los propios planes y políticas de igualdad de género, como en los organismos de igualdad impulsores de dichas políticas. Para valorar su impacto debemos tener en cuenta las consecuencias que las experiencias de evaluación tienen para las políticas y para los organismos de igualdad que las propician. Sin embargo, compartimos la idea de Scriven de creer que “las evaluaciones no deben ser evaluadas en términos de sus consecuencias reales (excepto por un historiador), sino sólo en términos de sus consecuencias previsibles si se usan apropiadamente” (Scriven, 1988: 230). Entre otras cosas, para valorar las consecuencias reales de una evaluación habría que dejar pasar el suficiente tiempo para consignar las posibles consecuencias de dicha evaluación también a medio y largo plazo. Por lo tanto, el concepto de impacto de las evaluaciones está inextricablemente unido al de uso; el impacto de las evaluaciones depende del uso que se haga de ellas.

Adoptando una visión amplia de la utilización de la evaluación, recordamos los cuatro tipos de uso de los que habla Carol Weiss (Weiss, 1998b: 23-24): uso instrumental (utilizando la evaluación para tomar decisiones), conceptual (utilizando la evaluación para cambiar la comprensión de lo que el programa o plan es, hace o pretende), persuasivo (sirviendo para movilizar apoyos para conseguir cambios necesitados en el programa), y por último, un uso que influye en otras instituciones,

³⁰ La norma U5 (Claridad del informe) del *Joint Committee* dice que “Los informes de evaluación deben describir claramente el programa que está siendo evaluado, incluyendo su contexto, y los propósitos, procedimientos y resultados de la evaluación, de manera que se ofrezca información esencial y fácilmente comprensible” (*Joint Committee* 94: 49).

más allá del programa o política evaluada (influyendo en las redes de política o *policy networks* de programas similares, en los profesionales relacionados con el problema que se aborde e incluso en los paradigmas de políticas y las agendas políticas). Así, aplicando estos cuatro usos al caso de las políticas de igualdad de género, deberíamos buscar el impacto de la evaluación no sólo en su uso instrumental, para la propia toma de decisiones sobre los planes evaluados, sino también en el uso conceptual, analizando si las evaluaciones realizadas han tenido algún tipo de impacto en lo que se refiere a una mayor comprensión sobre los planes y las políticas de igualdad. Además, se debería ver si la evaluación ha servido para movilizar apoyos para el plan (uso persuasivo) y si ésta influye en otras instituciones, en las redes de política (*policy networks*) o en la propia agenda de gobierno.

Pero, además de en los usos de los **resultados** de la evaluación, aquí se pretende analizar especialmente los usos de los **procesos de evaluación**, sirviendo éstos para introducir una cultura y una metodología de evaluación y para generar una mayor comprensión de la intervención pública en general, además de una mayor consciencia sobre cómo se abordan los problemas públicos³¹.

En este sentido, seguimos a Torres, Preskill y Piontek (1996: 2) cuando definen el **uso** de la evaluación como cualquier consecuencia o resultados de una evaluación y que incluye, al menos:

1) **Los efectos de los mismos procesos de evaluación**, incluida una mayor comprensión y apreciación de la evaluación. También podrían incluirse, entre estos efectos, el fortalecimiento o institucionalización de la función de evaluación, tanto en los propios planes, como en las instituciones que los impulsan -los organismos de igualdad-; asimismo, el aprendizaje sobre dicha función de evaluación, es decir, que por efecto de la evaluación, se hayan aprendido a evaluar mejor los planes de igualdad.

2) **Una mayor comprensión o nuevas perspectivas en una situación, programa u organización**. Es evidente que la reflexión sobre los planes que, en algunos casos, proporcionan las evaluaciones ayudan a tener una visión más global, consciente y profunda de lo que se está haciendo y cuáles son las principales dificultades y fortalezas de las políticas de igualdad.

3) **Las acciones que se llevan a cabo como resultado de la evaluación**. Además de buscar información sobre si se han modificado contenidos o estrategias, tanto globalmente (por ejemplo, dar más peso a un tipo de acciones que a otras) como en el nivel concreto de cada actuación (modificando cualquier aspecto de cada actuación o actividad a realizar), debido a los resultados de la evaluación, se deben buscar, además, otro tipo de acciones. Por ejemplo, aquellas que tienen más que ver con la propia estructuración del plan (concretar más las acciones, asignarles una delimitación temporal y unos responsables concretos para su

³¹ Nuestra insistencia tiene que ver con que, en general, la “utilidad” de los procesos de evaluación analizados en nuestra metaevaluación (aunque en muchos de los casos tampoco ésta ha sido muy alta) se encuentra más en este segundo aspecto que en el uso instrumental y directo de los resultados concretos de los estudios de evaluación que, en realidad, ha sido muy bajo en la mayor parte de los casos, como se verá en el análisis de cada una de las experiencias.

ejecución) o con su posterior evaluación (adquirir compromisos, aunque éstos siendo muy generales, respecto a la misma en los subsiguientes planes),

En definitiva, se trata de consignar algún indicio o “rastros” de utilización en estos tres sentidos para poder dar algo de luz sobre este aspecto, aunque se debe reconocer que el estudio del impacto de las experiencias de evaluación requeriría de una investigación específica que va más allá de los objetivos y posibilidades de este trabajo que aquí se presenta.

Una vez expuestos los criterios para el análisis y valoración de las once experiencias de evaluación, recordamos aquí que las técnicas que se utilizaron para la recopilación y el análisis de la información son la recopilación y análisis documental (esencialmente los propios planes, informes de evaluación, así como otra documentación relacionada -proyectos, pliegos de condiciones técnicas, fichas y registros, etc.-, y las entrevistas abiertas semiestructuradas realizadas a las responsables de los organismos de igualdad. Mucha de la información fue completada y contrastada a través de las entrevistas estructuradas personales y telefónicas realizadas a los y las responsables técnicos/as de la evaluación en los correspondientes organismos de igualdad.

La exposición de la información de las once experiencias de evaluación se hace a lo largo de los próximos tres capítulos (cinco, seis y siete), utilizando las tres grandes dimensiones a través de las que se han organizado los criterios de metaevaluación (planteamiento de las evaluaciones, elementos clave y utilización de los procesos de evaluación).

CAPITULO 5

EL PLANTEAMIENTO GENERAL DE LOS PROCESOS DE EVALUACIÓN: TIPOS, CONTEXTO, PROPÓSITOS, ESTRUCTURAS Y RECURSOS

INTRODUCCIÓN

A lo largo de éste y de los dos siguientes capítulos se desarrolla una descripción pormenorizada de los procesos de evaluación llevados a cabo para valorar la segunda generación de planes de igualdad autonómicos y estatales en España. Este capítulo cinco se centra en el planteamiento general de dichas evaluaciones, destacando cinco dimensiones de dicho planteamiento: los tipos de evaluación realizados, el contexto que rodea la situación y el proceso de evaluación, los propósitos y objetivos de dichos procesos, las estructuras institucionales existentes y utilizadas para la evaluación, así como los recursos invertidos en la misma.

Con el fin de ordenar la exposición de la información de las once experiencias metaevaluadas se utilizó, además del esquema propuesto de criterios de evaluación -expuesto en el capítulo anterior-, la clasificación o tipología según el agente evaluador. La elección de esta tipología es debida a que creemos que es la que marca mayores diferencias entre las experiencias de evaluación estudiadas. De esta manera, se clasificaron las once experiencias en evaluaciones externas, evaluaciones mixtas y evaluaciones internas.

Para fundamentar la elección de este criterio de clasificación, el primer punto que se desarrolla en el presente capítulo es la descripción de los diferentes **tipos de evaluación** utilizados en las once experiencias metaevaluadas. Se describen los tipos de evaluación utilizados según las cuatro tipologías básicas de evaluación señaladas en el capítulo dos: según el papel que cumple la evaluación -evaluación formativa o sumativa-, según el contenido o fase del programa o política -diseño, proceso y/o resultados-, según el momento en el que se evalúa -evaluación *ex-ante* o *ex-postfacto*-, y según el agente evaluador -evaluaciones internas o externas-.

Una vez definido y fundamentado el criterio de clasificación escogido, se plantea la información en las restantes dimensiones relacionadas con el planteamiento de las evaluaciones: contexto, propósitos, estructuras institucionales y, finalmente, recursos invertidos en las mismas. Para cada una de estas dimensiones, además de ofrecer un primer panorama general, se describen más pormenorizadamente cada una de las experiencias metaevaluadas, utilizando siempre el criterio de clasificación según el agente evaluador -evaluaciones externas, evaluaciones mixtas y evaluaciones internas-.

El análisis de los **contextos** en los que se generan los procesos de evaluación puede ofrecer claves sobre el tipo de respuesta evaluativa seleccionada en cada administración pública. Dentro de la dimensión del contexto evaluativo se contemplaron la antigüedad y la fortaleza del organismo de igualdad responsable, así como su adscripción a determinadas áreas gubernamentales, el partido político gobernante en el momento en el que se encargó, realizó y entregó la evaluación, la estabilidad y continuidad de las responsables de los organismos de igualdad, además de la existencia de experiencias previas en evaluación.

En segundo lugar, se analizan los **propósitos** de las evaluaciones y su claridad explícita o ausencia de la misma. Así, el estudio de las razones de la evaluación (tanto explícitas como implícitas) ayudará a entender el papel que la función de evaluación cumple en cada una de las instituciones que propician y ejecutan la evaluación y, en general, en las políticas públicas de igualdad de género. Las dos cuestiones que se tuvieron en cuenta en el análisis de estos propósitos u objetivos de los estudios de evaluación fueron su grado de claridad y explicitación, así como el grado de ambición en su planteamiento.

En tercer lugar, se describen las **estructuras institucionales** que permitieron, fomentaron y apoyaron la evaluación, es decir, cómo la evaluación se organizó institucionalmente. Una primera dimensión que se analiza al respecto es si existía o no previamente la mención explícita a su evaluación en el Plan aprobado; así vemos que, en esta segunda generación de evaluaciones, las menciones explícitas eran todavía pocas. También se analizan los tipos de estructuras que, en torno a los organismos específicos de igualdad, tenían alguna responsabilidad en el seguimiento y evaluación de los planes, bien sean los propios Consejos Rectores de los Institutos, estructuras “ad-hoc” para el seguimiento y evaluación de los planes u otras estructuras no específicas. También se señala, en este apartado, si los Consejos Regionales u otras estructuras de participación social tuvieron alguna implicación en los procesos de evaluación, así como si existía alguna vinculación u obligación formal con el poder legislativo, relacionada con la evaluación de los planes o los organismos de igualdad.

Por último, se expone un panorama general de los **recursos** económicos, técnicos y humanos empleados en las evaluaciones objeto de estudio; esto último puede resultar un indicador especialmente interesante del papel que la evaluación cumple en el proceso general de las políticas públicas de igualdad de género. Aunque resultó prácticamente imposible obtener datos fiables de las dimensiones que se habíamos establecido previamente (costes indirectos de la evaluación, presupuestos de los planes y de los organismos de igualdad), se realizaron estimaciones y descripciones generales de cada una de las experiencias metaevaluadas.

1. TIPOS DE EVALUACION REALIZADOS

Uno de los aspectos más importantes del planteamiento de una evaluación es el tipo de evaluación que se escoge para responder ante una determinada situación evaluativa. La situación evaluativa, en sus aspectos del contexto que rodea la situación, las estructuras y los recursos existentes para la evaluación, así como los propósitos que se plantean con dicha evaluación, se analiza en el punto siguiente de este capítulo.

La idea de comenzar analizando los tipos de evaluación realizados responde a una necesidad de clasificación de las once experiencias a metaevaluar. Esta clasificación, además de ser uno de los puntos clave de análisis, nos permite un segundo criterio para la exposición de la información de las once experiencias. Es decir, además de exponer la información según las tres grandes dimensiones previamente expuestas - planteamiento (capítulo cinco), elementos clave (capítulo seis) y utilización (capítulo siete)-, se utiliza también un criterio de clasificación -evaluaciones externas o internas- según el agente que realiza la evaluación. A continuación se fundamenta tal elección.

1.1. Un primer análisis de las once experiencias según los tipos de evaluación realizados

Uno de los criterios para la metaevaluación, mencionados en el punto anterior (punto A.c.), es el tipo de evaluación realizado. Este criterio se utiliza en este apartado y en este capítulo para establecer una primera diferenciación que nos permita clasificar las experiencias de evaluación, con el fin de exponer los resultados de una forma coherente y ordenada.

Como ya se ha comentado, aquí se van a utilizar cuatro clasificaciones básicas: según el papel que se quiere que cumpla la evaluación (evaluación formativa vs. sumativa), según el contenido o fase de la política que se evalúa (evaluación de diseño, proceso y/o resultados), según el momento en el que se evalúa (evaluación *ex-ante* o *post-facto*) y según el agente evaluador (evaluación interna o externa). En la tabla 5.0 se ofrece una visión general, aunque resulta conveniente realizar algunas matizaciones.

**TABLA 5.0: TIPOS DE EVALUACION
SEGUNDA GENERACION DE EVALUACIONES**

	Según el papel que cumple la evaluación	Según el contenido o fase	Según el momento	Según el agente evaluador
Instituto Mujer II (93-95)	Sumativa	Proceso y Resultados	<i>Post-facto</i>	Mixta Diseño interno, parte ejecución externa
Andalucía II (95-97)	Sumativa	Diseño, Proceso y Resultados	<i>Post-facto</i>	Externa
Aragón I (94-96)	Formativa/ Sumativa	Proceso y Resultados	<i>Post-facto</i>	Interna
Canarias I (95-96)	Sumativa	Proceso y Resultados	<i>Post-facto</i>	Externa
Castilla-La Mancha II (95-99)	Sumativa	Diseño, Proceso y Resultados	<i>Post-facto</i>	Externa
Castilla y León I (94-96)	Sumativa	Resultados	<i>Post-facto</i>	Externa
Cataluña II (94-96)	Sumativa	Diseño, Proceso y Resultados	<i>Post-facto</i>	Mixta Diseño con asesoría externa, ejecución interna
Galicia II (95-97)	Sumativa	Diseño, Proceso y Resultados	<i>Post-facto</i>	Externa
Madrid II (93-95)	Sumativa	Proceso y Resultados	<i>Post-facto</i>	Externa
La Rioja II (96-99)	Formativa/ Sumativa	Proceso y Resultados	<i>Post-facto</i>	Interna
País Vasco II (96-2000)	Formativa/ Sumativa	Diseño, Proceso y Resultados	<i>Post-facto</i>	Mixta Diseño con asesoría externa, ejecución interna

En primer lugar, y referente a la **clasificación según el papel que se quiere que cumpla la evaluación**, todas las evaluaciones realizadas tienen carácter sumativo, salvo la evaluación del I Plan aragonés y las del II PAPME vasco y II Plan riojano, cuyo carácter, en cierto modo, podría ser considerado como formativo. En el caso aragonés, el carácter formativo lo da la intención de realizar la evaluación para, sobre la misma marcha, establecer las prioridades del Plan para el año 1996. En el caso vasco (y de forma similar a lo que ocurre en La Rioja), el carácter formativo viene dado por el hecho de que la Comisión Interdepartamental debe realizar un primer informe de seguimiento al finalizar el primer trimestre de cada año, junto con una valoración en torno a la consecución de los objetivos previstos en el Plan. En buena lógica, dicha valoración se hace para rectificar lo que todavía sea posible para dicho año. También las memorias que cada año *Emakunde* presenta al Parlamento tienen un cierto carácter formativo respecto al Plan, ya que permite modificar cosas año a año, aunque creemos que en realidad su origen es más sumativo ya que hace hincapié en la función de rendimiento de cuentas y en hacer recapitulación de cada año.

Efectivamente, el hecho de recopilar información desde un principio y con carácter continuado no tiene por qué implicar necesariamente una evaluación formativa. En La Rioja, en la que también se plantean memorias anuales y en la que la evaluación tiene un carácter continuo, y desde el inicio del Plan, ésta tiene un carácter también más sumativo que formativo. Respecto a la evaluación de objetivos y acciones, se dice en el (II) Plan Integral de las Mujeres de La Rioja que “el objetivo de este tipo de evaluación es el análisis del grado de cumplimiento del Plan, es decir, en qué medida se han llevado a cabo o no las acciones propuestas”. Y sigue, “se efectuará un seguimiento del Plan mediante la recogida de información continuada, a modo de memoria, de las medidas o acciones realizadas” (subrayado mío). El tiempo pasado del verbo y la idea de “ir guardando” los datos recogidos a modo de memoria para su posterior evaluación, nos indica una intención más sumativa que formativa. Con la evaluación de la participación institucional se señala que se irá recogiendo la información, “analizando la evolución de su grado de implicación en el Plan durante su período de vigencia”. El análisis de la evolución requiere haber recogido información durante todo el período de vigencia para tener datos suficientes para valorar dicha evolución. La evaluación tendría carácter formativo si se insistiera en la intención de ir recogiendo información para ir evaluando y, por lo tanto, ajustando el Plan sobre la marcha, durante el mismo proceso de ejecución. Se trataría de utilizar la evaluación como un instrumento de conformación y ajuste del propio Plan, no como recapitulación para obtener información útil para el rendimiento de cuentas y para mejorar los planes futuros.

Respecto a la **clasificación según el contenido o la fase de la política que se evalúa**, es conveniente hacer una serie de puntualizaciones para cada categoría y cada proceso evaluador realizado.

En general, si se observa la tabla 5.0, se ve que se ha hecho muy poca **evaluación de diseño**. Hacer este tipo de evaluación implica valorar el Plan en su conjunto, incluyendo una valoración de su fundamentación, objetivos, coherencia interna, nivel de concreción, estrategias y metodologías empleadas para su formulación e implantación,

etc. **Sólo en dos comunidades autónomas se plantean este tipo de evaluación de forma explícita: Andalucía y Castilla-La Mancha.** De estas dos evaluaciones la más completa, es decir la que se plantea más parámetros del diseño, es la de Andalucía, ya que además de plantearse la coherencia y concreción de cada uno de los objetivos (como se hace en Castilla-La Mancha), se analiza y valora el diseño del Plan en su conjunto, incluyendo el concepto de igualdad al que se está respondiendo. Aunque no sea de forma tan explícita ni completa, tanto Cataluña como Galicia y el País Vasco hacen también parte de este tipo de evaluación, ya que analizan el Plan en su globalidad y recogen información y valoran los tipos de actividad finalmente realizados y ejecutados. Esto arroja una información que permite la reflexión y valoración sobre a qué tipo de actividades se les está dando mayor peso, incluso a qué personas destinatarias llega mejor o peor la actuación de los poderes públicos. Y parte de esta reflexión puede ser sobre el propio diseño de los planes.

Respecto a la *evaluación de proceso* vemos que todas las evaluaciones, con la excepción de Castilla y León, han incluido algún tipo de evaluación de proceso y esto es debido, sobre todo, al análisis del “grado de cumplimiento de las acciones” que se realiza en casi todas las evaluaciones. En algunos casos (como en la evaluación del II plan nacional -a partir de ahora denominada del Instituto de la Mujer -”IM”-, y como la de Andalucía), es denominada “evaluación de la implementación”, que viene a ser la evaluación de la ejecución del Plan. Efectivamente, la evaluación de proceso es la que valora el proceso de puesta en práctica del plan, incluido el proceso de progresión en la ejecución de las acciones, el ajuste al calendario, si se llega a las personas destinatarias previstas, las dificultades que se presentan en dicho proceso (problemas de información, tiempo o desplazamientos, Ballart, 1992: 110-111) y todo lo referente al funcionamiento de dicha puesta en práctica incluyendo, sobre todo, los procesos de coordinación con otros departamentos y entidades. Todo lo referente al funcionamiento de los diferentes foros y comisiones tiene especial interés en una evaluación de proceso de los planes de igualdad. En la mayoría de las evaluaciones analizadas, el grado de cumplimiento de las acciones se refiere a la comprobación de que las acciones propuestas se han ejecutado. Si esto se queda sólo en eso, sería más bien una muy simple evaluación de producto (resultados); pero si también se recoge qué acciones se han ido ejecutando en qué momentos del período de vigencia del Plan, entonces eso permite una cierta evaluación de proceso. Lo mismo ocurre con la participación de las diferentes instituciones a lo largo del proceso o con los colectivos destinatarios: estudiando el grado de cobertura, el sesgo en la población servida y la tasa de abandonos (Ballart, 1992: 108-109). En algún caso, como en la evaluación de la implementación del IM, además del grado de cumplimiento, recogen información sobre cómo se ha ejecutado el Plan refiriéndose exclusivamente a una descripción detallada de las diferentes actividades puestas en marcha en su ejecución. Las evaluaciones de proceso más completas, como se verá en más detalle cuando analicemos los criterios de evaluación utilizados, son las realizadas en Andalucía, Castilla-La Mancha, La Rioja y País Vasco; también, en cierto modo, Galicia, ya que ofrece explícitamente datos de la evolución de las actividades realizadas por año. Aún así, han existido algunas limitaciones para estas evaluaciones de proceso; en general, hay muchas más probabilidades de obtener buena información de proceso si se recoge la información de forma continuada y desde un principio y esto, como se verá en la próxima clasificación, es algo en lo que han fallado la mayoría de las evaluaciones analizadas.

La *evaluación de resultados* puede dividirse en la evaluación de *outputs* o productos y de *outcomes* o efectos. Parece especialmente conveniente hacer esta distinción para estudiar la evaluación de resultados que se ha realizado en las evaluaciones objeto de nuestro análisis.

Una cierta evaluación de *outputs* o productos se realiza en todos los procesos evaluativos estudiados, aunque en el caso de Castilla y León esta evaluación es prácticamente simbólica respecto a otros datos de resultados que se ofrecen. En cierto modo, y como se ha mencionado anteriormente, el análisis del “grado de cumplimiento” del Plan, podría ser considerado más bien como una evaluación de *outputs*, de lo que se ha hecho o generado por la puesta en marcha del Plan, es decir, de lo que ha producido. Saber que se han hecho tales actividades, de las que finalmente se ha beneficiado tal gente, que se han hecho tales materiales o campañas, etc. es, en realidad, una evaluación de los “productos” del Plan. En este sentido, todas las evaluaciones han contemplado los resultados de los planes. Otra cosa es que en algunos casos se limite simplemente a constatar que dichas actividades han ocurrido o se han ejecutado. Insistimos en que, en realidad, muchas de las evaluaciones denominadas de proceso o de implementación son más evaluaciones de productos finales (qué se ha hecho finalmente, quiénes han colaborado o participado) que del proceso de puesta en práctica y funcionamiento del Plan en general y de las medidas en particular (cuál ha sido el proceso por el cual se han transformado los inputs en los resultados previstos). La única experiencia de evaluación en la que se “reconoce” este tipo de evaluación como evaluación de productos (resultados) es la de Castilla-La Mancha, en la que se habla de “memoria o mapa de resultados”.

La evaluación de la calidad de dichos productos es mucho menos común, entre otras cosas, porque como se verá más adelante, no están definidos los criterios de calidad de las actuaciones. Por ejemplo, en un curso de formación, además de saber cuántas mujeres han pasado y de cuántas horas ha constado, habría que aplicar criterios de calidad del producto formativo en sí mismo: por ejemplo diseño coherente y adecuado a las destinatarias, objetivos realistas, buena organización, buen profesorado y metodología didáctica (y de cada uno de éstos habría que definir que es una buena organización y un buen profesor y metodología), material suficiente y de calidad, grado de satisfacción del alumnado, etc. El proyecto de “evaluación de la implementación” de Andalucía contemplaba la evaluación de calidad, definiéndola como el grado en el que se han cumplido los atributos que definen la calidad de la actividad en relación con unos estándares de calidad establecidos. También las evaluaciones de Cataluña, Galicia y País Vasco son algo más completas respecto a la evaluación de *outputs*.

Por último, y como parte de la evaluación de resultados, estaría la evaluación de *outcomes*, es decir de los efectos o impacto¹ del plan. En todas las evaluaciones se entiende como efectos sobre la población destinataria, aunque ésta es, por lo general, muy variada y amplia (a veces y para determinadas actuaciones de los Planes la población destinataria es la población en general, como por ejemplo, de una campaña de

¹ Algunos autores, como Alvira (1991) distinguen entre evaluación de resultados, que serían los efectos sobre la población destinataria, y evaluación de impacto, que serían los efectos de un programa sobre la sociedad en general. Aquí, se prefiere hablar de efectos (o impactos, indistintamente) directos (población destinataria) e indirectos (sobre la población en general). Si el programa va dirigido a la población en general, el impacto sobre la misma serían efectos directos.

sensibilización sobre el reparto de las responsabilidades domésticas). La mayoría de las evaluaciones analizadas no pretenden la evaluación de impacto o efecto (Aragón, Canarias, Castilla-La Mancha, Cataluña, Galicia y País Vasco) y otras, aunque lo pretenden, no lo consiguen. Es el caso del IM, Castilla y León y Madrid, en las que se realizan excelentes estudios de la situación social de la mujer, pero que no son evaluaciones de los efectos del Plan. Una de las grandes dificultades de la evaluación de programas es, precisamente, la evaluación del llamado “efecto neto” del programa, que son los efectos atribuibles al programa y no a otros factores. Es decir, el efecto neto de un Plan consiste en los efectos producidos por el mismo, es decir, lo que no hubiera ocurrido si el Plan no se hubiera ejecutado. La evaluación de efectos o impactos de un Plan debe, por lo tanto, procurar la identificación de dichos “efectos netos”. En los tres casos que se han señalado, se han medido diferentes variables relacionadas con la situación diferencial de las mujeres, respecto a los hombres (y que, lógicamente, tienen que ver con lo que pretenden los planes, como por ejemplo, si ha aumentado el número de mujeres universitarias o si han variado los usos del tiempo), y se han medido en el momento en el que comenzó el Plan y en el momento en el que finaliza. Se han comparado esas dos medidas y se ha identificado su evolución y si ha habido algún avance al respecto. Pero los estudios no dicen nada sobre si los Planes han tenido alguna influencia sobre dicho cambio, ya que los cambios pueden deberse a muchos otros factores, entre ellos la propia evolución de la sociedad. Es cierto que este tipo de estudios puede ser una muy buena base para las evaluaciones de efectos, pero no son evaluaciones de los Planes de Igualdad, sino de la situación de la mujer y su evolución en un determinado período de tiempo. La única experiencia analizada que pretende una evaluación de los efectos y que establece un diseño de evaluación que la permite es Andalucía, que hace cuatro diferentes encuestas, combinando diferentes criterios (preguntas) momentos y destinatarias (población general o usuarias/beneficiarias), además de utilizar también el estudio que sobre la situación social de la mujer andaluza se realizó en el año 1988.

Otra de las formas de hacer evaluación de efectos o impactos que se plantea en las evaluaciones analizadas es la evaluación de la percepción y opinión de los colectivos destinatarios o de las personas implicadas en la ejecución del Plan (normalmente, trabajadores o responsables de otros departamentos gubernamentales). Por lo general, la fundamentación de este tipo de evaluación es recoger la opinión o percepción de agentes clave sobre el Plan, sin especificar si se trata de opiniones o percepciones sobre el diseño, el proceso de ejecución o sus resultados. En el caso del IM, la “evaluación cualitativa” se concibe como parte de la evaluación de resultados, aunque los criterios sobre los que se recoge información no tienen nada que ver con los de la “evaluación cuantitativa”, que también forma parte de dicha evaluación de resultados. Con dicha evaluación cualitativa lo que se pretende es conocer la “percepción personal que del desarrollo y los efectos de las actuaciones han tenido los responsables de la coordinación desde los distintos ministerios, y la percepción social que, de los efectos de las actuaciones, han tenido sus destinatarias” (Alvaro et al. 1998: 38). Es decir, además del desarrollo, se pretende obtener la percepción sobre los efectos. En otros casos, como en Castilla-La Mancha, se plantea inicialmente recoger la percepción de los colectivos destinatarios y personas implicadas respecto al desarrollo del Plan. En la exposición de los resultados de la evaluación, sin embargo, se refleja cómo en el análisis se discierne bien entre las opiniones y percepciones sobre el diseño del plan, el desarrollo del mismo (cuestiones de organización, implicación de entidades y dotación

de recursos), los resultados y el impacto y nivel de conocimiento del Plan en la sociedad. En La Rioja, sólo se especifica que “al finalizar el Plan, se recogerán opiniones de la población destinataria respecto a las medidas establecidas por el Plan” y que se hará mediante grupos de discusión y entrevistas a la población destinataria y a agentes implicados en el Plan².

En tercer lugar, se analiza la **clasificación según el momento en el que se evalúa**. Desde este punto de vista, todas las evaluaciones son *post-facto* ya que todas se plantean una vez comienza (o ya ha comenzado) el Plan. Es decir, ninguna de las evaluaciones se plantea *ex-ante*, proponiendo, por ejemplo una evaluación del diseño del Plan, previa a su puesta en marcha, o un estudio piloto de una o varias medidas, antes de generalizarlas. Es más, y como ya se ha comentado anteriormente, uno de los grandes problemas de la evaluación, no es ya que se hagan *post-facto* sino que se comiencen a diseñar y ejecutar cuando ya el plan está muy avanzado. Esta circunstancia hace que el instrumento de evaluación pierda muchísima potencia, en primer lugar, por no estar diseñado paralelamente al plan y segundo, porque se pierde la oportunidad de recoger información muy valiosa al inicio y durante el proceso de ejecución de los planes. Las únicas evaluaciones que se plantean desde un principio son la evaluación del II PAPME vasco y la del (II) Plan Integral de La Rioja. El resto se plantean, por lo general, en un período no superior a un año de vencimiento del período de vigencia del correspondiente plan.

Finalmente, analizamos las evaluaciones con la **clasificación según el agente evaluador**. Según esta clasificación las evaluaciones pueden ser internas, o externas, si se realizan por parte de alguien que no tenga que ver ni con la financiación, ni con la formulación ni con la ejecución del plan evaluado. También cabe la posibilidad de procesos de evaluación mixtos, que recurran a un equipo externo sólo para una parte del proceso de evaluación; en realidad, todos estos procesos mixtos son, en realidad, evaluaciones internas en las que, aunque se recurre a una colaboración externa, el proceso global de evaluación es responsabilidad interna. Este es el caso del IM, Cataluña y País Vasco, aunque las características de estos procesos mixtos son distintos. En el caso del IM, el diseño de la evaluación es interno y sólo recurren a una empresa externa para la recopilación y análisis de la información de la evaluación de resultados (la evaluación de la implementación es totalmente interna). Incluso, y como se refleja en el pliego de condiciones técnicas, incluso el diseño de la “evaluación cuantitativa” y de la “evaluación cualitativa” (sobre todo el primero), incluyendo los criterios y las variables a estudiar, estaba ya realizado internamente y en realidad, lo que hace la empresa externa es ejecutarlo. En el caso de Cataluña y el País Vasco es al contrario, ya que para lo que los respectivos Institutos recurren a asesoramiento externo es para el diseño de la evaluación, haciéndose la ejecución internamente. En el caso de Cataluña, la dimensión e importancia del asesoramiento externo es mayor que en el caso del País Vasco.

² (II) Plan Integral de la Mujer de La Rioja, p.66

Las evaluaciones del I Plan aragonés y del II riojano son procesos netamente internos. El resto de las evaluaciones (Andalucía, Canarias, Castilla-La Mancha, Castilla y León, Galicia y Madrid) se realizan por parte de equipos externos en coordinación con los respectivos organismos de igualdad. Como ya se ha expuesto, en la tabla 5.0 se exponen la totalidad de los once procesos evaluativos estudiados de esta segunda generación de evaluaciones de planes de igualdad, según las cuatro clasificaciones básicas explicadas.

1.2. Los criterios para la clasificación

Una de las clasificaciones más claras y factibles de las experiencias de evaluación a analizar, es la de los tipos de evaluación realizados. Las cuatro tipologías básicas expuestas en el punto anterior no son mutuamente excluyentes. Es decir, cualquier experiencia de evaluación puede ser clasificada por todas y cada una de estas cuatro tipologías. Este hecho hace difícil el establecimiento de una única clasificación que contemple las cuatro tipologías. Esta es la razón por la que se propone la tipología según el agente evaluador para clasificar las once experiencias de evaluación de la segunda generación.

Las razones para escoger la tipología según el agente evaluador, son las siguientes:

1) Esta tipología es, junto con la tipología según el contenido, la más clara, es decir, la que más discrimina y distingue entre unas experiencias y otras, ya que las otras dos tipologías-según el momento y según el papel- prácticamente no discriminan: casi todas las experiencias son de carácter sumativo y *post facto*.

2) La tipología de contenido, que sería la otra clasificación más adecuada y factible, por un lado, contiene categorías que no son mutuamente excluyentes. Es decir, una misma experiencia de evaluación puede ser de diseño, proceso y resultados, lo que hace que no sea una tipología tan adecuada a efectos de clasificación. Por otro lado, lo que en algunas experiencias se autodenomina como evaluación de proceso, por ejemplo, nosotros lo clasificaríamos con otro tipo de evaluación. Por lo tanto, la tipología según el agente sigue siendo la más clara.

3) Finalmente, la tipología según el agente evaluador es la que refleja las opciones más conscientes, y desde un inicio, por parte de los organismos de igualdad responsables de las evaluaciones, sobre cómo se quieren llevar a cabo los procesos de evaluación. El resto de las opciones contempladas respecto a las restantes tipologías no parecen haber sido tomadas de forma tan consciente.

Por lo tanto, la clasificación queda hecha de la siguiente manera:

- Evaluaciones externas

- Evaluaciones mixtas
- Evaluaciones internas

EVALUACIONES EXTERNAS

Son las evaluaciones en las que se encarga a un equipo externo -es decir, que no tiene que ver ni con la financiación, ni con la planificación ni con la ejecución del plan evaluado- el estudio de evaluación en su totalidad: diseño, ejecución e informes.

Estas son:

Evaluación del **II Plan Andaluz (1995-1997)**

Evaluación del **I Plan Canario (1995-1996)**

Evaluación del **II Plan Castellano-Manchego (1995-1999)**

Evaluación del **I Plan Castellano Leonés (1994-1996)**

Evaluación del **II Plan Gallego (1995-1997)**

Evaluación del **II Plan Madrileño (1993-1995)**

Aunque estas seis experiencias comparten la característica de ser externas, son muy heterogéneas entre sí, como se podrá observar a lo largo de la exposición concreta de los datos. Además de las diferencias temporales, tanto del momento en el que se plantean (van desde la evaluación del I Plan canario que se realiza a finales del año 1996 a la del II castellano-manchego que se hace en la primera mitad del año 1999), como a la duración de los procesos (de los tres meses de la evaluación canaria a los dos años y medio de la andaluza), hay evidentes diferencias en las propias dimensiones de los estudios, costando algunos menos de dos millones de pesetas (estudio canario y castellano leonés) y, otros, como el andaluz, más de cincuenta.

Sin embargo, todos estos estudios de evaluación son definidos como externos por parte de los correspondientes organismos de igualdad, y se hace hincapié, precisamente, en este carácter externo y, por lo tanto, supuestamente más objetivo. Aunque, en la mayoría de los casos, la decisión de encargar una evaluación externa tiene más que ver con la incapacidad técnica de llevarla a cabo internamente (no sólo por falta de formación o especialización profesional sino, sobre todo, por falta real de recursos humanos que pudieran encargarse y dedicarle el tiempo necesario a ello), que por necesidad de una mayor objetividad.

En la tabla 5.0.a se exponen de nuevo las seis evaluaciones externas según los otros tres tipos de evaluación:

TABLA 5.0.a: TIPOS DE EVALUACION
Evaluaciones externas
SEGUNDA GENERACION DE EVALUACIONES

	Según el papel que cumple la evaluación	Según el contenido o fase	Según el momento	Según el agente evaluador
Andalucía II (95-97)	Sumativa	Diseño, Proceso y Resultados	<i>Post-facto</i>	Externa
Canarias I (95-96)	Sumativa	Proceso y Resultados	<i>Post-facto</i>	Externa
Castilla-La Mancha II (95-99)	Sumativa	Diseño, Proceso y Resultados	<i>Post-facto</i>	Externa
Castilla y León I (94-96)	Sumativa	Resultados	<i>Post-facto</i>	Externa
Galicia II (95-97)	Sumativa	Diseño, Proceso y Resultados	<i>Post-facto</i>	Externa
Madrid II (93-95)	Sumativa	Proceso y Resultados	<i>Post-facto</i>	Externa

EVALUACIONES MIXTAS

Son evaluaciones en las que se recurre a equipos externos en sólo una parte del proceso de evaluación. Puede ser que el diseño sea interno y parte de la ejecución externa (como es el caso del IM) o, por el contrario se pida asesoramiento externo en el diseño de la evaluación, aunque la ejecución se haga internamente (País Vasco y Cataluña).

Estas son:

Evaluación del **II Plan del IM -nacional- (1993-1995)**

Evaluación del **II Plan Catalán (1994-1996)**

Evaluación del **II PAPME -País Vasco- (1996-2000)**

Estas tres experiencias de evaluación comparten la característica de ser procesos mixtos en lo que a los agentes de la evaluación, es decir, que en algún momento del proceso de evaluación recurren a equipos externos. Sin embargo, aquí se sostiene la tesis de que, en realidad, son procesos internos de evaluación en los que requiere de colaboraciones externas en determinados momentos del proceso. En este sentido, estas tres experiencias también son diferentes entre sí ya que mientras el diseño del estudio de evaluación del II Plan central (IM) es interno y sólo se “externaliza” para el trabajo de campo y el análisis de determinada parte del estudio, en Cataluña y el País Vasco, sin embargo, recurren a los equipos externos para un asesoramiento desde el propio diseño del estudio aunque una buena parte de la ejecución de dicho estudio sea interna. También las experiencias de Cataluña y el País Vasco difieren entre sí: mientras en Cataluña, la colaboración del equipo externo -ESADE- resulta crucial para la definición del proceso de evaluación, en el País Vasco la colaboración del equipo externo - Universidad del País Vasco- se subsume más en un proceso interno ya iniciado con la evaluación del I PAPME y en una organización -*Emakunde*- con una función y una cultura de evaluación más o menos desarrollada.

En la tabla 5.0.b. se exponen estas tres experiencias según los otros tres tipos de evaluación.

TABLA 5.0. b.: TIPOS DE EVALUACION
Evaluaciones mixtas
SEGUNDA GENERACION DE EVALUACIONES

	Según el papel que cumple la evaluación	Según el contenido o fase	Según el momento	Según el agente evaluador
Instituto Mujer II (93-95)	Sumativa	Proceso y Resultados	<i>Post-facto</i>	Mixta Diseño interno, parte ejecución externa
Cataluña II (94-96)	Sumativa	Diseño, Proceso y Resultados	<i>Post-facto</i>	Mixta Diseño con asesoría externa, ejecución interna con colaboración externa
País Vasco II (95-98)	Formativa/ Sumativa	Diseño, Proceso y Resultados	<i>Post-facto</i>	Mixta Diseño con asesoría externa, ejecución interna con colaboración externa

EVALUACIONES INTERNAS

Son las evaluaciones que se realizan a través de procesos internos de evaluación, es decir, llevados a cabo y dirigidos, en su totalidad, por el personal técnico del organismos de igualdad correspondiente. Esto significa que todo el esfuerzo evaluativo se realiza por parte del propio organismo de igualdad.

Estas son:

Evaluación del **I Plan Aragonés (1994-1996)**

Evaluación del **(II) Plan Integral Riojano (1996-1999)**³

³ La experiencia de evaluación aquí analizada es el proceso interno de evaluación que se lleva a cabo en los primeros años de vigencia del Plan. También como parte de la evaluación del II Plan estaba contemplada una evaluación final, ésta de carácter externo (que, de hecho, se contrata en el año 2000). Todo el diseño y la recopilación y análisis de la información, así como la redacción de las memorias anuales se realiza de forma interna, salvo, y de forma excepcional, la redacción de la memoria anual de 1999, para la que se pide ayuda externa. En cualquier caso, nosotros mantenemos la experiencia de evaluación riojana como evaluación interna, ya que son los datos e información que hemos podido manejar (las aportaciones externas no estaban todavía disponibles cuando se cierra el trabajo de campo para esta investigación).

Sólo dos de las once experiencias analizadas pueden ser consideradas como evaluaciones internas puras y, en uno de los casos, La Rioja, se recurre a una colaboración externa al final del proceso.

En el caso de la evaluación aragonesa, ésta se realiza durante el último año de vigencia del Plan y la misma corresponde a un esfuerzo de conocer el “estado de la cuestión” del Plan por parte del nuevo equipo que llega ese año al organismo de igualdad. En el caso de La Rioja, la evaluación se concibe como una evaluación continua interna, desde el mismo inicio del Plan. Desde el inicio del Plan, también estaba contemplada la realización de una evaluación al finalizar el Plan (año 2000), para la que era previsible que se requiriera algún tipo de colaboración externa, como de hecho finalmente ocurre⁴. Sin embargo, dado que al inicio de este trabajo dicha colaboración externa no se había dado, y dado que el proceso de evaluación continua se ha mantenido de forma interna hasta el penúltimo año, se ha decidido mantener esta clasificación como evaluación interna.

Las dos evaluaciones internas aparecen en la tabla 5.0.c, en la que también se refleja su clasificación según las otras tres tipologías de evaluación.

TABLA 5.0. c.: TIPOS DE EVALUACION
Evaluaciones internas
SEGUNDA GENERACION DE EVALUACIONES

	Según el papel que cumple la evaluación	Según el contenido o fase	Según el momento	Según el agente evaluador
Aragón I (94-96)	Formativa/ Sumativa	Proceso y Resultados	<i>Post-facto</i>	Interna
La Rioja II (96-2000)	Formativa/ Sumativa	Proceso y Resultados	<i>Post-facto</i>	Interna

A partir de ahora se exponen los datos obtenidos por cada experiencia según los criterios de meta-evaluación establecidos, y también se utilizará esta clasificación, basada en la tipología de evaluaciones según el agente evaluador.

⁴ Además de la evaluación final también se pide a una empresa externa que realice la memoria anual del año 1999, ya que se considera que es mucho trabajo para asumir internamente y dado que la técnica que hasta entonces se había ocupado de las mismas estaba de baja maternal.

2. EL PLANTEAMIENTO DE LAS EVALUACIONES Y LAS ESTRATEGIAS EMPLEADAS: CONTEXTO, PROPOSITOS, ESTRUCTURAS Y RECURSOS

En los puntos que se desarrollan a continuación se presentan, en primer lugar, los datos de forma generalizada -panorama general- ya que los aspectos estudiados pertenecen al contexto en el que se generan los estudios y procesos de evaluación. De esta manera, el análisis de dichos aspectos puede precisamente ofrecer luz sobre el tipo de evaluación elegida, factor que sirve para su clasificación posterior en evaluaciones externas, mixtas e internas. En una segunda parte de cada uno de los puntos tratados se exponen datos de cada tipo de evaluación.

2.1. Los contextos evaluativos

Panorama general

En el año 1995, los planes de igualdad como herramienta básica y principal de las políticas de igualdad de género se habían generalizado en todas las comunidades autónomas. Salvo en Navarra (en donde precisamente en ese mismo año se crea el Instituto Navarro de la Mujer), en las restantes comunidades autónomas se habían ya aprobado planes de igualdad. También en todas las comunidades autónomas (salvo en La Rioja, cuya política de igualdad se enmarca en la Dirección General de Bienestar Social) existen organismos específicos de igualdad⁵. En las elecciones autonómicas del año 1995, algunos gobiernos de las comunidades autónomas cambian de signo político, dando lugar a un panorama autonómico mayoritariamente conservador. Este mismo cambio ocurrirá en las elecciones generales de marzo de 1996. Hasta entonces, las políticas de igualdad de género estaban identificadas con el PSOE y con los partidos nacionalistas en el Gobierno de Cataluña y el País Vasco. Incluso existía la creencia de que dichas políticas iban a ser drásticamente abortadas por el partido conservador (PP); al menos, el discurso político era contrario a la existencia de organismos específicos de igualdad (Valiente, 1995). Sin embargo, estas políticas son asumidas en todas las

⁵ En 1995, la Comunidad Autónoma de la Región de Murcia cambia su anterior Dirección General de la Mujer en una nueva y significativamente diferente "Dirección General de Política Social y Familia"; en 1997 se reconvierte en Dirección General de Juventud, Mujer y Familia y se adscribe a la Consejería de Presidencia.

comunidades autónomas, entre otros factores, por la influencia europea e internacional⁶

Es decir, existen dos factores clave para entender el contexto en el que se genera esta segunda generación de evaluaciones. En primer lugar, la *generalización de los planes de igualdad* en las comunidades autónomas, reforzados entre otros factores, por el panorama internacional. Y, en segundo lugar, los *cambios de gobierno en algunas de las comunidades autónomas*. Precisamente, una vez asumidas las políticas de igualdad de género por todos los gobiernos autonómicos (incluidos los de signo más conservador) como un terreno propio, el cambio de gobierno en algunas comunidades autónomas ha resultado, en algunos casos, un cierto factor facilitador de la evaluación, ya que las nuevas responsables necesitaban saber y aprehender que es lo que se había hecho hasta entonces.

En el análisis del contexto evaluativo se han tenido en cuenta las siguientes variables: **organismo** responsable, tipo y antigüedad del mismo; **departamento** al que está inscrito dicho organismo; **partido en el gobierno** en el momento de la evaluación y cambios producidos en dicho período; si han existido **cambios en las responsables** de los organismos de igualdad que han hecho y/o encargado la evaluación. Por último se ha tenido en cuenta la historia respecto a los planes y respecto a la evaluación de los mismos (es decir, si existía **experiencia evaluadora previa**). Estas son las variables que se contemplan de forma conjunta y resumida en la tabla 5.1. que se expone a continuación.

⁶ La IV Conferencia de Pekín se celebra en septiembre de 1995 y su Plataforma de Acción fomenta y avala la intervención de los poderes públicos a través de organismos específicos y planes de igualdad. Para un análisis en mayor detalle, véase el capítulo 3.

TABLA 5.1.
EL CONTEXTO EVALUATIVO
SEGUNDA GENERACION DE EVALUACIONES

	Organismo responsable y año creación	Partido en gobierno (1)	Adscripción a Departamento de (2)	Cambios de responsable	Planes de igualdad	Experiencia anterior
INSTITUTO DE LA MUJER	Instituto de la Mujer (1983)	PSOE PP (3/96)	Cultura Asuntos Sociales (88) Trabajo y Asuntos Sociales (3/96)	C. Bustelo C. Mez.-Ten (88) P. Gutiérrez (91) M. Subirats (93) C. Dancausa (96) P. Dávila (2000)	I (88-90) II (93-95) III (97-2000)	SI
ANDALUCIA	Instituto Andaluz de la Mujer (1988)	PSOE/PA	Presidencia	C. Olmedo (88) T. Jiménez (2000)	I (90-92) II (95-97)	SI
ARAGON	Instituto Aragonés de la Mujer (1993)	PP/PAR	Presidencia	C. Corral C. Sanromán (2/96) A.de Salas (2000)	I (94-96) II (97-2000)	NO
CANARIAS	Instituto Canario de la Mujer (1994)	CC/PP	Sanidad y Asuntos Sociales Empleo y Asuntos Sociales (95)	P. Monzón E. Nuez (95) R. Dávila (99)	I (95-96) II (97-2000)	NO
CASTILLA-LA MANCHA	Dirección General de la Mujer (1989)	PSOE	Bienestar Social	M.C. G ^a Hidalgo A. Sanroma (97) A. Navarro (2000)	I (90-94) II (95-99) III (99-2003)	NO
CASTILLA Y LEON	Secretaría (3) Regional de la Mujer (1994)	PP	Sanidad y Bienestar Social	A. Parra (95)	I (94-96) II (97-2000)	NO
CATALUÑA	<i>Institut Català de la Dona (89)</i>	(CiU)	Presidencia	N. Llorach (89) T. Fandos (2000)	I (89-92) II (94-96) III (98-2000)	SI
GALICIA	<i>Servicio Galego de Promoción da Igualdade do home e da muller (1991)</i>	PP	Trabajo y Servicios Sociales Familia, Mujer y Juventud (95)	C. Bianchi (95) A. Gallego (97) Aurora Montes (98)	I (92-94) II (95-97) III (98-2000)	NO
MADRID	Dirección General de la Mujer (1989)	PP	Presidencia Sanidad y Servicios Sociales (95)	I. Alberdi Carmen Sáez H. Corbi M.L. Avez.Durante A. Miura (95)	I (89-91) II (93-95) III (97-2000)	NO
LA RIOJA	Dirección General de Bienestar Social (1988)	PP	Salud, Consumo y Bienestar Social	S. Loza (95)	I (91-95) II (96-99)	NO
PAIS VASCO	<i>Emakunde (1988)</i>	PNV-EAJ/ PSOE/ EA	Presidencia	T. Arteaga (88)	I (91-94) II (95-98) III (99-2002)	SI

¹ En el momento de la evaluación.

² En esta y las siguientes columnas, aparece en negrita la situación en el momento de la evaluación.

³ En el año 2000 se convierte en Dirección General de la Mujer e Igualdad de Oportunidades.

Es evidente que las estructuras institucionales más fuertes del panorama autonómico, en lo que se refiere a las políticas públicas de igualdad de género, son Andalucía, Cataluña y el País Vasco: son los institutos creados con mayor antelación, son organismos autónomos bien colocados en el respectivo organigrama gubernamental (adscritos a presidencia) y que, además, han gozado de una mayor estabilidad, incluyendo las responsables, ya que tienen como directoras a las mismas personas desde su creación (en los tres casos, hace ya más de una década)⁷. Esto, lógicamente se refleja también en la capacidad evaluativa: sólo en estos tres casos, además de en el del IM, se tenía experiencia evaluativa previa. Estos tres casos son, además, los que presentan unos segundos procesos de evaluación más completos, y con mayor alcance y eficacia evaluativa, como se irá comprobando en las siguientes páginas.

El IM (IM) es una estructura institucional fuerte por ser el primer organismo de igualdad que se crea por ley en España, por ser el antecedente claro de los organismos autonómicos (sentando el precedente de crear organismos específicos y hacerlo a través de la figura del organismo autónomo), por el propio nivel gubernamental al que representa (que le coloca como impulsor y coordinador de las políticas de igualdad en el territorio nacional, y que le otorga una función de liderazgo ejercida a través del mecanismo de las Conferencias Sectoriales⁸) y porque ha sido la correa de transmisión de las influencias internacionales en los primeros años, además de ser el representante de las políticas de igualdad de género en dicho ámbito internacional. Su posición en el organigrama gubernamental, sin embargo, es más débil que la de algunos de sus homólogos autonómicos y, a medida que se desarrolla el Estado de las Autonomías en España y las comunidades autónomas desarrollan sus propias políticas de igualdad, el IM debe ir redefiniendo su espacio y su papel. El IM sufre, además, un cambio gubernamental sustancial en el año 1996, que hace que se renueve el equipo directivo, y un cambio de sede al poco tiempo. Estas circunstancias hacen que parte del equipo técnico y funcional también se renueve⁹.

Como se puede observar en la tabla 5.1, los **organismos** responsables de los procesos de evaluación, es decir, los que los realizan o encargan, son los organismos de igualdad que formulan e impulsan los planes de igualdad evaluados. En el año 1998, siete de esos once organismos de igualdad que llevan a cabo procesos de evaluación en

⁷ En el caso de Andalucía, la directora del Instituto Andaluz de la Mujer desde su creación, Carmen Olmedo, abandona su cargo a principios del año 2000, para presentarse y posteriormente ser elegida como diputada al Parlamento por el PSOE, en las elecciones generales de Marzo de dicho año. También cambia la directora del Instituto Catalán de la Mujer en dicho año.

⁸ Este mecanismo intergubernamental es siempre unidireccional, en el sentido de que es siempre el Ministerio correspondiente el que convoca a las comunidades autónomas, y no al contrario.

⁹ Este cambio de equipo es importante ya que uno de los factores que definen los primeros años del IM es un equipo técnico y funcional muy sensibilizado y comprometido con la causa feminista y la misión del IM. Su primera directora, Carlota Bustelo, puso especial empeño en formar un equipo con estas características persuadiendo a funcionarias de carrera con sensibilidad feminista para que se incorporaran al IM. Este nivel de implicación personal resulta sin embargo difícil de mantener a lo largo del tiempo a medida que las instituciones crecen y se consolidan y, sobre todo, cuando cambian los equipos responsables.

esta segunda generación eran organismos autónomos¹⁰. Sobre el total del nivel autonómico (17 comunidades autónomas con sus correspondientes organismos) y central (IM), el número de estos organismos autónomos es de ocho, es decir, el único organismo de igualdad, que siendo organismo autónomo, no realiza proceso de evaluación de un plan de igualdad es el Instituto Navarro de la Mujer, cuestión lógica ya que dicho Instituto se crea en el año 1995 y hasta 1998 no aprueba su I Plan de Igualdad, por lo que el planteamiento de su evaluación deberá ser necesariamente posterior. Con esto se constata, que aunque también se han dado procesos de evaluación en organismos con otro formato distinto al de organismo autónomo, el hecho de tener una estructura institucional fuerte parece que facilita la existencia de procesos de evaluación.

No es el caso en lo referente a la **adscripción** a Presidencia, aunque este factor sí resulta importante a efectos de la posición de los organismos de igualdad en el organigrama gubernamental y a su autoridad y su capacidad de persuasión. Paradójicamente, parecería que la adscripción al área de Bienestar Social podría provocar una relativa mayor tendencia a la evaluación. De hecho, se podría aventurar una posible explicación, ya que hay sectores de política que, en general, tienen una mayor cultura evaluativa; éste es el caso del área de los Servicios Sociales, en los que se pueden encontrar mayores experiencias sistemáticas de seguimiento y evaluación (sobre todo referidas a la definición de indicadores) que en otros sectores. Sin embargo, además de la simple presencia o ausencia de procesos evaluativos, se puede sostener que la adscripción al área de presidencia sí puede facilitar una visión y aplicación más global de dichos procesos de evaluación, como se verá más adelante.

Los **cambios de responsable** en los organismos de igualdad también son una variable interesante a examinar del contexto evaluativo. La ausencia de cambios, ya se ha visto, es un factor que da estabilidad al organismo y, por lo tanto, hay mayores probabilidades de continuidad en sus acciones, incluida una política de evaluación. En algún caso, como es el del Instituto Aragonés de la Mujer, un vacío temporal de casi ocho meses entre el cese de una directora y el nombramiento de la siguiente, afectó al hecho de tener que plantearse la evaluación muy avanzado el Plan y con poco tiempo para su ejecución. Va a ser importante, además, la sensibilidad que cada responsable tenga personalmente hacia los temas de evaluación¹¹. En el caso del IM, aunque se pueden observar varios cambios de responsable durante la época PSOE, en realidad existe una continuidad; incluso, el IM se ve reforzado en 1988 con la creación del Ministerio de Asuntos Sociales, al que se adscribe el IM, ya que Carlota Bustelo, primera directora del mismo, pasa a ser la Subsecretaria de dicho Ministerio y tanto Matilde Fernández como Cristina Alberdi (primera y segunda ministra de dicho nuevo Ministerio) tienen un fuerte compromiso feminista (Valiente, 1996). Es decir, en la época del PSOE hay una cierta continuidad. Aún así, es muy probable que una mayor estabilidad en las responsables, hubiera permitido una política de evaluación más decidida que hubiera jugado un papel más importante del que ha tenido en el fomento,

¹⁰ En agosto de 1999 se crea el Instituto Asturiano de la Mujer, que no llega a ser un organismo autónomo, sino un órgano desconcentrado adscrito a la Consejería de Presidencia, creado por decreto y no por ley. En el año 2000 se crea el Instituto Balear de la Mujer, éste sí como organismo autónomo.

¹¹ Para un análisis pormenorizado del discurso que las responsables tienen sobre la evaluación, véase capítulo 6.

apoyo y coordinación de las evaluaciones de los diferentes planes autonómicos¹². En cualquier caso, el proceso de evaluación del II Plan nacional se produce a lo largo del cambio político de gobierno, ya que comienza antes del cambio y finaliza después del mismo; la continuidad de dicho proceso, en este caso, responde en parte a la continuidad del Consejero Técnico de Estudios del IM, que es quién impulsa y da continuidad al proceso de evaluación.

Efectivamente, el papel de los responsables técnicos en los procesos de evaluación suele ser importante y puede suplir, en algunos casos, la ausencia o cambio de liderazgo político. En Madrid, por ejemplo, en el que se dan importantes cambios con las elecciones autonómicas de 1995¹³, y se da un cambio de responsable, Consejería y sede, el diseño de la evaluación del II Plan ya estaba formulado por el equipo anterior. Aunque el pliego de condiciones técnicas establecido para el concurso público sobre el estudio de evaluación sale en el mes de noviembre de 1995, éste estaba realizado por el equipo de funcionarios previamente al cambio y se puede observar que dicho pliego establece unas condiciones técnicas muy concretas¹⁴.

El contexto evaluativo de las evaluaciones externas

Los contextos evaluativos en los que se realizan las seis experiencias externas de evaluación son muy heterogéneos entre sí y no especialmente favorables según los criterios de autonomía y posición gubernamental del organismo de igualdad correspondiente. De los seis organismos de igualdad que impulsan estos procesos de evaluación, sólo tres son organismos autónomos. De estos tres institutos, sólo Andalucía goza de una historia larga (desde 1988, ya que el Servicio Gallego es de 1991 y el

¹² Lo único que hace el IM al respecto es ya en Junio del año 1998, después de que varias comunidades autónomas reclamaran apoyo técnico en los temas de evaluación, en el que se organiza un seminario de dos días y medio de duración en la Universidad Internacional Menéndez Pelayo en Santander, sobre “Evaluación de Políticas de Igualdad de Oportunidades entre mujeres y hombres” y que dirigen la profesora Rocío Fernández Ballesteros (Universidad Autónoma de Madrid) y Mariano Alvaro (Consejero Técnico de Estudios del IM). Aunque este curso denota interés en el tema, esta acción es concreta, tardía y aislada en el tiempo y posteriormente no tiene continuidad en ningún seminario permanente o grupo de trabajo.

¹³ En Madrid, gana las elecciones autonómicas de 1995 el Partido Popular, hay una reestructuración importante de la Administración Autonómica y la Dirección General de la Mujer pasa de la Consejería de Presidencia a la recién creada Consejería de Sanidad y Servicios Sociales (que resultaba de la fusión de las anteriores Consejerías de Salud y de Integración Social).

¹⁴ Memoria explicativa para la contratación de un estudio para “elaborar un sistema de indicadores para evaluar el II Plan de Igualdad” (Dirección General de la Mujer. Noviembre 1995). En el caso de Madrid el estudio de evaluación se divide en dos fases que se sacan a dos concursos públicos, uno en noviembre de 1995 y el otro en junio de 1996. Las condiciones del concurso de junio de 1996 son, lógicamente, una continuación del primero (Pliego de Prescripciones Técnicas que regula la contratación de Consultoría y asistencia para la evaluación del II Plan de Igualdad de Oportunidades para las Mujeres de la Comunidad de Madrid).

Instituto Canario de 1994), de estabilidad en su equipo responsable y de una buena posición en el organigrama gubernamental (el Instituto Andaluz de la Mujer está adscrito a la Consejería de Presidencia, el Servicio Gallego de Igualdad a la Consejería de Familia, Mujer y Juventud, y el Instituto Canario a la Consejería de Empleo y Bienestar Social).

Las otras tres experiencias se encargan por parte de las Direcciones Generales de la Mujer de Castilla- La Mancha y de Madrid y por parte de la Secretaría Regional de la Mujer de Castilla y León, todas ellas en el área de Bienestar Social.

Parece, por lo tanto, que recurrir a procesos de evaluación totalmente externos, podría ser una forma de introducir la evaluación en contextos no especialmente favorables. Por un lado, es una forma de otorgar credibilidad a los procesos y, por otro, es también una forma de “externalizar” el trabajo de evaluación cuando no existe capacidad o recursos humanos internos disponibles para ellos (éste es claramente el caso de Canarias y, en cierto modo, el de Andalucía, ya que la primera experiencia andaluza de evaluación había costado mucho esfuerzo interno sin haber dado frutos especialmente gratificantes).

El contexto evaluativo de las evaluaciones mixtas

Los tres organismos que realizan procesos internos mixtos de evaluación -IM, Instituto Catalán de la Mujer e Instituto Vasco de la Mujer -*Emakunde*- son organismos autónomos, creados por ley. Son, además, de los organismos más antiguos, por lo que, en este sentido, nos encontramos ante un contexto evaluativo fuerte y estable.

Salvo el IM, que está adscrito al área de Asuntos Sociales, los otros dos tienen, además, una posición privilegiada en el organigrama gubernamental ya que se sitúan en el área de Presidencia. Como ya se ha visto antes, el IM, a pesar de no estar en el área de presidencia tiene un contexto institucional fuerte por su propia condición nacional y por haberse fortalecido de una forma excepcional en sus primeros diez años de vida.

Respecto a la estabilidad, tanto el Instituto catalán como el vasco han tenido, además, la misma responsable desde su creación, en el año 1989 y en el año 1988 respectivamente. Por otro lado, tanto en una Comunidad Autónoma como en la otra también se ha dado una estabilidad en el nivel político. Aunque dicha estabilidad es menor en el caso del IM, vemos que esta cuestión no afecta sustancialmente a la evaluación. De hecho, durante el proceso de evaluación del segundo plan se da un cambio en el partido de gobierno en el nivel central (marzo de 1996), pero éste no parece afectar sustancialmente al proceso de evaluación salvo en algún pequeño retraso. Esto es debido a que la evaluación se concibe, en este caso, con un carácter casi exclusivamente técnico (y como se analizará más adelante, ésta es precisamente, una de sus grandes debilidades); al mantenerse pues estable el responsable técnico encargado del estudio, el cambio político no afecta tanto.

Además, en los tres casos, existía una experiencia de evaluación previa. En el caso del IM y de Cataluña dichas primeras experiencias tuvieron muchas limitaciones (ver el capítulo 4); en el caso del País Vasco, la experiencia previa en evaluación había tenido mayores dimensiones y había sido mucho más contundente y continuada que en los otros dos casos.

En general, nos encontramos, en los casos de las evaluaciones mixtas, con los contextos evaluativos más propicios para la evaluación. Tanto en el caso de Cataluña como, y especialmente, del País Vasco, la apuesta de evaluación es fuerte. En cierto modo, se podría decir que la elección de realizar procesos internos en los que se involucra de forma especial el organismo de igualdad pero que, además, recurren a ayuda externa, es una forma de apostar por el desarrollo de la función de evaluación y de una cultura evaluativa en torno a las políticas de igualdad de género. Este quizá no sea tanto el caso del IM, en el que, como se ha comentado antes, la función de evaluación se relega más a una función técnica de estudio o investigación. Sin embargo, se podría decir que, incluso con sus limitaciones, el estudio de evaluación del II Plan nacional, cuenta con un contexto evaluativo favorable para ello.

El contexto evaluativo de las evaluaciones internas

Los dos casos de evaluaciones internas cuentan con un contexto evaluativo radicalmente distinto. En un caso, como es el aragonés, se cuenta con un organismo de igualdad que, aún siendo muy joven -se crea en 1993- y no muy estable en lo referente a la situación política autonómica y a sus responsables, es sin embargo potente, desde el punto de vista de su estructura -es un organismo autónomo- y de su posición en el organigrama gubernamental, ya que está adscrito al área de Presidencia. Son precisamente, y en este caso, los cambios políticos los que generan la realización de la experiencia de evaluación del I Plan aragonés que aquí se analiza, ya que como ya se ha dicho con anterioridad, esta evaluación responde a la necesidad de las nuevas responsables de saber qué se estaba haciendo.

En el caso de La Rioja, sin embargo, se trata de la estructura institucional más débil para un organismo de igualdad, ya que se trata de una unidad administrativa no específica -la Dirección General de Bienestar Social- y los temas de igualdad -más concebidos como temas de promoción de la mujer- se encuentran inmersos en una estrategia general de servicios sociales. Sin embargo, el hecho de que en dicha Dirección General ya se hubieran realizado y evaluado otros planes integrales -de minusvalías, por ejemplo- facilita el planteamiento de evaluación realizado para el II Plan integral de la Mujer analizado en este trabajo.

2.2. La claridad en los propósitos y las funciones de las evaluaciones

Panorama general

Uno de los factores más importantes a analizar en cualquier evaluación es la claridad o ambigüedad del planteamiento de dicha evaluación a través de sus propósitos y de las funciones que se espera cumplan dicha evaluación. Normalmente estos propósitos se formulan de manera explícita o bien en los pliegos de condiciones técnicas, si el estudio de evaluación sale a concurso público, o bien en los proyectos de evaluación si es por convenio o adjudicación directa; pero esto es en los casos de evaluación externa (este tipo de documentos no existen en casos de evaluación interna). Lógicamente, los propósitos y objetivos de un proceso de evaluación deberían hacerse explícitos en los propios informes que se generan a raíz de dichos procesos; pero esto no siempre es así, como veremos en el punto referido a los informes. En algunos casos nos ha resultado difícil encontrar de forma explícita los propósitos u objetivos de la evaluación. Sólo suelen aparecer reflejados claramente cuando se tiene que justificar, fundamentar o explicar el estudio de evaluación. Este es el caso de alguna comparecencia ante el poder legislativo correspondiente o de la presentación de ponencias sobre los estudios de evaluaciones en foros especializados (por ejemplo, el seminario sobre evaluación de políticas de igualdad organizado por el IM en la UIMP en el año 1998).

A pesar de la falta de claridad en los propósitos, prácticamente todas las experiencias de evaluación metaevaluadas mencionan, de una manera u otra, alguna o todas las funciones básicas de la evaluación: mejora, rendimiento de cuentas o responsabilidad y conocimiento o iluminación para acciones futuras. Algunas experiencias parecen hacer más hincapié en la función de rendimiento de cuentas (IM, Cataluña, y País Vasco), otras en la de ser una base para elaborar el siguiente plan (Canarias, Castilla-La Mancha, también Cataluña) y otras simplemente en conocer qué es lo que ha ocurrido (Aragón, La Rioja y también Canarias). Respecto a los objetivos, hay algunas experiencias que los presentan algo más integrados o interdependientes entre sí (Andalucía, Galicia y País Vasco) y otras en las que los objetivos se definen por separado, según las diferentes partes de las que se compone el estudio de evaluación. Hay evaluaciones mucho más ambiciosas en sus propósitos (IM, Andalucía y Madrid) y otras mucho más modestas (Aragón, Canarias, Castilla-La Mancha, La Rioja). Es interesante también destacar cómo algunas de la experiencias exponen, a la vez de sus propósitos, sus limitaciones (Castilla-La Mancha, Cataluña y, en cierto modo, País Vasco) siendo más realistas, e incluso generando una mayor credibilidad que el resto. Por último también interesa señalar cómo únicamente el País Vasco¹⁵, tiene entre sus

¹⁵ En cierto modo, también Madrid y Castilla y León plantean la idea de construir una metodología permanente de indicadores, que permita la evaluación de planes futuros. Aunque, en realidad, más que a la evaluación de planes, se refieren a la construcción de un sistema de indicadores para medir de forma

propósitos cuestiones relacionadas con la evaluación en sí misma, es decir, que les interesa y pretenden desarrollar la función de evaluación a la que dan, además, prioridad, entre otras funciones como organismo de igualdad.

Por último, se quiere señalar que, como ya se ha visto en el capítulo 1 de este trabajo, en los procesos de evaluación son importantes no sólo los propósitos explícitos, sino que también resulta interesante analizar la existencia de posibles propósitos implícitos o encubiertos. Este tipo de propósitos son muy difíciles de analizar y hacer explícitos y requerirían de unos estudios de cada uno de los casos hechos mucho más en profundidad de lo que este trabajo puede abarcar. Sin embargo, nos parece percibir una intención, en el caso del IM y en los casos de Andalucía, Madrid, y Castilla y León, de “aprovechar” la situación de evaluación para hacer investigaciones sobre la situación social de la Mujer, cuestión indudablemente necesaria y útil para la formulación de políticas de igualdad de género, pero que no pueden ser calificadas como evaluaciones de los planes de igualdad propiamente dichas. En el caso de Andalucía, el estudio sobre la situación social de la mujer se combina con la evaluación del II Plan. En los restantes casos mencionados los estudios de evaluación se convierten, en sí mismos, en estudios sobre la situación social de la mujer, minimizando la parte de evaluación de los planes supuestamente evaluados (en algún caso siendo prácticamente inexistente).

Los propósitos y objetivos de las evaluaciones externas

Si los propósitos de los estudios de evaluación son heterogéneos entre sí, tanto en lo que respecta a su grado de ambición como a su grado de claridad y explicitación, esta heterogeneidad es especialmente alta si se observan los propósitos y objetivos de las seis evaluaciones externas.

Respecto al diferente grado de ambición, hay que tener en cuenta que entre estas seis experiencias se encuentran estudios que cuestan menos de dos millones de pesetas (Canarias y Castilla y León) y alguno de más de cincuenta (Andalucía). Por lo tanto es lógico que haya diferencias a este respecto. En general, parece que cuando los estudios son especialmente pequeños, los equipos externos “se curan en salud” y señalan unos propósitos acordes con las dimensiones de su estudio. En algunos casos, como en Galicia, los objetivos iniciales son más ambiciosos de lo que la dimensión real del estudio luego resulta y en otros, como en Castilla- La Mancha, los objetivos son más realistas y ajustados a la dimensión del estudio y al proceso que realmente se lleva a cabo.

En algunos casos, los objetivos y propósitos se hacen explícitos más claramente, especialmente si existe un convenio (Andalucía) o unos pliegos de condiciones técnicas

continuada la situación de la mujer y las desigualdades de género, más al estilo de “observatorios de igualdad”. En cualquier caso, no es comparable con la preocupación y apuesta decidida por la función de evaluación del País Vasco.

(Madrid) para la realización del estudio. En casi todas las experiencias analizadas, los objetivos se explicitan por parte del equipo externo Andalucía, Canarias, Castilla- La Mancha, Castilla y León y Galicia, pero en otras, como Madrid, están previamente establecidos ya que se abre un concurso público para la evaluación. Mientras en Andalucía, Castilla y León y Madrid se hace hincapié en propósitos relacionados con la investigación, en otras se hace en el control y rendimiento de cuentas o la sistematización (Canarias y Castilla- La Mancha) y en otras en el análisis de los planes para la mejora de los mismos (Galicia y, en cierto modo también Canarias y Castilla- La Mancha).

Analizando los propósitos y objetivos de cada experiencia más en detalle:

La evaluación del **II Plan andaluz (1995-1997)** es muy ambiciosa respecto a sus objetivos. En el proyecto de evaluación que hace el IESA (Instituto de Estudios Sociales Avanzados de Andalucía, un centro mixto de la Junta de Andalucía y del CSIC) se señala que se pretende realizar una evaluación global del Plan (“evaluar todas y cada una de las fases del Plan, con el fin de que los responsables del mismo tengan un conocimiento fidedigno tanto de su correcta ejecución y futuro impacto, como de la adecuación del diseño” p. 17). Realmente es el único proyecto de evaluación (junto con el que se realiza para evaluar el II Plan castellano-manchego) que se plantea explícitamente realizar, además de una evaluación de la implementación o ejecución y de los resultados e impacto, una evaluación del diseño y de los objetivos del Plan. Esta evaluación se plantea de forma global, dejando claro que, aunque las distintas fases de la evaluación se traten por separado, todas están íntimamente relacionadas entre sí. En el propio proyecto de evaluación se contempla también la realización de un estudio sobre la situación social de la mujer en Andalucía en 1998, para comparar con otro estudio realizado en 1988. Con este estudio, realizado a través de una encuesta, se pretende analizar los cambios producidos en la situación social de las mujeres andaluzas en dicho período de tiempo; aunque forma parte del diseño de evaluación (entre otras encuestas y recopilación variada de información), se deja claro que es un apoyo más, y no pretende ser la evaluación de resultados. Textualmente, “el objetivo de este estudio, además de servir de soporte para la recogida de indicadores para la evaluación del II Plan Andaluz para la Igualdad de las Mujeres, es conocer la opinión de las mujeres andaluzas sobre sus vidas en relación con la sociedad en la que están inmersas, sus problemas, sus actitudes y sus necesidades” (IESA, Avance del Informe de Evaluación, julio 1997, p. 63, subrayado mío).

Además de este planteamiento, la entidad que realiza la evaluación externa (IESA), pretende proporcionar al estudio el “máximo rigor científico en un doble sentido: aplicando los conocimientos sociológicos más avanzados y abordando la evaluación como un proyecto científico de investigación” (p. 18). Es decir, además de ser una evaluación que sirva como apoyo a las decisiones políticas, pretende ser una evaluación que genere conocimiento científico. Desde nuestro punto de vista, hacer una evaluación con el máximo rigor no tiene por qué implicar necesariamente abordarla como un proyecto científico de investigación, ya que los propósitos de la investigación

son distintos (o, al menos, las prioridades son distintas) a los de la evaluación¹⁶. Como se verá más adelante, creemos que este propósito de investigación puede haber sido precisamente uno de los factores por los que la evaluación del II Plan andaluz ha generado problemas de coordinación interna-externa y se ha retrasado en su entrega, haciendo claramente peligrar su objetivo último de utilidad para el proceso de toma de decisiones.

El Instituto Canario de la Mujer encarga un estudio externo del **I Plan canario (1995-1996)**, al que se denomina “auditoría”, y en ningún momento aparece el término de evaluación. Es posible que la utilización del término “auditoría” tenga que ver con la intención de realizar un estudio que haga hincapié en la comprobación de lo que se ha hecho hasta el momento, más que una verdadera valoración del Plan. En el informe de auditoría, el equipo técnico que la realiza (Asesores en Imagen Corporativa) se plantea el análisis de las actuaciones desarrolladas por el ICM como por el resto de los organismos, departamentos y Consejerías, pretendiendo: 1) “Conocer el alcance de las acciones realizadas” (no explicitan a qué se refieren con el concepto de “alcance”, si es referido a las personas o grupos destinatarios, impacto, etc.), así como 2) El grado de cumplimiento de los objetivos previstos. Como propósito explícito el informe de auditoría “ha sido elaborado con la pretensión de convertirse en manual de referencia para los debates (sobre el II Plan de Igualdad)” (p.6).

Ya se ha visto que la evaluación del **II Plan Castellano-Manchego (1995-1999)** se realiza poco tiempo antes de que se elabore y apruebe del III Plan y que el mismo equipo externo que realiza el estudio de evaluación es también el equipo que asesora y coordina el proceso de elaboración de dicho III Plan. Es decir, la intención inicial es utilizar la evaluación como base para elaborar el siguiente Plan, aunque como se verá posteriormente, la premura del calendario no va a permitir que esto ocurra en su plenitud. El equipo que propone la evaluación es consciente de las limitaciones con que se encuentran (además de un plazo de realización muy corto, la ausencia de procesos de seguimiento y evaluación intermedia del Plan y la falta de una estructura de coordinación de los distintos organismos implicados en su desarrollo), y, en consecuencia, pretende ser realista en los propósitos y objetivos planteados. Para el equipo externo que la realiza (Fundación Mujeres) “la propuesta de evaluación que se realiza aportará como valor primordial la sistematización de toda la información existente sobre los niveles de ejecución de las actuaciones del Plan (...) y, como valor añadido, un análisis del diseño del Plan y una valoración de su desarrollo hasta este momento”. La propuesta se establece en 3 ejes: 1) Análisis del diseño del Plan, 2) Memoria/mapa de resultados y 3) Valoración del desarrollo del Plan. El objetivo general que proponen es el de conocer el grado de consecución de las acciones y, como específicos, 1) Detectar las bondades y dificultades que hayan podido surgir como consecuencia del propio diseño del Plan, 2) Conocer el grado de implantación del Plan: nivel de consecución y resultados obtenidos y 3) Posibilitar el diseño de las estrategias de intervención para facilitar el diseño de un nuevo Plan, sucesor del actual. Como se

¹⁶ Véase el punto 1.3.1. del primer capítulo.

puede observar, resulta curioso que el objetivo general no incluya el primer objetivo específico, ya que hacer una evaluación de diseño (que tiene que ver más con la evaluación de la idoneidad) va más allá y es claramente diferente a analizar si los objetivos de dicho plan se han cumplido (evaluación de la eficacia)¹⁷.

La evaluación del **I Plan Castellano Leonés (1994-1996)** se plantea abiertamente como un “análisis comparativo de la situación de las mujeres en Castilla y León”. Este análisis comparativo se realiza entre el estado de las mujeres antes de y después del periodo de vigencia del Plan, entre la situación de Castilla y León respecto al conjunto nacional y entre los hombres y las mujeres. Claramente no está planteada como una evaluación, desde sus objetivos, pero es denominada y expuesta como evaluación del I Plan¹⁸. El enganche con el I Plan aparece en la pág. 5 del resumen de la evaluación. Textualmente dice que “en este marco de referencia, se puede valorar la incidencia que han tenido las acciones llevadas a cabo por la Junta de Castilla y León a lo largo de los años incluidos en el Plan”¹⁹. No dice cómo se va a valorar dicha incidencia y a lo largo del estudio no aparece ninguna información que señale algún tipo de relación entre los cambios acaecidos en las mujeres en el periodo de vigencia del Plan y el propio Plan. Es decir, ¿cómo se puede llegar a la conclusión de que dichos cambios son debidos al Plan y no a otros factores? No se recoge ninguna información al respecto. Curiosamente el anteproyecto de evaluación que realiza “Eurodoxa” (empresa externa que realiza la evaluación), aunque se planteaba un estudio mucho más amplio y de mayor alcance, también presentaba unos objetivos en los mismos términos. Se decía que con el proyecto de evaluación se pretendía medir: el cambio social, la mejora de las condiciones de las mujeres, la inserción laboral y la mejora educacional, la atención a los colectivos de mujeres más necesitados y la mejora del conocimiento y la información. Es definitiva, todos los aspectos estaban relacionados con el estudio de la situación de las mujeres y no con la evaluación del Plan. Esto era contradictorio con lo que luego se planteaba en el mismo proyecto, ya que se proponía, además de la evaluación con fuentes administrativas y secundarias (que es lo que finalmente se hace), y de una “evaluación cuantitativa” para estudiar al situación social de la mujer a través de una encuesta, una evaluación cualitativa a través de entrevistas y grupos de discusión (con un diseño muy similar al del IM) y cuyo objetivo era evaluar las acciones realizadas, es decir, la

¹⁷ La información y las citas están extraídas del documento “Proyecto de evaluación del II Plan de Igualdad de Oportunidades. “Las Mujeres de castilla-La Mancha hacia el siglo XXI” (Fundación Mujeres. Octubre 1998).

¹⁸ En la publicación del I Plan, aparece un apartado en el que, bajo el título de “La responsabilidad sobre la evaluación del Plan”, se expone que la Secretaría Regional de la Mujer, se responsabilizará de la elaboración de un Informe de Seguimiento del Plan, que cubrirá “las iniciativas lanzadas, los recursos aplicados, la colaboración con otros organismos, los cambios generados por el Plan en los programas en los que ha hecho referencia a estudios de situación sociológica, y las áreas e iniciativas de éxito o de fracaso en el desarrollo del Plan”. Es decir, los objetivos iniciales sí eran algo más ambiciosos.

¹⁹ Documento: Secretaría Regional de la Mujer (Junta de Castilla y León), “Resumen de la Evaluación del Plan Integral de Igualdad de Oportunidades para la Mujer en Castilla y León 1994-1996. 78 pp. Sin fecha.

evaluación del Plan y no de la situación de las mujeres castellano leonesas.²⁰ Es decir, previamente existían dos objetivos: el estudio de la situación de las mujeres en la región y, aunque no planteado de forma clara, la valoración de las actuaciones del II Plan. En el estudio que finalmente se realiza se pretende únicamente el primero de los objetivos.

Con la evaluación del **II Plan gallego (1995-1997)** se pretendía:

- 1) “Dar respuesta, por parte de la administración, a las exigencias internas de mejora de la calidad de los servicios prestados.
- 2) Proporcionar elementos de juicio para la toma de decisiones en el ámbito de las políticas de igualdad de la Consejería.
- 3) Dar cuenta a la sociedad gallega del rendimiento de recursos dedicados a la promoción de la igualdad.
- 4) Mejorar y corregir las deficiencias de los planes de igualdad anteriores” (los subrayados son míos)²¹

Como objetivos concretos de la evaluación gallega, se plantean conocer la cantidad y calidad de las actividades desarrolladas, detectar sus puntos débiles y sus ventajas para mejorar, detectar nuevas necesidades, “utilizar los resultados de la evaluación para redefinir, en su caso, los planes estratégicos de actuación, con el fin de mejorar las prestaciones” y “proporcionar tanto a las usuarias de los servicios como a las entidades que los financian una información objetiva y fiable de la calidad de las prestaciones”. En definitiva, están presentes las tres funciones de la evaluación: mejora, rendimiento de cuentas e iluminación de acciones futuras.

El **II Plan madrileño (1993-1995)** se evalúa por parte de una empresa externa en dos fases diferentes, que dan lugar a dos concursos públicos entre los que median ocho meses. Al igual que en Castilla y León (se realiza por parte de la misma empresa externa –Eurodoxa–), se asemeja más a un estudio de la situación social de la mujer madrileña que a una evaluación del II Plan, aunque las intenciones iniciales y, que se reflejan tanto en los pliegos de condiciones técnicas como en los proyectos e introducciones de los informes de evaluación, son claramente evaluativas. El primer concurso público responde al título de “Asistencia técnica para la elaboración de un sistema de indicadores para evaluar los resultados obtenidos con el II Plan de Igualdad de Oportunidades para las mujeres de la Comunidad de Madrid” (noviembre 1995) y el segundo al de “Consultoría y asistencia para la evaluación del II Plan de Igualdad de Oportunidades para las mujeres de la Comunidad de Madrid” (julio 1996). Ya en el primer pliego de condiciones técnicas se señala la necesidad de valorar y evaluar las acciones contempladas en el II Plan, para poder medir su efecto. Incluso se llega a decir

²⁰ Documento: EURODOXA, S.A., “Anteproyecto para la Evaluación del Plan Integral de Igualdad de Oportunidades para la Mujer en Castilla y León 1994-1996. 14 pp. Sin fecha.

²¹ Tríptico: Evaluación del II Plan de Igualdad de Oportunidades para las Mujeres Gallegas. Información”

que los indicadores deberán “servir para medir el resultado obtenido con el esfuerzo inversor y ayudar a mejorar y, si fuera necesario, a modificar las acciones contempladas hasta la fecha”. El objeto del contrato es “la elaboración de un sistema objetivo y fiable de evaluación de los resultados obtenidos por la aplicación de las políticas de acción positiva integradas en el II Plan” (subrayados míos). El segundo contrato tiene por objeto la implementación de dicho sistema de indicadores el cuál, textualmente, “permitirá verificar en qué grado se han desarrollado medidas pertinentes para el logro de los objetivos propuestos en dicho Plan al tiempo que facilitará el conocimiento de medidas alternativas acordes con aquellas actividades mediante cuya aplicación no se hayan logrado los resultados previstos y proporcionará un modelo analítico para seguir y evaluar anualmente las actividades a desarrollar por la Dirección General de la Mujer”²². Es decir, lo que se pedía en los pliegos de condiciones técnicas era claro: el propósito inicial era la evaluación del II Plan y no el estudio de la situación de la mujer madrileña durante el periodo de vigencia de dicho Plan, que fue lo que posteriormente se realizó. Con esa evaluación se pretendía, además, no sólo verificar el cumplimiento del Plan, sino obtener alternativas para mejorar. Más adelante veremos como en el informe de evaluación no se ofrecen ni valoraciones sobre los datos recogidos ni recomendaciones, sugerencias o alternativas. Sin embargo, la ambigüedad en los propósitos de la evaluación queda patente cuando los responsables políticos y técnicos declaran que, además de la evaluación, el objetivo último era obtener datos sobre la situación de la mujer madrileña. En la publicación de los resultados de la evaluación, en su presentación, se mencionan varios propósitos: “evaluar los resultados e incidencia del II Plan”, “elaborar las herramientas necesarias para evaluar sucesivos planes de igualdad”, “conocer el grado de cumplimiento”, “comparar la situación actual de las mujeres madrileñas con la que existía en el año 1992” y “cotejar los datos referidos a nuestra Comunidad con los correspondientes al conjunto de España”. Como se puede observar, se entremezclan propósitos referidos a una evaluación -que no se realiza- y a una investigación sobre la situación social de la mujer, hecha en forma de “observatorio” periódico -que es lo que, en parte se realiza-. De hecho, la fundamentación que encontramos en dicha presentación es que “el diseño de las políticas de igualdad resultará más eficiente si se identifican cuáles son los problemas que van superándose a medida que la sociedad se transforma y cuáles son los que ofrecen resistencia o incluso se agravan” (DGM. *Las mujeres en la Comunidad de Madrid*. p.3-4). Es decir, tiene más que ver con una fundamentación referida a la necesidad de conocer la realidad sobre la que se actúa y la evolución del problema público sobre el que se pretende actuar, y no tanto con la necesidad de conocer los efectos de la actuación de los poderes públicos sobre dicho problema, ni con la necesidad de mejorar dicha actuación, ni con el rendimiento de cuentas. A pesar de la legitimidad, necesidad y clara conveniencia de esta función de investigación continua sobre el problema de la (des)igualdad de género, creemos que el hecho de que la DGM se haga responsable del estudio de la situación de las mujeres madrileñas no le debería eximir de la responsabilidad de evaluar sus actuaciones.

²² Documentos: Memorias explicativas y Pliegos de Condiciones Técnicas que regirán la contratación de “Asistencia Técnica para la elaboración de un sistema de indicadores para evaluar los resultados obtenidos con el II Plan de Igualdad de Oportunidades para las Mujeres de la Comunidad de Madrid” y “Consultoría y asistencia para la evaluación del II Plan de Igualdad de Oportunidades para las Mujeres de la Comunidad de Madrid”. Dirección General de la Mujer. 6 y 8 páginas respectivamente. Septiembre de 1995 y Mayo de 1996 respectivamente.

Los propósitos y objetivos de las evaluaciones mixtas

El **IM** se plantea la evaluación como una forma de dar cumplimiento a dos preceptos: el primero, referido a las competencias atribuidas al IM en la ley de creación del mismo (16/1983, de 24 de Octubre), por la que corresponde a este organismo el seguimiento de los Planes de Igualdad de Oportunidades de las Mujeres, a través del Consejo Rector; el segundo, referido al compromiso que se hacía en el propio II Plan, por el que se decía que al término de su período de vigencia debía elevarse un informe de evaluación al Consejo de Ministros (II Plan: p.139). También se menciona la función de facilitar la toma de decisiones en los planes subsiguientes. Los objetivos de la evaluación se concretan de la siguiente forma, en una ponencia presentada por el equipo técnico de evaluación en un seminario en la UIMP: “1) Conocer cómo se pone en marcha el Plan correspondiente, qué tipo de actividades se han emprendido para dar cumplimiento al Plan, 2) Estudiar qué resultados se han conseguido con la ejecución de los planes, o mejor²³, qué variaciones se han producido a lo largo de los años de vigencia de los planes, en aspectos considerados básicos e importantes, 3) Analizar cómo perciben la puesta en marcha de los planes personas o grupos que han estado relacionadas o relacionados muy directamente con la realización de las actividades llevadas a cabo y 4) Averiguar cuáles han sido los puntos débiles de la ejecución de los planes, así como los fuertes, con el fin de elaborar una serie de consideraciones o recomendaciones a tener en cuenta a la hora de elaborar el siguiente plan, por las personas responsables de las políticas de igualdad (...)” (Alvaro, De la Fuente y Aranda, 1998: 2). Este último objetivo demuestra la intención también de que la evaluación cumpla la función de mejora, además de la del rendimiento de cuentas y la iluminación para los planes futuros que ya se ha mencionado. Sin embargo, y como se verá a lo largo del análisis, no se realizan demasiadas acciones que puedan cumplir dicho objetivo: ni en los informes de evaluación ni en la síntesis de resultados de la Evaluación del II Plan, publicada por el IM, aparece ningún apartado de conclusiones ni recomendaciones.

La finalidad primaria de la evaluación del **II Plan catalán** fue proporcionar al Gobierno de la Generalidad y al Parlamento de Cataluña una visión global y sistematizada sobre los resultados conseguidos con el II Plan. Además, esto sirve también como punto de partida para la elaboración del III Plan. Es decir, sobre todo rendimiento de cuentas y, secundariamente, iluminación de acciones futuras. En concreto se quería dar respuesta a cuatro cuestiones básicas: ¿Qué se ha hecho en los tres años de vigencia del Plan (grado de cumplimiento del Plan)?, ¿Quién ha participado en la ejecución de las diferentes acciones?, ¿A qué colectivos se han dirigido las acciones del Plan? y ¿Qué recursos se han destinado?. En este estudio de evaluación se

²³ En la formulación de este objetivo se ve claramente la justificación que se intenta hacer de convertir un estudio de la situación social de la mujer en evaluación de resultados del II Plan. Claramente las dos partes del objetivo no son iguales: no es lo mismo estudiar qué resultados se ha conseguido con la ejecución del Plan que estudiar qué variaciones se han producido a lo largo de los años de vigencia del Plan. Esta contradicción se analizará con mayor profundidad más adelante.

reconocen las limitaciones y se insiste en que el énfasis de la evaluación ha sido más descriptivo que valorativo. Además, “respecto a la cuestión de cuál ha sido el impacto social del Plan, en pocas ocasiones se ha dispuesto de la evidencia empírica suficiente para pronunciarse claramente en términos de efectividad de las acciones y del grado de cobertura de la demanda social”. A pesar de este carácter modesto, se considera que la evaluación ofrece elementos de aprendizaje y reflexión (función de mejora y aprendizaje) a los diferentes actores institucionales y sociales implicados en la política de igualdad de género en Cataluña²⁴.

Finalmente, y en lo referente a la evaluación del **II PAPME vasco**, la principal finalidad a la que se hace referencia es la del cumplimiento de la encomienda que en dicho Plan se realiza al Instituto Vasco de la Mujer (*Emakunde*) de impulso, seguimiento y evaluación del Plan. De hecho, *Emakunde* se plantea dos objetivos primordiales en relación a dichas funciones: 1) Impulsar la participación activa de la Administración Pública, tanto en la ejecución como en la evaluación del II PAPME y, 2) Profundizar y mejorar la metodología de evaluación. En definitiva, es el único organismo específico de igualdad que tiene la evaluación de los planes como verdadera prioridad y, entre sus propósitos, se encuentran objetivos que tienen que ver con la evaluación en sí misma. Como objetivos de evaluación, se plantean: dar a conocer las actuaciones que se realizan desde los tres niveles de la Administración Pública Vasca en relación al Plan, analizar en qué medida están siendo o no ejecutadas las acciones propuestas en el II PAPME, analizar desde una perspectiva cualitativa las actividades realizadas por los diversos organismos implicados en la ejecución del Plan y realizar una valoración del conjunto de las actuaciones con objeto de orientar la ejecución del II PAPME en toda su integridad durante su período de vigencia. Es decir, tienen presentes tanto funciones de rendimiento de cuentas y de mejora, además con la idea de orientar sobre la marcha el propio Plan evaluado²⁵. En este caso, los propósitos de la evaluación están más claros que en el resto de las experiencias. De hecho, se plantean la “evaluación de la propia política pública” (ventajas y desventajas de la estrategia seguida y características específicas de las acciones desarrolladas) y renuncian explícitamente a la “evaluación del impacto de la política” (de los resultados que esta política ha tenido en función de los objetivos que se había propuesto conseguir) ya que “existían limitaciones que se estimó impedirían hacer una evaluación del impacto rigurosa cuestión que (...) ocurre también en el conjunto de los países de nuestro entorno”. En la evaluación del II PAPME se incorpora, además, un nuevo objetivo, que también se hace explícito claramente y que es aumentar el nivel de compromiso de la Administración Pública Vasca en la evaluación, aumentando su participación y mejorando e incrementando la información aportada por la misma. También son claras al diferenciar la evaluación de otro tipo de investigación, que también realizan: “No

²⁴ Documento: Institut Català de la Dona, “Informe d’avaluació del II Pla d’actuació del Govern per a la igualtat d’oportunitats per a les dones” Versió revisada 25/4/1997. 242 pp.

²⁵ Todas las referencias de Emakunde, “Memoria de Actuación de los poderes públicos en 1997 en ejecución del II Plan de Acción Positiva para las Mujeres en la Comunidad Autónoma de Euskadi” 159 pp. más anexo.

obstante Emakunde realiza anualmente, aunque no como parte integrante de la evaluación, un informe sobre la situación de las mujeres en Euskadi y su evolución”²⁶.

En definitiva, parece que los objetivos de Cataluña y el País Vasco tienen que ver con el desarrollo de la propia función de evaluación y no sólo con los objetivos del estudio en sí y parecen, además, algo más ajustados y realistas a las dimensiones de los estudios que en el caso del IM.

Los propósitos y objetivos de las evaluaciones internas

En el **Instituto Aragonés de la Mujer**, el estudio sobre el grado de cumplimiento que se realiza sobre el I Plan se enmarca dentro del proceso de seguimiento de dicho Plan, y se realiza poco antes de que el periodo de vigencia del Plan finalizara (en abril de 1996). Es decir, no tiene grandes pretensiones y, por lo tanto, tiene dos propósitos: por un lado establecer prioridades para lo que restaba de año, y por otro, saber qué acciones se habían realizado, ya que durante el año 1995 se da un cambio de responsable y un vacío de varios meses hasta que es nombrada la nueva directora.

La evaluación del **(II) Plan Integral de la Mujer de La Rioja** se enmarca en el proceso interno de seguimiento y evaluación de dicho Plan. Con ello se quiere obtener una información de manera continuada sobre los objetivos y las acciones (grado de cumplimiento del plan a través de recopilación sistemática y continuada de la información), los organismos implicados y la participación institucional (recopilación a través de fichas), así como la opinión de la población destinataria (al final del Plan, se harán entrevistas en profundidad y grupos de discusión). En La Rioja sí parecen tener clara la diferencia entre evaluación del plan e investigación sobre la situación de la mujer, ya que se proponen, además de las evaluaciones, la creación de un “Observatorio sobre la situación de las Mujeres en La Rioja” que consistirá en la elaboración de un sistema de indicadores que actúe de observatorio permanente y que permita conocer la evolución de la población destinataria en el tiempo de vigencia del Plan²⁷.

Es decir, los propósitos en estos dos casos son muy distintos: mientras en el caso aragonés el caso es cubrir unas necesidades concretas de información y que responden a una situación y contexto determinado, en el caso de La Rioja, se concibe la evaluación como algo inherente a la propia intervención, es decir, como una fase más de los planes, pretendiendo además normalizarse en la dinámica propia de dichos planes.

²⁶ Documento sobre evaluación, presentado por Emakunde al Seminario sobre Evaluación de Políticas de Igualdad en la UIMP, Santander, Junio 1998. p.6-7.

²⁷ Gobierno de LR, Plan Integral de la Mujer 1996-1999. p. 66-67.

2.3. Las estructuras institucionales de la evaluación

Panorama general

El marco institucional para la evaluación de planes de igualdad en España es todavía incipiente, aunque los procesos de evaluación de dichos planes están haciéndose crecientemente más explícitos. En este apartado analizamos cuál ha sido el marco institucional en la segunda generación de evaluaciones de planes de igualdad. Para este análisis del marco institucional se toman en cuenta dos factores: por un lado, si se menciona y se hace un compromiso explícito de evaluación en el plan evaluado y, por otro, el tipo de estructuras creadas alrededor de los organismos o los planes de igualdad y que tienen alguna responsabilidad en el seguimiento y la evaluación de los planes. También se tiene en cuenta la estructura funcional y técnica de los organismos específicos de igualdad, ya que sobre el personal técnico recae el peso ejecutivo de los procesos de evaluación. Conocer las dimensiones de la plantilla de estos organismos en el momento en el que se realizan las evaluaciones analizadas, ofrece una visión más completa de las estructuras institucionales que soportan los procesos de evaluación, bien sean internos o externos.

La **mención explícita sobre evaluación** es todavía incipiente en esta segunda generación de planes evaluados, a excepción del II PAPME vasco. Ya se ha visto que los compromisos explícitos sobre evaluación aumentan sustancialmente en la tercera generación de planes. Aunque tampoco la existencia de una mención explícita asegura necesariamente que exista un compromiso real ni que se cumpla dicho compromiso.

De los once planes evaluados de esta segunda generación que es objeto de nuestra meta-evaluación, sólo se hace referencia explícita a la evaluación en cinco de ellos. Esta referencia, además, suele ser una simple mención (salvo en el caso del País Vasco, en la que la mención es extensa y específica), a diferencia de la tercera generación de planes en los que se suele dedicar un apartado específico al sistema de evaluación. Las menciones a la evaluación que aparecen en cinco de los once planes metaevaluados son²⁸:

El IM, en su cuarto apartado del Acuerdo del Consejo de Ministros, dice: “Antes de finalizar el período de vigencia previsto para el Plan, la Ministra de Asuntos Sociales elevará un informe de evaluación al Consejo de ministros en el que expondrá el grado de cumplimiento alcanzado, las actuaciones puestas en práctica y las causas que hubieran motivado la no realización de algunas de las actuaciones previstas. Este informe incluirá un análisis de la incidencia social del Plan”.

²⁸ Todas las referencias y citas en este apartado aparecen en los correspondientes planes evaluados.

En el tercer apartado del Acuerdo del Consejo Rector del Instituto Canario de la Mujer, sobre el I Plan 95-96) se dice que dicho Consejo acuerda: *“Encomendar a la Dirección del Instituto Canario de la Mujer las funciones de impulso, seguimiento, coordinación, y junto con el Consejo Rector, la evaluación de las medidas contenidas en el Anexo del Acuerdo”*.

En el I Plan de Castilla y León, sí existe un apartado específico sobre “Evaluación y seguimiento del Plan”, en el que se especifica que será responsabilidad de la Secretaría Regional de la Mujer la elaboración de un Informe de Seguimiento del Plan. Posteriormente menciona el “Informe de Evaluación”, refiriéndose al mismo informe de seguimiento y dice que *“será sometido a la Junta de Castilla y León (contemplándose, además su difusión en el ámbito de la Dirección General de Servicios Sociales), y será informado a la opinión pública en el marco de las iniciativas de comunicación antes descritas”*. También se menciona que dicha evaluación deberá ser realizada anualmente.

También en el II Plan de La Rioja²⁹, existe un apartado específico para la evaluación, en el que se especifican tres tipos de evaluación: de los objetivos y acciones, de la participación institucional y de la opinión de la población destinataria, es expresa cuál va a ser la metodología general y los órganos de seguimiento y control del Plan (Consejo de Bienestar Social, Consejo Sectorial de la Mujer y Comisión Interinstitucional de Bienestar Social).

Finalmente, en el Acuerdo del Consejo de Gobierno Vasco por el que se aprueba el II PAPME, en su tercer artículo se acuerda: *“Encomendar a la Dirección del Instituto Vasco de la Mujer (EMAKUNDE) las funciones de impulso, seguimiento y evaluación generales del II Plan de Acción Positiva para las mujeres en la Comunidad Autónoma de Euskadi (a continuación se especifica el modo de efectuar dichas funciones)”*³⁰

En lo que respecta a las **estructuras** responsables del seguimiento y la evaluación, se pueden establecer seis tipos diferentes: 1) Consejos Rectores, 2) Estructuras creadas “ad-hoc”, 3) Estructuras no específicas de evaluación, 4) Consejos de la Mujer y 5) Organismos que deben informar a los correspondientes poderes legislativos. Se analizan a continuación cada uno de estos tipos:

1) Algunos **Consejos Rectores** de los Institutos de la Mujer (IM, Aragón, Canarias y País Vasco) son los principales responsables de la coordinación, impulso, seguimiento y evaluación de los planes, aunque en el caso de Canarias existe también la Comisión de Igualdad con responsabilidades en evaluación, y en el País Vasco, una Comisión Interdepartamental.

2) **Estructuras creadas “ad-hoc”** para la coordinación, impulso, seguimiento y/o evaluación de los planes de igualdad. Esto ocurre en Andalucía, Galicia, Madrid y el País Vasco. Sólo en Galicia se menciona explícitamente la palabra “evaluación”, es decir es una estructura específica de evaluación; las otras mencionan coordinación,

²⁹ El (II) Plan Integral de la Mujer de LR (1996-1999) no es un Plan aprobado por Consejo de Gobierno sino por el Consejo de Bienestar Social.

³⁰ En la “Propuesta de Acuerdo por el que se aprueba el II Plan de Acción Positiva para las Mujeres en la Comunidad Autónoma de Euskadi”, cuyo acuerdo adoptó el Consejo de Gobierno, en su Sesión celebrada el 29 de Diciembre de 1995, se especifican de forma muy amplia tres apartados respecto a este artículo tercero: uno referido al “Ejercicio de las funciones de coordinación, seguimiento y evaluación del Plan”, otro sobre “Métodos de coordinación, evaluación y seguimiento del Plan” y finalmente “Disposiciones comunes al ejercicio de las funciones de coordinación, seguimiento y evaluación”, en el que se incluyen hasta los aspectos que deben recoger los informes de evaluación.

impulso, seguimiento y control. Con la excepción de Madrid, en donde la comisión incluye una representación de grupos de mujeres y de los gobiernos locales, estas estructuras están integradas exclusivamente por representantes de los departamentos gubernamentales implicados en el plan.

Estas estructuras son distintas en su naturaleza. Algunas de ellas están más formalizadas, como son los casos de Andalucía, Madrid o el País Vasco (en dónde además se especifican claramente sus funciones), en los que dichas estructuras aparecen mencionadas y estructuradas en los propios acuerdos de gobierno que aprueban los planes correspondientes. En Galicia, sin embargo, la estructura es más interna, habiendo sido creada a raíz del comienzo del proceso de evaluación y por recomendación del equipo externo que lo llevó a cabo. También es distinto el rango de los miembros gubernamentales de dichas comisiones, siendo en Andalucía necesariamente viceconsejeros y quedando en las otras sin especificar, por lo que la representación suele quedarse en un nivel más técnico que político.

3) **Estructuras no específicas de coordinación y seguimiento de los planes:** en esta categoría se encuentran Canarias, Castilla y León y Cataluña. Con la excepción de la “Comisión para la Igualdad de la Mujer” en Canarias, en la que también hay representación de sindicatos y grupos de mujeres, el resto son estructuras gubernamentales. Aunque estas estructuras no son específicas de la función de evaluación, su existencia es importante para la misma, ya que son estructuras generales de coordinación intergubernamentales y son normalmente estructuras de apoyo de los organismos de igualdad, en todas sus funciones, incluida la evaluación.

4) **Consejos de la Mujer:** aunque normalmente no se suelen mencionar explícitamente responsabilidades en la evaluación de los planes por parte de estas estructuras (salvo en el caso de La Rioja), suelen cumplir también un papel simbólico respecto a las mismas. Por ejemplo, en el caso del Consejo Regional de Castilla y León, aunque no hay ninguna función explícita en su orden de creación, en su seno hay un grupo de seguimiento y control del II Plan, aunque no lo hubo para el I Plan. En el caso de Cataluña y La Rioja, los correspondientes Consejos tienen reconocida la función de seguimiento (en el caso de Cataluña, indirectamente, a través del Instituto Catalán de la Mujer). Pero el hecho de que exista un Consejo Regional de la Mujer o similar no implica que tenga funciones de seguimiento y evaluación de los planes (como es el caso de Castilla- La Mancha).

5) En algunas comunidades autónomas existen **vínculos con el parlamento o asamblea legislativa** correspondiente. En el caso de Cataluña, existe en el Parlamento catalán una “Comisión de Seguimiento del Proceso de Equiparación Hombre-Mujer” que es bastante activa pidiendo comparecencias a la presidenta del ICD. La existencia de comisiones relacionadas con el tema de la igualdad o la mujer también se dan en otros parlamentos o asambleas autonómicas, y en algunos casos se han producido también comparecencias relacionadas con los estudios de evaluación (por ejemplo, Andalucía³¹ o Madrid), aunque su actividad es más irregular. En Cataluña existe una comparecencia el 6 de mayo de 1997, en la que la presidenta del ICD explicaba la

³¹ “Comisión Permanente para la Promoción de la Igualdad de Oportunidades y Protección de los Derechos de las Mujeres”, creada en la 4ª legislatura, el día 4.9.1994 (B.O.P.A. nº 16 de 1994).

evaluación del II Plan³². En ella hablaba tanto de los resultados de la evaluación como de la metodología que se había utilizado para la misma. La existencia de una estructura que fomenta este esfuerzo de explicación y rendimiento de cuentas es sin duda un acicate más para los procesos de evaluación y refuerza la función de evaluación.

En el caso del País Vasco, el vínculo es más directo ya que en la propia estructura del organismo de igualdad (*Emakunde*) existe una vinculación y un compromiso explícito con el legislativo dado que en la propia ley de creación de *Emakunde* (Ley 2/88, de 5 de febrero) se señala que se debe presentar una memoria anual al Parlamento sobre las actuaciones de los poderes públicos en materia de mujer. Esta circunstancia ha obligado al organismo, desde un principio, a recopilar información sistemática sobre sus actuaciones.

Curiosamente, también existe una Comunidad Autónoma, Baleares, en la que existen ciertos vínculos formales con el poder legislativo con respecto al II Plan pero en la que, sin embargo, éste no ha sido evaluado. El Parlamento de las Islas Baleares aprueba unas propuestas de resolución respecto al II Plan, entre las que figuran que insta al Gobierno a darle cuenta del grado de cumplimiento del mismo y a presentar una memoria valorando los resultados del Plan. Parece, por lo tanto, que los vínculos formales con los organismos de igualdad, como es el caso del País Vasco, son más efectivos que los que se establecen a través de los propios planes.

6) Finalmente, existen **estructuras no específicas de igualdad** a las que el organismo de igualdad debe informar sobre el plan, como es el caso del Consejo de Bienestar Social en La Rioja, que es el que aprueba el Plan (a diferencia de las comunidades autónomas en las que se aprueba por Acuerdo de Consejo de Gobierno). También existe en este caso la Comisión Interinstitucional de Bienestar social, que también cumple un papel en el seguimiento y evaluación del (II) Plan Integral de la Mujer de La Rioja. Este caso es muy distinto y refleja cómo la política de igualdad forma parte de la política de Bienestar Social y, por lo tanto, utiliza sus estructuras³³.

Como se puede observar, entre los casos estudiados, existe una Comunidad Autónoma, Castilla- La Mancha, en la que no existe ningún tipo de estructura para la coordinación, seguimiento y evaluación del Plan.

También resulta interesante analizar como, en algunos de los casos de los planes que no han sido evaluados (y por lo tanto, no son objeto de nuestra meta-evaluación), sí existía una mención a la evaluación y una mínima estructura institucional para la misma. Este es el caso del II Plan asturiano (1993-1995), en el que, bajo el apartado de “coordinación y seguimiento”, además de proponer la creación de una “Comisión Interinstitucional de elaboración, coordinación y seguimiento del II Plan” (es decir, una

³² ICD. “Compareixença de la Presidenta de l’Institut Català de la Dona davant la Comissió de Seguiment del Procés d’Equiparació Home-Dona”. (Parlament de Catalunya, 6 de Maig de 1997). Documento interno. 24 pp.

³³ Aunque no sea una Comunidad Autónoma que se esté analizando en esta segunda generación de evaluaciones, conviene señalar aquí que este es también el caso de lo que ocurre en Murcia, con el Consejo Regional de Servicios Sociales.

estructura “ad-hoc”), se menciona la creación de un sistema de evaluación, la evaluación periódica de las acciones emprendidas, una evaluación del Plan a mitad de su ejecución y una evaluación final. Ninguna de estas cuestiones llega a realizarse, probablemente debido, en parte, a los cambios políticos y de estructura institucional del Principado de Asturias. En los primeros años de desarrollo de las políticas de igualdad, Asturias fue una de las comunidades autónomas más activas al respecto; sin embargo, la estructura institucional con la que contó nunca fue lo suficientemente fuerte como para subsistir a los cambios políticos. Algo muy similar ha ocurrido en la Comunidad Autónoma Valenciana, en la que los avatares políticos e institucionales, no han permitido la estabilidad suficiente para llevar una política de igualdad coherente y continuada durante unos años³⁴.

Estas estructuras son incipientes y todavía débiles en la mayoría de los casos. No se puede concluir una relación causa-efecto de un tipo de estructura con un mayor grado de evaluación. Otro problema añadido al establecimiento de conclusiones es que, además de la existencia de un tipo u otro de estructura institucional formal, hay que estudiar su funcionamiento real. Un ejemplo de esto es el estudio de Valiente sobre el IM, en el que destaca que la Comisión de Igualdad tuvo serios problemas y que el Consejo Rector apenas se reunía (Valiente, 1996). Los límites de este trabajo no permiten profundizar en cada uno de los once casos estudiados y, por lo tanto, insistimos en la necesidad de que existan estudios de caso sobre las políticas de igualdad autonómicas, como es el caso del trabajo de Granados (1999). Sin embargo, sí parece haber estructuras más adecuadas que otras para asegurar una evaluación. Sin ánimo de establecer unas conclusiones definitivas, creemos que se pueden observar las siguientes tendencias:

1) Los fundamentales garantes de los procesos de evaluación son los propios organismos de igualdad. En el caso de los organismos que son autónomos, los Consejos Rectores o de Dirección tienen un papel clave respecto a la función de evaluación. Cuanto más fuerte sea la estructura de dichos organismos, mayores posibilidades de que la estructura institucional para la evaluación sea también más fuerte.

2) Las estructuras creadas “ad-hoc” para coordinar, impulsar, seguir y evaluar los planes son ventajosas porque implican formalmente al resto de los departamentos en una tarea en la que su aportación de información va a ser clave. Pero se corre el peligro de que se queden en una mera estructura formal y son menos estables institucionalmente.

3) Las estructuras no específicas para los planes tienen, a veces, una fortaleza institucional mayor que las estructuras “ad-hoc” (son incluso, más permanentes), pero también tienen otras funciones que cumplir, lo que puede hacer que la función de evaluación de los planes se diluya. En general, pueden cumplir un papel muy

³⁴ El organismo de igualdad creado en 1988 no tenía rango de Organismo autónomo, rango que adquiere con la Ley 5/1992, de 30 de Junio, aunque es suprimido un año y medio más tarde en la Ley 6/1993, de 31 de Diciembre, de Presupuestos de la Generalidad Valenciana para el ejercicio 1994. Para más detalles, véase el capítulo 2 de este trabajo.

importante en la coordinación, seguimiento y evaluación de las políticas de igualdad en general.

4) Por ahora, la presencia de la sociedad civil y el movimiento de mujeres (a través de los Consejos de la Mujer u otras estructuras) en los procesos de seguimiento y evaluación de los Planes es meramente simbólica.

5) Los vínculos con el poder legislativo en los procesos de seguimiento y evaluación son siempre impulsores de los mismos. Si el grado de compromiso es importante y está vinculado directamente al organismo de igualdad, como es en el caso del País Vasco, se convierte en un factor crucial para impulsar, desarrollar y mantener procesos de evaluación.

Finalmente, la **estructura del personal** de los organismos de igualdad a grandes rasgos puede completar la visión de la relativa fortaleza de la estructura institucional de la evaluación, aunque el personal en concreto que se dedica a las tareas de evaluación en cada organismo es analizado más adelante en el punto referido a los agentes de las evaluaciones. Lógicamente, existe una gran diferencia entre el *staff* del IM y el de sus homólogos autonómicos y, entre éstos, también existen diferencias sustanciales. El *staff* del IM en 1993 era de 178 personas, incluyendo todo tipo de personal, organizadas en tres subdirecciones generales (Cooperación, Estudios y Documentación y Programas) y una secretaría general, además de un gabinete de relaciones internacionales y otro de medios de comunicación³⁵. Las plantillas de los organismos autonómicos van desde las 49 personas del Instituto Andaluz de la Mujer en 1993 a las 4 personas dedicadas al programa de mujer en la Dirección General de Bienestar Social de La Rioja o a las también 4 que afrontan la evaluación del I Plan canario (que luego se convertirán en 17 a partir de noviembre de 1996). Estos datos se reflejan a continuación en la tabla 5.2.

³⁵ Información extraída de IM. 1994. *10 años del Instituto de la Mujer*. Madrid: Instituto de la Mujer. pp.169-172.

**TABLA 5.2.: ESTRUCTURA INSTITUCIONAL
SEGUNDA GENERACION DE EVALUACIONES**

	Mención expresada de evaluación en el Plan	Estructuras responsables del seguimiento y la evaluación de los Planes	Plantilla en organismos de igualdad
Instituto Mujer II (93-95)	Sí	Consejo Rector	178 (en 1993)
Andalucía II (95-97)	No	Comisión del Plan para coordinar e impulsar (viceconsejeros)	49 (en 1994) *
Aragón I (94-96)	No	Consejo Rector	19 (en 1996)
Canarias I (95-96)	Sí	Consejo Rector Comisión para la Igualdad de la Mujer	4 (hasta 9/ 96) y 17 a partir 11/96)
Castilla-La Mancha II (95-99)	No	No existe ninguna estructura Dirección General de la Mujer	10 (en 1998)
Castilla y León I (94-96)	Sí	Comisión Interconsejerías para temas de igualdad de la mujer	Sin datos
Cataluña II (94-96)	No	<i>Comissió Interdepartamental per a la Igualtat d'Oportunitats per a las Dones</i> Comisión Parlamentaria de Seguimiento del Proceso de Equiparación Hombre-Mujer <i>Consell Nacional de Dones</i>	Sin datos
Galicia II (95-97)	No	Comisión de Seguimiento para la evaluación del II Plan	Sin datos
Madrid II (93-95)	No	Comisión para el impulso y seguimiento del Plan	Sin datos
La Rioja II (96-99)	Sí	Consejo de Bienestar Social Consejo Sectorial de la Mujer Comisión Interinstitucional de Bienestar Social	4 (Mujer) 45 (D.G. Bienestar Social)
País Vasco II (1995-1998)	Sí	Consejo de Dirección Comisión Interdepartamental Parlamento Vasco	22 (en 1997)

* Dato extraído de Granados, E. "El Instituto Andaluz de la Mujer: organización, presupuesto y actuaciones". Ponencia presentada en el II Congreso de la AECPA. Salamanca 1997.

Las estructuras institucionales para las evaluaciones externas

En las evaluaciones externas existen algunas estructuras institucionales que facilitan o al menos cumplen alguna función respecto a los procesos de evaluación. En el caso de Castilla- La Mancha no existe ninguna estructura institucional y en el de Galicia, la creación de una “Comisión de evaluación” interna resulta uno de los pilares para la recopilación de información necesaria para la realización del estudio. En el resto de los casos, las estructuras de coordinación existentes cumplen un papel tangencial respecto a los procesos de evaluación.

En el caso de Canarias, en la ley de creación del ICM (Ley 1/1994, de 13 de Enero) no se establece para el **Consejo Rector** ninguna función relacionada con la evaluación, ni de las políticas ni de los planes de igualdad. Sin embargo, el acuerdo de aprobación del I Plan atribuye al Consejo Rector, juntamente con la dirección del ICM, la evaluación de las medidas del Plan.

De las seis experiencias externas, se crean **estructuras ad-hoc** en tres de ellas.

En el caso de Andalucía, se crea la Comisión del Plan en el artículo dos del Acuerdo del Consejo de Gobierno que aprueba el II Plan. Su misión explícita es la de coordinar e impulsar el Plan, por lo que se supone que también tendrá algún papel respecto al seguimiento y la evaluación. A esta Comisión, además, se le pretende dar un cierto carácter político, ya que sus componentes tienen rango de viceconsejeros. El artículo 2 dice textualmente:

“Al objeto de coordinar e impulsar las medidas previstas en el Plan, se crea una Comisión que estará presidida por el Consejero de Presidencia y de la que formarán parte los Viceconsejeros de Industria, Comercio y Turismo, Trabajo y Asuntos Sociales, Salud, Educación y Ciencia, Cultura, así como la Directora del Instituto Andaluz de la Mujer”

En Galicia, la “Comisión de evaluación” es una comisión interna, es decir, no tiene una estructura avalada formalmente, por lo que se supone que su autoridad y reconocimiento formal va a depender del liderazgo que ejerza el Servicio de Igualdad (y que precisamente sufre dos cambios de directora durante el proceso evaluador). Por lo tanto, en Galicia, la única estructura realmente específica sobre la evaluación, es débil desde el punto de vista formal y surge a raíz del inicio del proceso evaluador y por necesidades de recopilación de información. Esta Comisión está compuesta por personas designadas por las distintas Consejerías y otros organismos autonómicos, y se crea con el fin de “iniciar un proceso de participación de la Xunta de Galicia en la

evaluación de los Planes (...), concienciar a la Administración de la necesidad de dicha evaluación y de coordinar la recogida de la información necesaria”³⁶.

En Madrid, y de forma parecida a lo que ocurre en Andalucía, se crea una Comisión “ad-hoc” para el Plan. Esta Comisión parece haber tenido un papel más bien simbólico en el proceso de evaluación, aunque sí consta que se tiene una reunión con dicha Comisión antes de la presentación del informe. El artículo tercero del Acuerdo del Consejo de Gobierno que aprueba el II Plan dice textualmente:

“Crear, para el impulso y seguimiento de las medidas previstas en este Plan, una Comisión en la que estarán representados, en la forma que se establezca por orden de la Consejería de Presidencia, la administración autonómica, el movimiento asociativo, los agentes sociales de Madrid y la Federación de Municipios de Madrid”

También encontramos algunas **estructuras no específicas** de evaluación, pero que cumplen un cierto papel en las mismas, tanto en Canarias como en Castilla y León.

La “Comisión para la Igualdad de la Mujer” Canaria se crea, junto con el ICM, como órgano consultivo de apoyo del mismo (Ley 1/1994, de 13 de Enero) y se desarrollan su estructura y funciones en el Decreto 1/1995, de 13 de Enero. Esta Comisión está integrada por la directora del ICM, en calidad de presidenta de la misma, cinco representantes de la Administración autonómica, tres representantes de sindicatos, tres de organizaciones empresariales y siete propuestos por las entidades que trabajen específicamente en programas a favor de la igualdad y promoción de la mujer. Entre sus funciones no se encuentran ninguna específicamente relacionada con la evaluación, aunque una de ellas es la de “Elaborar informes sobre las medidas relacionadas con la igualdad de la mujer, a iniciativa propia o a petición del Consejo Rector o del Gobierno de Canarias”, que podría interpretarse con una cierta relación. Aunque en el I Plan (cuya evaluación estamos analizando) no se le otorga formalmente ninguna función de evaluación a dicha Comisión, esto sí se hace en el II.

En Castilla y León, la “Comisión Interconsejerías para temas de igualdad de la mujer”, que se constituye por “responsables de temas de la mujer en cada Consejería (1 o 2 por Consejería en función de sus atribuciones)” (no se especifica más), tiene por funciones además de la coordinación, el intercambio de información y la movilización de programas, la “movilización de la elaboración del Informe anual sobre iniciativas para la igualdad para la Mujer en Castilla y León (informe de seguimiento de iniciativas del Plan, que deberá ser coordinado técnicamente, a nivel de gestión de datos y elaboración, por la Secretaría Regional de la Mujer”³⁷. Pero no se dice nada específicamente sobre el tema de evaluación.

³⁶ Tríptico “Evaluación del II Plan de Igualdad de Oportunidades para las Mujeres Gallegas. Información”.

³⁷ Junta de Castilla y León, “Plan Integral de Igualdad de Oportunidades para la Mujer en Castilla y León”. pp. 219-220.

Como se ha visto antes, aunque los **Consejos Regionales de la Mujer**, podrían cumplir una cierta función de seguimiento de los planes (como es en el caso de Castilla y León, en el que existe un grupo de seguimiento y control del II Plan, aunque no lo hubo durante la evaluación del I Plan que es la que se analiza en este trabajo), no parecen cumplir ningún papel importante en los procesos de evaluación externos de la segunda generación. El Consejo Regional de la Mujer de Castilla-La Mancha³⁸ no menciona entre sus funciones ningún aspecto referido al seguimiento o evaluación de los planes u otras actividades de la Dirección General de la Mujer (sólo se refiere a “participar en los programas de la DGM”). De hecho, esta estructura no parece haber tenido un papel importante en el proceso de evaluación del II Plan castellano-mancheño.

Las estructuras institucionales para las evaluaciones mixtas

En el caso del **IM**, en el II Plan, sólo se menciona que su Consejo Rector realizará las funciones de impulso y coordinación de las medidas previstas en el Plan. Es decir, no se menciona explícitamente la función de evaluación en ningún lugar y, como se ha visto antes, la responsabilidad de la evaluación recae sobre la Ministra de Asuntos Sociales que es la que deberá “elevar” el informe de evaluación al Consejo de Ministros. En el IM se creó en 1988 una “Comisión Consultiva para el seguimiento del Primer Plan de Igualdad”, que como se ha visto en la exposición de dicha primera evaluación, tuvo problemas en su funcionamiento. En la evaluación del II Plan no se crea una estructura semejante, y la responsabilidad de la evaluación recae sobre el propio Instituto y su Consejo Rector.

En el caso del **País Vasco**, sí se atribuye directamente al Consejo de Dirección de *Emakunde* las funciones de impulso, coordinación, seguimiento y evaluación. De hecho, se especifica, en el Acuerdo del Consejo de Gobierno que aprueba el II PAPME, que dicho Consejo de Dirección será el que establecerá los criterios y pautas que deberán seguir los métodos de coordinación, seguimiento y evaluación de las acciones a desarrollar por el conjunto de las Administraciones Públicas Vascas.

En el País Vasco existe, además, la “Comisión Interdepartamental para la coordinación de la ejecución del PAPME”, que es una estructura creada “ad-hoc” para el seguimiento y evaluación del Plan. Esta Comisión que tiene formalmente acotadas sus funciones debe, textualmente:

*“1) Aprobar al inicio de cada ejercicio un documento-programa en que se señalen las acciones concretas del Plan a efectuar durante el mismo, los recursos disponibles y los medios utilizables.
(...)”*

³⁸ Decreto 35/1990, de 13 de Marzo (D.O.C.M. nº 18, 21/3/1990).

2) Realizar, *al finalizar el primer trimestre*, un primer informe de seguimiento de las acciones desarrolladas y de los medios y recursos utilizados, junto con una valoración en torno a la consecución de los objetivos previstos en el Plan.

3) Aprobar, *al finalizar el ejercicio*, el informe anual de seguimiento con las mismas características que el anterior al que se sumarán tanto un informe de evaluación realizado por la Dirección del Instituto como las propuestas de adaptación de los objetivos y acciones que se estimen convenientes por parte de la Dirección del Instituto para su consideración por el Consejo de Dirección de Emakunde y, en su caso, el Consejo de Gobierno.³⁹

En **Cataluña**, coexiste la “Comisión Interdepartamental para la Igualdad de Oportunidades de las Mujeres” con el “Instituto Catalán de la Mujer”. La Comisión es un “órgano de coordinación e impulso de las acciones que promueven la no discriminación entre hombres y mujeres y la igualdad de derechos que sean competencia de los departamentos de la Generalidad” (decreto 209/1993 del 31 de agosto). Es un órgano de máxima representación gubernamental, ya que el presidente de dicha Comisión es el presidente de la Generalidad y la componen un representante de cada uno de los departamentos, nombrados por cada uno de los Consejeros. No tiene una función explícita en evaluación, pero tiene la de coordinar y fomentar las acciones, por lo que también tiene un papel, aunque éste sea simbólico respecto a dicha función.

Además, el *Consell Nacional de Dones de Catalunya* es un órgano consultivo del Instituto Catalán de la Mujer (ICD), y reúne a representantes de las entidades y colectivos que trabajan a favor de la igualdad. El Pleno del Consejo, entre otras, tiene las funciones de debatir e informar sobre el plan anual de actividades del ICD y de aprobar el informe anual sobre el estado y las perspectivas de la situación de la igualdad de oportunidades. Por lo que, lógicamente, va a estar interesado en la evaluación de los Planes de Actuación.

Respecto a los vínculos con los parlamentos, el caso del País Vasco es el vínculo más estable ya que existe a través del organismo de igualdad y no del Plan. Por la ley de *Emakunde* el organismo es el que está obligado y toma la iniciativa de rendir cuentas al Parlamento. En los otros casos, la iniciativa parte de las diferentes comisiones legislativas relacionadas con los temas de igualdad o de mujer; por lo que la actividad parlamentaria al respecto depende de lo activo o fuerte de dichas comisiones y siempre es a petición de las mismas, y no del organismo ejecutivo.

En general, las estructuras institucionales para la evaluación en las experiencias mixtas son fuertes en el caso de los organismos de igualdad autonómicos, destacando el caso del País Vasco, y no tanto en el caso del IM en el nivel central.

³⁹ “Propuesta de Acuerdo por el que se aprueba el II Plan de Acción Positiva para las Mujeres en la Comunidad Autónoma de Euskadi”, cuyo acuerdo adoptó el Consejo de Gobierno, en su Sesión celebrada el 29 de Diciembre de 1995.

Las estructuras institucionales para las evaluaciones internas

En el caso aragonés, no existe ninguna estructura “ad-hoc” para la evaluación y las funciones de seguimiento y evaluación de los planes son directamente encomendadas al Consejo Rector del Instituto Aragonés de la Mujer.

En La Rioja, existe un Consejo Sectorial de la Mujer, que depende del Consejo de Bienestar Social y que tiene por objeto, entre otras funciones, el seguimiento del Plan. Como ya se ha señalado, la política de igualdad de género en La Rioja está limitada al marco de las políticas de bienestar social, siendo el tema de mujer uno entre otros. Por lo tanto, las estructuras institucionales existentes son muy diferentes a las del resto de las comunidades autónomas.

2.4. Recursos empleados en las evaluaciones

Panorama general

Finalmente, otro factor importante para valorar el planteamiento que se hace de una evaluación, es sin duda, los recursos que en ella se invierten, ya sean económicos o humanos. En el caso de las evaluaciones externas, resulta fácil la ponderación de los gastos directos, ya que suelen quedar consignados como tal. Más difícil es la medición de los gastos indirectos que provocan las situaciones de evaluación, no sólo en los procesos de evaluación interna sino también de evaluación externa. En las evaluaciones internas prácticamente todos los gastos que se generan son de carácter indirecto, pero también en los procesos de evaluación externa es interesante medir dichos gastos, es decir, el esfuerzo invertido por el correspondiente organismo de igualdad en la coordinación con el equipo externo, más el tiempo interno empleado en aportar y recoger información, así como en recibir, revisar, ponderar, analizar y procesar los resultados de una evaluación.

La medición de estos gastos indirectos es muy difícil hacerla de manera fiable si no se han establecido previamente unos criterios y un sistema para la recogida de información (tiempo empleado por cada trabajador/a en el proceso de evaluación) que permita consignar los datos de forma periódica y de una forma frecuente (como máximo semanalmente), ya que la estimación del tiempo dedicado a una tarea ya realizada es muy subjetiva y depende de muchos factores. Si existe personal técnico dedicado a tiempo completo y exclusivo a la evaluación es más fácil, pero éste no es el caso en los organismos de igualdad estudiados: prácticamente todo el personal técnico con responsabilidades en las tareas de evaluación se dedican, además, a otras tareas distintas de la evaluación. En el inicio de este trabajo se intentó realizar una estimación de estos gastos indirectos en cada una de las once experiencias metaevaluadas, pero ésta resultó prácticamente imposible. Dadas las limitaciones de este trabajo, por lo tanto, se hace sólo una aproximación a los gastos indirectos de los procesos de evaluación.

Con el fin de realizar una ponderación de los recursos empleados en términos relativos, se trató de obtener información no sólo de los recursos empleados en los procesos de evaluación, sino también de los recursos empleados en el plan evaluado así como los presupuestos del organismo de igualdad durante el período de vigencia del plan correspondiente. Esto ofrecería una idea relativa del esfuerzo invertido en los procesos de evaluación de los planes de igualdad.

Los criterios que se quisieron utilizar fueron los siguientes:

Gastos directos: se contemplan sumadas las cantidades totales sumando los diferentes contratos en casos de que haya más de uno (por ejemplo, en IM, se suman los 7 millones del estudio realizado en 1993 con los 7,5 del estudio cuantitativo más los 3,5 del estudio cualitativo de 1996).

Gastos indirectos: En este caso nos encontramos no sólo con una falta real de datos (en muchos casos los responsables técnicos no eran capaces de hacer una estimación del tiempo dedicado a la evaluación) sino de un problema de fiabilidad de los mismos. Además de la subjetividad y relatividad de las estimaciones cuando éstas se realizaban, para hacer un cálculo válido de los gastos indirectos reales de los procesos de evaluación, habría también que incluir el porcentaje de los gastos generales de cada organismo de igualdad correspondiente a la función de evaluación, así como también del esfuerzo empleado en cada departamento gubernamental (distinto al organismo de igualdad) en el proceso de evaluación. Estas dificultades nos hicieron desistir de dar cifras aproximadas, aunque sí hemos querido dar una idea general sobre si el esfuerzo empleado ha sido mucho o poco.

Presupuesto del plan: El criterio tomado ha sido el presupuesto calculado sobre el monto invertido en el Plan no sólo por el organismo de igualdad correspondiente, sino por cada uno de los departamentos e instituciones gubernamentales que han participado en él. Lógicamente, en la ejecución de un plan de igualdad se invierte más dinero de otros departamentos gubernamentales distintos al organismo de igualdad. La cantidad refleja la suma de los presupuestos anuales de los años correspondientes al período de vigencia del plan evaluado, cuando éstos han sido calculados. El problema es que este monto total no ha sido calculado en la mayoría de los casos de esta segunda generación de evaluaciones.

Presupuesto de los organismos de igualdad durante los años de vigencia del plan evaluado. Esta cifra es más fácil de obtener y es más fiable que la del presupuesto de los planes. Con este dato se puede calcular el porcentaje de lo que ha supuesto la evaluación (al menos en sus gastos directos) respecto al organismo de igualdad. Esta información es importante ya que son los organismos de igualdad los que tienen la responsabilidad de la evaluación de los planes. El gasto relativo que se haga en evaluación en cada organismo da una idea aproximada del peso que se le da a dicha función. Para unificar criterios, se proponía comparar el gasto de la evaluación con los presupuestos correspondientes a los años de vigencia del Plan, aunque en la mayoría de los casos, la evaluación se hace en años posteriores a dicha vigencia.

La obtención de estos datos nos ha resultado muy difícil y, en la mayoría de los casos, no se han podido obtener de forma completa. Los gastos indirectos fueron, como ya se ha comentado, imposibles de calcular. Los presupuestos de los planes no llegaron a estar nunca calculados en la mayoría de los planes de segunda generación evaluados. Por último, los presupuestos de los organismos de igualdad han sido más fáciles si éstos eran autónomos, pero no así cuando se trataba de una dirección general o similar en la que su presupuesto no aparece en una partida única. Por lo tanto, se renuncia a establecer los cálculos pretendidos y sólo se pretenden realizar estimaciones globales en cada uno de los apartados siguientes referidos al conjunto de las once experiencias.

En general, los recursos invertidos en los procesos de evaluación de segunda generación han sido relativamente muy bajos, con la excepción en cierto modo del País Vasco en donde la función de evaluación en *Emakunde* está bastante desarrollada y en donde se han invertido gran cantidad de esfuerzos internos en ella. En términos absolutos, los gastos relacionados con estudios de evaluación son más bien bajos (mirar tabla en el apartado de evaluaciones externas), pero incluso cuando en algún caso el gasto directo absoluto parece alto, como en el caso de Andalucía, éste resulta bajo en

términos relativos (0,97% respecto al presupuesto del IAM y 0,07% respecto al presupuesto del Plan).

Los recursos para las evaluaciones externas

Los seis estudios de evaluación externa analizados tienen dimensiones muy distintas, como se ve en la tabla 5.3, en la que se reflejan los costes directos de dichos estudios (es decir, lo que los correspondientes organismos de igualdad han pagado por ellos a los equipos externos que los han realizado). En esta misma tabla se refleja también lo que supone dicho coste en términos relativos: por un lado, el % respecto al presupuesto total del Plan evaluado y, por otro, respecto a lo que han sido los presupuestos del organismo de igualdad que ha encargado dicho estudio durante los años de vigencia del Plan.

**TABLA 5.3 COSTES DIRECTOS⁴⁰
DE LOS ESTUDIOS DE EVALUACION EXTERNA
SEGUNDA GENERACION DE EVALUACIONES**

ANDALUCIA II (95-97)	55.980.000
CANARIAS I (95-96)	Menos de 2.000.000
CASTILLA-LA MANCHA⁴¹ II (95-99)	5.500.000
CASTILLA Y LEON I (94-96)	1.950.000
GALICIA II (95-97)	3.000.000
MADRID II (93-95)	19.500.000

⁴⁰ Ni en el IM ni en Madrid se incluyen en los gastos directos la publicación de los resultados de la evaluación.

⁴¹ Castilla- La Mancha: Este es un cálculo aproximado, hecho por parte de las responsables de la evaluación, ya que se firma un mismo convenio con la Fundación Mujeres por la evaluación del II Plan y la asesoría y proceso de elaboración del II Plan, por un total de 15.000.000 ptas.

Como se puede observar, en términos absolutos hay mucha diferencia entre los estudios analizados. Por un lado, estarían los dos estudios más pequeños, los de Canarias, Castilla y León, tanto en dimensiones como en coste. En segundo lugar estarían los estudios de Castilla-La Mancha y Galicia, con dimensiones parecidas, aunque los costes sean algo diferentes. Después estaría el estudio madrileño que cuesta un total de 19.500.000 ptas. y, por último, diferenciado tanto en dimensiones como en coste, se encontraría el estudio de evaluación del II Plan andaluz.

	Costes directos de la evaluación	Presupuesto del Plan (miles de ptas.)	% gasto Evaluación / Plan	Presupuesto del organismo de igualdad (miles de ptas.)	% gasto Evaluación / organismo
ANDALUCIA II (95-97)	55.980.000	80.401.000	0,07	5.765.006	0,97

Aunque en términos absolutos algunos estudios de evaluación puedan parecer de alto coste, resulta interesante contemplar dichas cifras en términos relativos. Seleccionando el estudio más caro (con grandes diferencias sobre el resto), es decir el andaluz, se puede apreciar que, en términos relativos, el coste de la evaluación no es tan alto: representa menos del 1% del gasto del IAM durante los tres años de duración del Plan y sólo el 0,07% respecto al presupuesto del Plan, cantidad que parece desde luego muy pequeña para su evaluación.

Además, en los cálculos hechos habría que tener en cuenta los presupuestos de los años en los que se realiza la evaluación que no siempre coinciden con los de vigencia del Plan. Por ejemplo, en el caso de Andalucía, el gasto de la evaluación se contrae para los años 1997, 1998 y 1999 (y sin embargo nosotros comparamos con los presupuestos sumados de 1995, 1996 y 1997, años de vigencia del Plan); esto quiere decir que en realidad, el porcentaje de gasto dedicado a la evaluación sería aún menor (0,81% en lugar de 0,97%), ya que la suma de los presupuestos 97-98-99 del IAM (6.907.727.000 ptas.)⁴² es bastante mayor que la de 95-96-97 (5.792.819.000 ptas.).

También habría que tener en cuenta las prórrogas que en muchos casos han tenido los planes evaluados. En el mismo caso de Andalucía, el Plan se extiende dos años más, por lo que el porcentaje de lo que ha supuesto la evaluación externa descendería si añadiéramos el presupuesto autonómico dedicado al Plan de Igualdad en los años 1998 y 1999. Este caso también ocurre en Madrid, ya que el segundo plan se extiende durante más de un año antes de la aprobación del tercero.

⁴² No disponemos de los presupuestos de 1999, por lo que la suma la hacemos equiparándolos a los de 1998. Los datos los extraemos de Granados (1999).

Los recursos empleados en las evaluaciones mixtas

Al ser procesos mixtos de evaluación, no sólo se trata de calcular los costes directos de las colaboraciones externas sino de “traducir” en costes indirectos el esfuerzo y los recursos invertidos en los procesos internos. Como ya se ha comentado, el cálculo de estos costes indirectos resulta imposible hacerlos de forma más o menos fiable, por lo que se harán apreciaciones de forma global.

En el caso del **IM**, en el que la colaboración externa en el estudio de evaluación asciende a 18 millones, habría que tener en cuenta la dedicación de la Subdirección de Estudios al proceso de evaluación. Una parte del estudio, la llamada evaluación de la implementación, se realizó plenamente de forma interna y dicha subdirección contaba con un equipo, consistente en una becaria durante los años de 1993 y 1994 (que en estimación del Consejero Técnico de Estudios le dedicó un 80% a la evaluación) y dos becarias durante los años 1995 y 1996 (con una estimación de una dedicación del 30% de su tiempo laboral cada una)⁴³. A eso habría que añadirle la dedicación del propio consejero técnico que comparte otras funciones, además de la evaluación (su estimación de dedicación fue más o menos de una semana al diseño del estudio, tres a los indicadores más el tiempo dedicado a la redacción de los informes). Si tenemos en cuenta que la plantilla del IM era de 178 personas en 1993, se puede decir que el esfuerzo interno relativo dedicado al trabajo de evaluación del II Plan fue bastante bajo.

Aunque sobre la evaluación del II plan realizada en el **Instituto Catalán de la Mujer** hemos obtenido pocos datos concretos al respecto, parece que la evaluación del II Plan supuso bastante esfuerzo interno por parte del equipo técnico del Instituto, a pesar de la colaboración externa que tuvo ESADE -equipo externo- en la misma. Sin embargo, esto es una apreciación extraída de las entrevistas que realizamos a la directora y a la responsable técnica de la evaluación, ya que no pudimos tener datos referidos a la dimensión y coste de la ayuda externa brindada.

En cualquier caso, la diferencia es grande con los recursos internos empleados en **Instituto Vasco de la Mujer**, en el que, a pesar de contar con una plantilla estable de no más de 25 personas, la dedicación a los procesos de evaluación ha sido grande y continua durante los cuatro años de vigencia del Plan. La colaboración externa en el País Vasco -realizada por el Departamento de Psicología Social y Metodología de las Ciencias del Comportamiento de la UPV- se estableció en un convenio anual de unas 2.800.000 ptas. cada año aproximadamente, lo que asciende a un monto de 11.200.000 ptas. en total en lo que a costes directos de la colaboración externa se refiere. Sin

⁴³ Hay que tener en cuenta que ese mismo equipo es el que elabora la publicación periódica de “Mujeres en cifras”, que requiere bastantes energías y recursos y que no puede ser considerado como trabajo de evaluación de los planes.

embargo, en estimación de la técnica dedicada al tema de evaluación en *Emakunde* esta cantidad es muy baja comparada con los esfuerzos internos dedicados a ello.

Los recursos empleados en las evaluaciones internas

Finalmente, en el caso de los procesos de evaluación interna, resulta especialmente difícil medir los costes de los mismos. Sin embargo, y haciendo una apreciación general, para la evaluación del I Plan aragonés, dadas sus dimensiones tan pequeñas y la premura de tiempo con la que se realizó, no se emplearon demasiados esfuerzos ni se comprometieron muchos recursos.

Para el II Plan riojano, sin embargo, el esfuerzo ha sido más continuado aunque tampoco fue especialmente grande ya que las funciones de evaluación se desarrollan en el seno de una Dirección General no específica en la que, de 45 personas, sólo 4 están específicamente asignadas al área de mujer. Las tareas de evaluación recaen, sobre todo, en la socióloga que cumple sus funciones de técnica para toda la Dirección General y no sólo para el área de mujer.

CAPÍTULO 6

LOS ELEMENTOS CLAVE DE LA EVALUACIÓN: AGENTES, MOMENTOS, CRITERIOS Y METODOLOGÍA

INTRODUCCIÓN

En el capítulo anterior se tratan cuestiones referidas al diseño y a la conceptualización de los procesos de evaluación, abordando el **planteamiento** de dichos procesos y las estrategias evaluativas seguidas para la evaluación de los planes. En este capítulo se discuten cuestiones relacionadas con el proceso de puesta en práctica de la evaluación, analizando el manejo que se realiza en las experiencias metaevaluadas de algunos **elementos clave** de los procesos evaluativos, como son los agentes que participan en la evaluación, los momentos y el calendario de la misma, así como los criterios utilizados y la metodología y los procedimientos seguidos. También se hace referencia a si estos elementos clave estaban contemplados previamente en el diseño de los procesos de evaluación.

El primer elemento clave analizado en este capítulo es el que hace referencia a los **agentes** involucrados en los procesos de evaluación analizados: quiénes participan en el diseño y ejecución de la evaluación y cuál es el grado de participación de los actores críticos de las políticas de igualdad en los procesos de evaluación. La primera cuestión tenida en cuenta fue la referida a los equipos técnicos que diseñaron y dirigieron los procesos de evaluación. En el caso de las evaluaciones internas y las mixtas, se analizaron los recursos humanos de los organismos de igualdad que se dedicaron y qué medida a los procesos de evaluación y se tuvo en cuenta, en las evaluaciones mixtas, el tipo y la magnitud de la colaboración externa en cada caso. Para las evaluaciones externas, se describen el tipo y las características generales de los equipos externos que realizaron las evaluaciones, así como los procedimientos de colaboración y coordinación con los organismos de igualdad. Además de a los equipos técnicos, este apartado hace referencia a otros posibles actores que hubieran podido participar activamente en algunas de las fases de evaluación, haciendo especial referencia a la sociedad civil y al movimiento de mujeres, así como a técnicos y profesionales de otras unidades administrativas implicadas en la elaboración y desarrollo de los planes evaluados.

Seguidamente se trata el tema de los **momentos** y los tiempos en los que se realiza la evaluación; el calendario, la temporalidad y la oportunidad en el tiempo son cruciales en las evaluaciones y, por lo tanto, el análisis de este factor resulta siempre especialmente esclarecedor. Se identificaron y analizaron cuatro elementos temporales diferenciados: en primer lugar, en qué momento del período de vigencia del plan a evaluar se planteó y comenzó el proceso de evaluación; en segundo lugar, la duración de los procesos de evaluación desde su inicio hasta la entrega del informe final; en

tercer lugar, el tiempo transcurrido entre la finalización del período de vigencia del plan y la entrega del informe final y, por último, los retrasos que se produjeron sobre el calendario inicial de los estudios de evaluación.

En tercer lugar, se analiza el contenido de las evaluaciones, es decir, qué se evalúa, o para decirlo más propiamente, cuáles han sido los **criterios e indicadores** empleados en cada evaluación. Los criterios e indicadores de una evaluación representan dos cuestiones fundamentales: por un lado, las variables estudiadas y, por otro, la base lógica en la que se fundamentan los juicios de valor que se establecen posteriormente en el proceso de evaluación. De ahí la necesidad de que se hagan convenientemente explícitos. Además del análisis global realizado de, en general, la poca importancia y “espacio” concedidos a los criterios de evaluación, en cada caso se analiza la claridad con la que se explicitan en los informes finales, además de relacionar dichos criterios, una vez identificados, con la tipología de evaluación según el contenido, es decir, saber si son criterios o indicadores que hagan referencia al diseño, a la ejecución o a los resultados (tanto de productos o *outputs* como a efectos o *outcomes*) del plan evaluado.

Por último, se analizan los **enfoques metodológicos** practicados en cada una de las once experiencias de evaluación, así como las **técnicas** de recopilación y tratamiento de la información utilizadas. Al igual que ocurre con los criterios, tampoco se encontraron en los informes, explicaciones profusas sobre la elección del tipo de técnicas o procedimientos metodológicos empleados en las evaluaciones, aunque en algunos casos se dan especificaciones técnicas de algún instrumento concreto. En este apartado se describen los instrumentos o técnicas de recopilación y análisis de información que se emplearon en cada evaluación, destacando los registros o fichas para recoger información producida por la propia gestión del plan. En menor medida se utilizaron cuestionarios, entrevistas y grupos de discusión, así como la realización de análisis de fuentes secundarias en algún caso.

Para la exposición de la información en este capítulo se utiliza, al igual que en el anterior, la clasificación o tipología de evaluaciones según el agente que evalúa, tratando de forma separada las evaluaciones externas (seis experiencias), las evaluaciones mixtas (tres experiencias), y las evaluaciones internas (dos experiencias).

1. AGENTES Y PARTICIPANTES EN LOS PROCESOS DE EVALUACIÓN

En este apartado, se analizan, por un lado, los equipos técnicos que diseñan y ejecutan las evaluaciones y, por otro, quiénes han participado en los procesos de evaluación.

Respecto a la primera cuestión, los **equipos técnicos** que dirigen las evaluaciones, se tratarán separadamente las evaluaciones externas y las evaluaciones internas, diferenciando en estas las evaluaciones mixtas, las cuales han contado con la colaboración de un equipo externo y las evaluaciones internas. Lógicamente, éste es precisamente el aspecto que más varía de un tipo de evaluación a otra.

La segunda cuestión, sin embargo, referida a los **participantes en los procesos de evaluación** puede ser analizada de una forma más general, aunque también describiremos los diferentes casos según la clasificación utilizada según el agente evaluador. En este punto queremos señalar que utilizar un determinado colectivo como fuente para obtener la información necesaria para realizar la evaluación no significa necesariamente que dicho colectivo haya participado en el proceso de evaluación. Con esto queremos decir que no se puede afirmar la participación en una evaluación de un colectivo al que, por ejemplo, se le ha preguntado su opinión sobre algo en un cuestionario. Por participación entendemos la implicación activa en alguna o varias partes del proceso de evaluación: diseño del estudio de evaluación, definición o selección de los criterios de evaluación, selección y construcción de los instrumentos de recopilación de información, recopilación de la información, análisis y valoración de la misma, establecimiento de recomendaciones, redacción del informe, difusión de los resultados.

Respecto a **quiénes han participado activamente en alguna de las fases de la evaluación**, en general, se podría decir que los procesos de evaluación analizados han tenido un carácter bastante “técnico”, dirigiendo los procesos de evaluación los equipos técnicos, bien sean éstos internos o externos, desde los organismos de igualdad. En un segundo plano, y a mucha distancia respecto al grado de implicación, habría que situar a técnicos de otros departamentos gubernamentales que han participado, en mayor o menor medida en el diseño y ejecución de la recopilación de información en sus respectivos departamentos. Cuando existe una estructura institucional “ad-hoc” para el seguimiento del Plan, parece más evidente la participación interdepartamental en las tareas de evaluación, especialmente si ha sido una estructura creada específicamente para el proceso de evaluación. Si no existen este tipo de estructuras, la participación queda bastante más desdibujada.

Debido al carácter “técnico” de las evaluaciones, el gran ausente de los procesos de evaluación es el movimiento de mujeres. Su participación, en el que caso que se haya dado, ha sido meramente simbólica, haciendo una presentación de los resultados de la evaluación en la estructuras correspondientes. Las estructuras con representación de la

sociedad civil (como los Consejos de la Mujer) no parecen haber tenido una importancia crucial en las evaluaciones que se han metaevaluado.

En cualquier caso, y todavía en esta segunda generación de evaluaciones, resulta difícil pensar en una participación amplia de personas, grupos y colectivos que no hubieran participado previamente en su planificación y formulación.

1.1. Agentes y participantes en los procesos de evaluación externa

En el caso de las evaluaciones externas, se trata de **equipos técnicos externos** los que han diseñado y ejecutado las evaluaciones. Por lo tanto, el peso de la responsabilidad técnica de las mismas recae sobre dichos equipos externos. En la tabla 6.1. a. se señalan los equipos externos y el procedimiento de colaboración utilizado en cada caso.

La mayoría de los equipos externos (no es el caso del IM, de Canarias y de Castilla- La Mancha), tienen origen o procedencia universitaria. En las evaluaciones de Castilla y León y de Madrid, que son realizadas por el mismo equipo, aunque es la empresa EURODOXA S.A., se insiste en que es el equipo del Prof. Martín Serrano (incluso en la presentación de la publicación sobre los resultados de la evaluación del II Plan madrileño), es decir, se apela a su condición de profesor universitario. Así, en la búsqueda y selección de los equipos parece haberse tenido en cuenta, como es lógico en cualquier proceso de evaluación, el factor de credibilidad de dichos equipos.

Respecto a la especialidad principal de los equipos externos, salvo en el caso de Castilla- La Mancha, en dónde se han seleccionado especialistas en igualdad y género, en el resto de las evaluaciones se han seleccionado especialistas más bien en metodología de evaluación y/o investigación. Precisamente, en las entrevistas realizadas, alguna responsable se queja de la dificultad de encontrar equipos especialistas en ambas dimensiones: por un lado, en evaluación de programas y, por otro, en cuestiones de género. Además de esta dificultad para encontrar equipos especializados en esta doble dimensión, creemos que existe un problema previo que es la especialización en evaluación de programas: al no existir una formación específica y especializada en este área en España, los equipos expertos tienden a estar más especializados en metodología de investigación social más que en evaluación de programas y políticas públicas.

Los procedimientos para la colaboración externa son variados: concurso público, convenios de colaboración, o adjudicación directa si la cantidad percibida por contrato por parte del equipo externo es menor a 2.000.000 ptas. Los convenios se establecen con entidades no empresariales, bien sean instituciones universitarias o Fundaciones. Los convenios permiten, además, un compromiso plurianual, cuestión más difícil con los concursos públicos. En cualquier caso, parece conveniente que el compromiso contractual tenga la posibilidad de una cierta perspectiva en el tiempo. Como se verá en

el siguiente epígrafe, uno de los grandes problemas de esta segunda generación de evaluaciones es que, en la mayoría de los casos, se ha decidido muy tarde hacer la evaluación. Una evaluación debe estar prevista desde el principio y esto probablemente implique, máxime si se decide realizar una evaluación externa, la presencia y participación de los equipos externos desde el inicio y hasta el final. Los procedimientos de contratación disponibles en las administraciones públicas, sobre todo si se trata de empresas con ánimo de lucro, están limitados a este respecto o, al menos, dejan mucho más en el aire la estabilidad de un mismo equipo durante todo el tiempo que puede durar un proceso de evaluación de un plan de igualdad.

**TABLA 6.1. a: EQUIPOS EXTERNOS DE EVALUACION
EVALUACIONES EXTERNAS
SEGUNDA GENERACION DE EVALUACIONES**

	Equipos externos responsables de la evaluación o que participan en ella	Procedimiento para la colaboración
Andalucía II (95-97)	Instituto de Estudios Sociales Avanzados (IESA) de Andalucía	Convenio
Canarias I (95-96)	Asesores en Imagen Corporativa S.A.	Adjudicación directa
Castilla-La Mancha II (95-99)	Fundación Mujeres (con la colaboración de WG, Consultoras para el Desarrollo)	Convenio
Castilla y León I (94-96)	EURODOXA S.A. (Prof. Martín Serrano)	Adjudicación directa
Galicia II (95-97)	Universidad de la Coruña (Prof. Amparo Almarcha)	Convenio
Madrid II (93-95)	EURODOXA S.A. (Prof. Martín Serrano)	Concurso público

Los equipos internos cumplen también un papel muy importante, no sólo en las evaluaciones internas, sino también en la necesaria coordinación del organismo de igualdad con los equipos externos en el caso de las evaluaciones externas. En la mayoría

de los casos, más que tratarse de un equipo interno, suele tratarse de una persona con la responsabilidad técnica de la evaluación, cuyo papel suele ser muy importante. Sin embargo, y en general, en las evaluaciones externas de esta segunda generación la implicación de los equipos internos en los procesos de evaluación no ha sido muy grande: precisamente, la contratación de un equipo externo “libera” a unos equipos internos que, además de la evaluación, deben cumplir muchas otras funciones (normalmente asociadas a fomento de la investigación, recopilación de datos sobre igualdad, documentación y publicaciones, etc.). Aún teniendo en cuenta esta tendencia general, existen diferencias entre las seis experiencias externas: en algunos procesos de evaluación se “ha dejado hacer” en mayor medida al equipo externo, como es el caso de Canarias, Castilla-La Mancha y Castilla y León. En el resto, se ha llevado una coordinación algo más estrecha y continuada equipo interno-externo durante el proceso, como es el caso de Andalucía, Madrid, y de forma especial Galicia, aunque en algunos de los casos, como es Andalucía, esta coordinación ha fallado a lo largo del proceso debido en parte, a los cambios producidos en el propio equipo externo.

Respecto a **quiénes han participado activamente en alguna de las fases de la evaluación**, ya se ha comentado que, en general, los procesos de evaluación tienen un carácter esencialmente “técnico”. Esto es especialmente evidente en las evaluaciones externas, en las que los verdaderos protagonistas de los procesos son los equipos externos contratados.

La participación en los procesos de evaluación de técnicos de otros departamentos o unidades gubernamentales se produce de forma algo más evidente cuando existe una estructura institucional “ad-hoc” para el seguimiento y/o evaluación del Plan, especialmente si ha sido una estructura creada específicamente para el proceso de evaluación, como es en el caso de Galicia. Sin embargo, y especialmente en el caso de las evaluaciones externas, el grado de implicación de otras unidades gubernamentales en la evaluación ha sido generalmente muy bajo.

Respecto a la participación de la sociedad civil (movimiento de mujeres), y como ya se ha visto en el capítulo anterior, las estructuras formales de participación suelen estar separadas, por lo que la ausencia de participación de las ONG en los procesos de evaluación resulta más patente; es decir, por un lado hay una comisión de representantes gubernamentales y por otro, de colectivos y asociaciones de mujeres, a excepción del caso de la Comisión de la Igualdad de Canarias y la Comisión para el impulso y seguimiento del Plan de Madrid. Además, en los dos casos mencionados, estas comisiones no parecen haber jugado un papel importante en las evaluaciones metaevaluadas.

1.2. Agentes y participantes en los procesos de evaluación mixta

En las evaluaciones mixtas, también se cuenta con **equipos técnicos externos** que colaboran en algún momento del proceso de evaluación, ya sea en su diseño (Cataluña), su ejecución (IM) o cuestiones concretas de ambos procesos (País Vasco). A diferencia de las evaluaciones externas, el peso de la responsabilidad técnica de las evaluaciones recae sobre los equipos internos de los organismos de igualdad aunque también colaboren equipos externos. En la tabla 6.1. b. se señalan los equipos externos que han participado en estos tres procesos mixtos de evaluación y el procedimiento de colaboración utilizado en cada caso.

En el caso del IM, el diseño del estudio de evaluación corresponde al Consejero técnico de Estudios, que cuenta con el apoyo del departamento de estudios y la ayuda de becarias de dicho departamento. La intervención del equipo externo se limita a la ejecución de parte del estudio, haciendo el trabajo de campo y el análisis de las llamadas evaluaciones cuantitativas y cualitativas. En Cataluña se recurre también al equipo externo de ESADE para el mismo diseño del proceso de evaluación y en el País Vasco, el equipo externo colabora a lo largo de todo el proceso aunque el grueso del diseño de la evaluación ya estaba hecho por *Emakunde* para la evaluación del I PAPME; sin embargo, se cuenta con el equipo externo para algunos cambios en el diseño (por ejemplo, redefinición de las categorías en las que se clasifican las actuaciones del Plan).

**TABLA 6.1. b: EQUIPOS EXTERNOS DE EVALUACION
EVALUACIONES MIXTAS
SEGUNDA GENERACION DE EVALUACIONES**

	Equipos externos responsables de la evaluación o que participan en ella	Procedimiento para la colaboración
Instituto Mujer II (93-95)	CUANTER, S.A.	Concurso público
Cataluña II (94-96)	ESADE	Sin datos
País Vasco II (96-2000)	Dpto. de Psicología Social y Metodología de las Ciencias del Comportamiento. Universidad del País Vasco (Profesoras Maite Erro y Maribel Pizarro)	Convenio

En dos de los casos de evaluación interna mixta, los equipos externos tienen origen universitario. Este no es el caso de la evaluación del II Plan nacional en el que la

colaboración externa es más “técnica” limitándose al trabajo de campo y el análisis de un diseño ya realizado de forma interna.

Los equipos internos cumplen en este tipo de evaluaciones un papel especialmente importante. Como ya se ha dicho, en el IM el responsable de la evaluación ha sido el Consejero Técnico de Estudios, quien asesora todos los estudios que se realicen en el IM y no sólo los de evaluación. En los casos de Cataluña y el País Vasco, además de una responsable, ha habido sendos departamentos con sus correspondientes equipos dedicados a la evaluación, con mayor dedicación y personal en el País Vasco que en Cataluña.

Respecto a **quiénes han participado activamente en alguna de las fases de la evaluación**, al igual que en las evaluaciones externas, el proceso de evaluación llevado a cabo por el IM ha tenido un carácter muy “técnico” y, por lo tanto, la participación de otros agentes no técnicos ha sido prácticamente nula. Ni tan siquiera ha sido importante la participación de otros agentes gubernamentales ya que ésta se concebía siempre a petición del propio IM y no había ninguna estructura de coordinación creada para ello.

Algo distinto ha sido en el caso de Cataluña y, sobre todo, en el del País Vasco, en donde la evaluación de los planes se ha concebido como algo no puramente técnico ni tangencial, sino como una función crucial en *Emakunde*. En el caso del País Vasco, además, se le da especial importancia a la participación de los diferentes departamentos de las administraciones públicas vascas y a la Comisión Interdepartamental como vehículo de dicha participación.

Las evaluaciones mixtas tampoco son una excepción en lo que se refiere a la participación, en los procesos de evaluación, del movimiento de mujeres.

1.3. Agentes y participantes en los procesos de evaluación interna

En el caso de las evaluaciones internas, no existe ninguna participación de equipos técnicos externos. Los equipos internos que desarrollan las dos evaluaciones son, en cualquier caso, pequeños (compuestos por una o dos personas).

En el caso de las evaluaciones internas, las responsables directas en los organismos correspondientes sí cuentan algo más con la participación de otros agentes gubernamentales para proporcionar información, aunque una implicación algo más activa de las otras unidades gubernamentales sigue siendo el gran “caballo de batalla” también en las evaluaciones internas.

Tampoco las evaluaciones internas cuentan con la participación, en los procesos de evaluación, del movimiento de mujeres.

2. MOMENTOS DE LOS PROCESOS DE EVALUACIÓN

El segundo elemento clave de las evaluaciones viene determinado por los momentos de dichas evaluaciones. Efectivamente, los aspectos temporales en los procesos de evaluación resultan cruciales dado su carácter tan aplicado y claro apego al contexto concreto en el que se realiza la evaluación.

En este punto se identifican cuatro elementos temporales distintos y que son los siguientes. En primer lugar, es importante señalar en qué **momento** del período de vigencia del plan evaluado se plantea y comienza la evaluación del mismo. En segundo lugar, se analiza la **duración** de los procesos de evaluación desde su inicio hasta la entrega del informe final. En tercer lugar, se consigna también cuánto **tiempo pasa entre la finalización del período de vigencia del plan y la entrega del informe final** y, por último, se analizan los **retrasos** sobre el calendario inicial de los procesos de evaluación. El criterio seguido para el momento en el que se plantea la evaluación ha sido consignar el momento en el que existe un proyecto escrito. Este suele ser distinto del momento de comienzo, en el cual se consigna la fecha en la que falla el concurso, se firma el convenio o se comienza de hecho a establecer criterios de evaluación, diseñar instrumentos de recopilación de información, etc. El momento de finalización viene determinado por la entrega del informe final.

Todos estos elementos se reflejan en las tablas 6.2, que se dividen en 6.2.a para las evaluaciones externas, 6.2.b para las mixtas y 6.2.c para las internas.

2.1. Momentos de los procesos de evaluación externa

Como se puede observar en la **tabla 6.2. a**, en todas las evaluaciones externas, el **planteamiento** de las mismas se realiza alrededor del último año de vigencia del plan que se va a evaluar: varía entre el año y medio antes de Galicia a los cuatro meses antes de Castilla y León. Es decir, lo antes que se plantea una evaluación es a mitad del período de vigencia del plan, como en Galicia. Esto es una circunstancia que va a condicionar mucho los procesos de evaluación y que se reconoce como una fuente importante de limitaciones desde el inicio de los procesos (el reconocimiento de dichas limitaciones aparece ya en algunos de los proyectos de evaluación -IESA para Andalucía, Fundación Mujeres para Castilla-La Mancha). Además, esta circunstancia es también un indicador de que los planes no fueron elaborados, en general, en términos de resultados esperados, por lo que resulta más probable el hecho de que existan, en ellos, objetivos, medidas y actuaciones que estén formulados de forma vaga o general. Pensar en la evaluación al final de los planes (o al menos, cuando éstos llevan ya una larga

vida) acarrea muchos problemas y hace que las evaluaciones pierdan una parte importante de su potencial como instrumento de mejora y rendimiento de cuentas.

TABLA 6.2. a: TIEMPOS Y MOMENTOS DE LAS EVALUACIONES EXTERNAS SEGUNDA GENERACION DE EVALUACIONES

	Momento en el que se plantea la evaluación	Duración		Tiempo entre final plan e informe final	Retrasos
		Principio	Final		
Andalucía II (95-97)	Nov. 1996 (a 13 meses)*	Marzo 97	Julio 99 (28 meses)**	1 año, 6 meses	Sí, el informe era para 3/99
Canarias I (95-96)	Julio 1996 (a 6 meses)	Julio 96	Oct. 96 (3 meses)	Se presenta antes	No
Castilla-La Mancha II (95-99)	Octubre 1998 (a 15 meses)	Oct. 98	Julio 99 (6 meses)	Se presenta antes	Sí, el informe era para 1/99
Castilla y León I (94-96)	Sept. 1996 (a 4 meses)	Sept. 96	Marzo 97 (6 meses)	3 meses	No
Galicia II (95-97)	Julio 1996 (a 18 meses)	Nov. 96	Feb. 98 (15 meses)	3 meses	Sí, el informe era para 10/97
Madrid II (93-95)	Enero 1995 (a 12 meses)	Nov. 95	Dic. 96 (13 meses)	1 año	Sí, se retrasa el inicio

Media***

11,3 meses

Media****

11,8 meses

* Se refiere al número de meses que faltan para que se finalice el periodo de vigencia del plan

** El informe final es devuelto a la consultora externa para que se efectúen correcciones y no se entrega hasta septiembre de 1999

*** La media de las once experiencias de evaluación, exceptuando las evaluaciones del País Vasco y La Rioja, es de 11,8 meses.

**** La media de las once experiencias de evaluación, exceptuando las evaluaciones del País Vasco y La Rioja, es de 13,2 meses.

Respecto a la **duración** de las evaluaciones, los procesos de evaluación son también bastante heterogéneos entre sí y van desde los casi dos años y medio de la evaluación andaluza a los tres meses de la evaluación canaria. Lógicamente las dimensiones de los procesos evaluativos son también distintas y, por lo general, a mayor complejidad de los procesos, mayor duración de los mismos. Sin embargo, aquí cabe analizar también el sentido de urgencia con el que se plantean las evaluaciones, ya que éste varía profundamente dependiendo del contexto y situación evaluativa determinada.

Por ejemplo, la urgencia con la que se plantea la evaluación del II Plan andaluz es mucho menor que con la evaluación del II Plan castellano-manchego. En el caso de Andalucía, la posibilidad de una evaluación estaba planteada desde un principio del plan, ya que existían precedentes de evaluación del I plan; incluso se tenía claro que se iba a tratar de una evaluación externa. El caso es que ésta no se plantea como tal hasta la mitad del período de vigencia y el proyecto de evaluación que se plantea entonces es muy ambicioso y requiere de tiempo por delante, entre otras cosas, porque incluye un estudio de la situación social de la mujer andaluza. En principio esto no parece causar problemas y, de hecho, la evaluación se plantea, desde un principio, a dos años vista y siendo conscientes de que los resultados de la evaluación no van a estar hasta 15 meses después del final del período de vigencia del plan: en realidad, esta situación no genera grandes problemas, ya que el II plan andaluz se “prorroga” un año más. En definitiva, en Andalucía la evaluación condiciona los tiempos del plan.

Sin embargo, los tiempos de la evaluación del II Plan en Castilla-La Mancha¹ están claramente determinados por el contexto político: en junio de 1999 hay elecciones autonómicas y parece que la Dirección General de la Mujer quiere dejar aprobado un nuevo plan de igualdad antes de las mismas, incluso aunque el III Plan quede aprobado antes de que haya acabado el periodo de vigencia del II², y provocando la finalización del II plan antes de tiempo, como de hecho ocurre. Estas “prisas” hacen que la evaluación tenga que concebirse de forma necesariamente rápida y con una duración corta ya que sus resultados deben estar disponibles para la formulación del III Plan³. En este caso, ocurre al contrario que en Andalucía, dictándose los tiempos de la evaluación a través de los tiempos de los planes correspondientes y dándole prioridad a estos últimos sobre su evaluación. Si se tuviera que hacer una clasificación a este respecto, diríamos que los procesos de evaluación habidos en Canarias y Castilla y León tienen un perfil parecido al de Castilla-La Mancha, dando prioridad a los planes por encima de la evaluación. Por el contrario, en los casos de Galicia y Madrid, parece que las evaluaciones condicionan más los tiempos y el calendario de los planes.

Además de en Castilla-La Mancha, también en Canarias se presenta la evaluación antes de que acabe el periodo de vigencia del plan evaluado (en el caso de Canarias se trata sólo de dos meses antes). En el resto de los casos, los informes de evaluación se presentan siempre algún tiempo después de la finalización, aunque también existe variedad respecto a estos tiempos: en Castilla y León y Galicia, se puede decir que, prácticamente, se presentan a tiempo (entre un mes y tres meses después); sin embargo, en el caso de Madrid, y, sobre todo Andalucía, se presentan más de un año más tarde.

¹ La evaluación del II Plan Castellano-Manchego es mucho más tardía y más reciente en el tiempo que el resto de las experiencias de evaluación, ya que se realiza a comienzos del año 1999 y el resto se concentra más bien en los años 1996 y 1997. Es una evaluación que se hace en un periodo corto de tiempo y en el marco de la elaboración del III Plan.

² Este solapamiento de tiempos es la primera vez que ocurre, ya que, por lo general, suele ocurrir lo contrario, existiendo lagunas de tiempo entre el final de un plan y la aprobación del siguiente.

³ De hecho, y como se verá más adelante, los resultados de la evaluación no se consiguen formalmente antes de la elaboración del III Plan, y, en la realidad, los procesos de evaluación del II Plan y elaboración del III transcurren de forma prácticamente paralela. Sin embargo, se consigue una relativa utilidad de los procesos de evaluación para la elaboración del siguiente plan, en el sentido de que es el mismo equipo de personas (la Fundación Mujeres) el que realiza las dos tareas, facilitándose enormemente la transferencia de información.

Esta información se analizará también más adelante cuando se estudie la utilidad y utilización de las evaluaciones.

Por último, parece interesante analizar los **retrasos** habidos en los procesos de evaluación. En la mayoría de estos procesos se producen retrasos sobre el calendario inicial y esto afecta a la globalidad de la evaluación, así como a su potencial utilidad. Los retrasos en Andalucía se producen a lo largo de todo el proceso de evaluación. En el caso de Castilla-La Mancha, el retraso produce que se apruebe el III Plan con anterioridad a que esté listo el informe de evaluación del II Plan. En el caso de Galicia, los retrasos son debidos, principalmente, a las dificultades en la recopilación de información, pero se trata de un retraso que, aun siendo de cuatro meses, no resulta especialmente significativo, ya que se presentan los resultados a los tres meses de que haya finalizado el plan. Finalmente, en el caso de Madrid, lo que ocurre es que se retrasa el inicio del proceso de evaluación debido a que coincide con el período de elecciones autonómicas y con un cambio de gobierno y de partido en el gobierno: aunque la evaluación se había planteado al inicio del año 1995 no se resuelve su ejecución hasta finales de ese mismo año, una vez han pasado las elecciones (junio 1995) y se dan los cambios de estructura, responsables y equipos en la Dirección General de la Mujer. Este retraso inicial es el que va a caracterizar todo el proceso de evaluación, aunque el calendario de la evaluación externa se cumpla sin retrasos.

2.2. Momentos de los procesos de evaluación mixta

En la **tabla 6.2. b** se reflejan los momentos en los que se plantea la evaluación, su duración, el tiempo transcurrido entre el final del plan y la presentación del informe final y los retrasos, de las evaluaciones mixtas

TABLA 6.2. b: TIEMPOS Y MOMENTOS DE LAS EVALUACIONES MIXTAS SEGUNDA GENERACION DE EVALUACIONES

	Momento en el que se plantea la evaluación	Duración		Tiempo entre final plan e informe final	Retrasos
		Principio	Final		
Instituto Mujer II (93-95)	Julio 1994 (a 18 meses)*	Enero 95	Marzo 97 (27 meses)	1 año, 3 meses	No
Cataluña II (94-96)	Enero 1996 (a 12 meses)	Junio 96	Mayo 97 (11 meses)	5 meses	Sí hay retrasos (aprox. 6 meses)
País Vasco II (96-2000)	Desde el inicio	Evaluación continua, con memorias anuales		-	Las memorias anuales se presentan 1er semestre año siguiente

* Se refiere al número de meses que faltan para que se finalice el período de vigencia del plan

El caso de la evaluación del II PAPME vasco (al igual que la del II plan riojano) es una excepción en lo que se refiere a los momentos de la evaluación, ya que ésta se plantea de forma continua y desde el inicio del período de vigencia. Los tiempos se determinan por la recopilación y el procesamiento continuo de información y por la presentación de memorias anuales. En este caso, además, tenían ya una experiencia acumulada y un rodaje con la evaluación del I PAPME, de forma que la función de evaluación se establece de forma continuada respecto al plan anterior.

En las evaluaciones mixtas los momentos en los que se plantea la evaluación (a 18, a 12 meses y desde el inicio) son sensiblemente anteriores a los momentos en las que se plantean las externas, que parecen más precipitados (media de 11,3 meses)⁴. Esto quiere decir que parecen ser procesos más conscientes y que demuestran, en cierto modo, una función de evaluación más normalizada en los organismos de igualdad correspondientes. También son procesos de evaluación que duran más en el tiempo, con una media de 19 meses (sin tener en cuenta el País Vasco) que, comparada con la media de las evaluaciones externas (11,8 meses), resulta mucho mayor. Esto es normal, desde

⁴ Véase la tabla 5.6. a.

el punto de vista que los procesos internos de evaluación suelen extenderse más en el tiempo que los externos.

La evaluación que se realiza del II Plan nacional desde el IM se presenta mucho tiempo después de la finalización del plan (un año y 3 meses), pero curiosamente no se retrasa sobre el calendario inicial de evaluación, ya que ésta se concibe con retraso y se plantea su grueso una vez ha finalizado dicho plan. Parece, por lo tanto, que no existía una preocupación especial por obtener los resultados para utilizarlos para la elaboración del siguiente plan (de hecho, la síntesis de los resultados de la evaluación se publica a la vez que se aprueba el III Plan nacional). En Cataluña, sin embargo, aunque sí existe un cierto retraso sobre el calendario oficial (aunque se plantean la evaluación a un año vista y son relativamente rápidos en su ejecución, se van dando los retrasos típicos en procesos internos de evaluación que se acumulan en un retraso general de unos seis meses), todavía da tiempo a que los resultados sean utilizados para la elaboración del III Plan catalán que se aprueba, finalmente, a primeros de enero del año 1998. De hecho, parece que la propia evaluación condiciona los tiempos de elaboración del III, al menos formalmente. Sin embargo, en el IM evaluación del II plan y elaboración del III parecen procesos independientes.

En el País Vasco, los tiempos y momentos de evaluación están más normalizados, existiendo una evaluación continua que produce unas memorias anuales que son presentadas al Parlamento Vasco en el primer semestre del año siguiente.

2.3. Momentos de los procesos de evaluación interna

Como en el caso de la evaluación del II PAPME vasco, la evaluación del II plan riojano es una excepción en lo que se refiere a los momentos de la evaluación, ya que ésta se plantea de forma continua y desde el inicio del período de vigencia. También en este caso los tiempos se determinan por la recopilación y el procesamiento continuo de información y por la presentación de memorias anuales. En el caso riojano, el primer año de vigencia del plan -1996- se emplea, en una buena parte, en establecer los criterios e indicadores para la evaluación.

En el caso de Aragón se da prioridad al plan por encima de la evaluación, ya que ésta última se realiza adaptándose a los tiempos de aquél. En la tabla 6.2. c se reflejan los tiempos y momentos de las evaluaciones internas.

TABLA 6.2. c: TIEMPOS Y MOMENTOS DE LAS EVALUACIONES INTERNAS

SEGUNDA GENERACION DE EVALUACIONES

	Momento en el que se plantea la ev.	Duración		Tiempo entre final plan e informe final	Retrasos
		Principio	Final		
Aragón I (94-96)	Abril 1996 (a 8 meses)	Abril 96 -Enero 97 (10 meses)		1 mes	No
La Rioja II (96-99)	Desde el inicio	Evaluación continua, con memorias anuales		-	Se retrasa algo el inicio en el 96

3. CRITERIOS E INDICADORES DE EVALUACIÓN: FUENTES Y PROCESO DE ESTABLECIMIENTO

Como tercer elemento clave de los procesos de evaluación, se analizan los criterios de evaluación utilizados. Sin ánimo de quitarle importancia al resto de los elementos tratados, creemos que éste es precisamente el elemento más importante y claramente crucial en la definición de los procesos de evaluación.

El contenido de las evaluaciones, es decir, el QUE se evalúa, se concreta a través de los criterios e indicadores de evaluación que representan dos cuestiones fundamentales: por un lado, las variables estudiadas y, por otro, la base lógica en la que se fundamentan los juicios de valor que se establecen posteriormente en el proceso de evaluación. En general, en la práctica totalidad de las experiencias de evaluación metaevaluadas es muy débil la fundamentación de la concreta elección y definición de los criterios e indicadores utilizados. En muchos casos hemos tenido que recurrir a los instrumentos de recopilación utilizados para poder extraer (suponer) las variables estudiadas, ya que en los informes de evaluación no se habían especificado los criterios de evaluación en la mayoría de las experiencias estudiadas.

Con el fin de exponer los criterios de evaluación (supuestamente) seguidos en los procesos de evaluación, de una forma más sintética, se han confeccionado las tablas 6.3., dividiéndolas de nuevo por evaluaciones externas (6.3.a), mixtas (6.3.b) e internas (6.3.c).

3.1. Criterios e indicadores en las evaluaciones externas

A continuación se expone la **tabla 6.3.a.** en la que aparecen los criterios de evaluación para cada una de las seis experiencias externas.

**TABLA 6.3. a.: CRITERIOS Y PREGUNTAS DE EVALUACION
EVALUACIONES EXTERNAS**

	Criterios o preguntas
Andalucía II (95-97)	<p>Evaluación de la conceptualización y el diseño:</p> <ol style="list-style-type: none"> 1. ¿Qué fines se persiguen en las distintas áreas del II Plan? ¿Están bien relacionados con el fin último del II Plan, esto es, la igualdad de oportunidades? 2. ¿Son adecuados los objetivos planificados en cada área para conseguir esos fines? ¿Están bien seleccionados los indicadores? 3. ¿Responden esos objetivos a necesidades sociales que se trata de abordar? <p>Evaluación de la ejecución: hasta qué punto el Plan está alcanzando a las beneficiarias y en qué medida el desarrollo del Plan responde a su diseño inicial. Se atiende a: 1. Justificación institucional del objetivo 2. Planificación de acciones para lograr el objetivo 3. Implementación o ejecución efectiva de las acciones.</p> <p>Evaluación de la eficacia e impacto: (<i>En el proyecto</i>)</p> <ol style="list-style-type: none"> A) ¿Ha sido el Plan efectivo en alcanzar sus objetivos iniciales? B) ¿Se pueden explicar los resultados del Plan como consecuencia de un proceso alternativo (I Plan, otras campañas..)? Es decir ¿Son los efectos del Plan debidos al Plan y no a otros factores? C) Análisis coste-beneficio (<i>En el informe final</i>). Se dice que se analiza el grado de satisfacción de las usuarias/beneficiarias de II Plan con respecto a los servicios, actividades y programas del Plan (Beneficio subjetivo que ha generado el Plan)
Canarias I (95-96)	<p>Grado de cumplimiento de los objetivos previstos y alcance de las acciones (en realidad se trata de relacionar las diferentes acciones ejecutadas ordenadas según las áreas del Plan y si las han realizado el ICM u otras instituciones canarias). Se consigna el tipo de actuación según una clasificación de contenido con 4 categorías: actuaciones informativas y formativas, iniciativas culturales, ayudas económicas y otro tipo de actuaciones.</p>

<p>Castilla-La Mancha II (95-99)</p>	<p>Evaluación del diseño: Sobre los objetivos y las actuaciones, A) Coherencia, B) Concreción, C) Posibilidades de realización y D) Puntos fuertes y débiles (tipo de programa, colectivos destinatarios, cooperación entre entidades, integración de recursos, entorno y realidad de las mujeres, “pistas” para desarrollar las actuaciones y continuidad de los programas)</p> <p>Memoria de resultados de ejecución del Plan: Mapa de resultados obtenidos en función de los objetivos, a través de la actuación de los diferentes departamentos de la Junta implicados.</p> <p>Valoración cualitativa del desarrollo del Plan: Opiniones de las personas implicadas en el desarrollo del Plan sobre los niveles de ejecución, el nivel de implantación en la sociedad y los avances conseguidos. Se recogen impresiones sobre diferentes aspectos de la ejecución del II Plan, como el diseño de los contenidos, el proceso de puesta en marcha desde el punto de vista de la organización, el nivel de implicación de las diferentes entidades, la dotación de recursos, la valoración de los resultados obtenidos, el nivel de conocimiento e impacto en la sociedad y propuestas y sugerencias que deberían ser líneas de futuro en el desarrollo de políticas de igualdad en la región.</p>
<p>Castilla y León I (94-96)</p>	<p>Cambios acaecidos en la situación de las mujeres de Castilla y León, durante el período de vigencia del PIIOM: en la disponibilidad de empleo, en la demanda y las condiciones de empleo, en la facilitación de empleo, en la evolución de la salud, en las anomias padecidas por las mujeres y en la oferta y la demanda de educación entre las mujeres. También se consigna información sobre las acciones culturales y de cooperación, la promoción del asociacionismo y la participación institucional en el marco del PIIOM.</p> <p>Se hace un análisis comparativo de la situación de las mujeres en Castilla y León, combinando los siguientes criterios:</p> <ul style="list-style-type: none"> - El estado de las mujeres antes de y durante el período de vigencia del plan - La situación en Castilla y León respecto al conjunto nacional - El cotejo de los datos referidos a las mujeres, con los datos referidos a los hombres

<p>Galicia II (95-97)</p>	<p>Grado de cumplimiento de las acciones según las áreas del II PIOM (asociacionismo, educación y cultura, formación y empleo, ordenamiento jurídico, salud y servicios sociales)</p> <p>Evolución de las actividades por años</p> <p>Distribución según tipo de actividad (información y formación; programas, planes, proyectos y servicios; campañas; convenios, decretos y convocatorias; estudios, publicaciones y folletos; subvenciones, ayudas y becas; datos estadísticos)</p> <p>Clasificación de las actividades según el carácter de las actuaciones (directas si se ponen en marcha para cumplir los fines del Plan, indirectas si responden a otros objetivos pero atienden indirectamente)</p> <p>Clasificación de las actuaciones según el contenido (información, divulgación, sensibilización; formación y educación; empleo, inserción laboral e integración; investigación y análisis de la situación social de las mujeres; y tratamiento de problemas sociales y de salud)</p> <p>Clasificación de las actividades según ámbitos de impacto /destinatarios (población general, centros y asociaciones, mujeres y profesionales)</p> <p>Participación institucional y entidades responsables de las actuaciones</p> <p>Ámbito geográfico en donde se realizan las actuaciones</p> <p>Financiación de las actividades</p> <p>Otros aspectos: canales de difusión y de información de las actividades, material utilizado, directivas europeas en relación a las actividades, convenios existentes para la realización de actividades.</p> <p>Valoración de las actuaciones por parte de los responsables de las mismas: grado de participación y nivel de respuesta de las destinatarias, impacto en los medios de comunicación, grado de consecución de los objetivos previstos, contribución potencial a las políticas de igualdad, repercusión para la sensibilización de las mujeres y grado en el que se enmarcan en las prioridades del PIOM.</p> <p>Se elaboran una serie de indicadores para cada objetivo y cada acción.</p>
--------------------------------------	---

<p>Madrid II (93-95)</p>	<p>No se especifican criterios de evaluación. Los indicadores utilizados son:</p> <p>I. Características demográficas: estructura de edades y género de la población en la CM, evolución de las composiciones familiares.</p> <p>II. Integración de las mujeres en el mercado laboral: situación en relación con el empleo; actividad según sectores económicos, según tramos de edades; características de la inactividad; dependencia del hogar respecto de los ingresos de la mujer que trabaja; paro según tramos de edades; relación existente entre la forma de conseguir y buscar trabajo. Todos estos indicadores se presentan por cada género y en comparación con el conjunto nacional.</p> <p>III. La evolución de la calidad de vida y el empleo del tiempo de las mujeres: el tiempo que pasaba en 1993 y el que pasa en 1996 en distintos lugares; percepción de las relaciones con la pareja y con las hijas e hijos; percepción de la situación económica familiar; vinculación de las tareas domésticas con otras ocupaciones; ocupación de las mujeres en tareas domésticas, en atender a sus familiares, a necesidades personales, tiempo libre; implicación en las tareas domésticas del resto de los miembros de la familia; feminización y masculinización en la elección de profesiones; audiencia de los medios de comunicación de masas en la CM. De todos estos datos se presentan datos de evolución 1993-96. Mujeres beneficiadas por los programas sociales en servicios especializados (DGM); características de las mujeres atendidas en los servicios jurídicos y motivos de las consultas.</p> <p>IV. La atención a la salud integral de las mujeres: causas de muerte de mujeres en la CM; mortalidad por tumores; SIDA; accidentes laborales con baja entre mujeres.</p> <p>V. La participación política y social de las mujeres: participación femenina en los parlamentos autonómicos; % de alcaldesas; proporción de profesionales y personal directivo entre las personas activas; asociaciones registradas con subvención.</p> <p>VI. La violencia que afecta a las mujeres: violencia y malos tratos; nº de delitos y faltas contra las mujeres en el ámbito familiar, y según relación de la persona agresora con la víctima; abusos y malos tratos, según el lugar donde se producen; distribución por edades de delitos; delitos de agresión sexual; asistencia a mujeres violadas; población reclusa femenina penada y según su condición procesal; la delincuencia en menores según los motivos de detención y diferencias según género.</p> <p>VII. La escolarización y los niveles de instrucción reglada: población según nivel de instrucción y género; población escolarizada por clases de enseñanza; universitarios graduados/as, por áreas de enseñanza;</p> <p>VIII. El analfabetismo y la formación de adultos: proporciones de analfabetismo declarado; población que sigue cursos para adultos/as.</p> <p>IX. Imagen de la DGM: cometido principal que le asignan a la DGM</p>
-------------------------------------	--

La evaluación del **II PIOM andaluz** es la única que se plantea la evaluación del diseño del plan, una verdadera (aunque quizá no completa) evaluación de la ejecución o implementación y una evaluación de los efectos netos del Plan, distinguiendo, además, entre productos o *outputs* y efectos o *outcomes*. Sin embargo, los criterios sobre la eficacia y el impacto del Plan están muy poco explicitados en el proyecto de evaluación, a diferencia de los criterios de los otros dos tipos de evaluación. Es decir, preguntarse en general si el plan ha sido efectivo sin especificar cuáles son los efectos que se buscan, significa quedarse en un plano muy abstracto. De hecho, el proyecto inicial menciona la medición de efectos y pretende, por lo tanto, muchas más cosas de lo que posteriormente aparece como evaluación de impacto en el informe final (grado de satisfacción de las usuarias). Sin embargo, y como se ha dicho antes, es importante y clarificadora la introducción del concepto de efectos netos y la distinción entre productos y efectos. Como se dice en el propio proyecto de evaluación “el Plan puede haber sido muy productivo por el gran número de campañas que haya realizado, los folletos que haya publicado y las actividades que haya puesto en marcha. Sin embargo, los resultados (*outcomes*) pueden ser nulos, negativos o causados por factores ajenos al Plan” (pag. 9). En el Avance del Informe de Evaluación, presentado al IAM en Julio de 1997, en el que se expone la evaluación del diseño y los proyectos de evaluación de la ejecución y de eficacia e impacto, de nuevo, se especifica mucho el diseño de recopilación de la información, pero se sigue sin especificar realmente cuáles van a ser los criterios de evaluación a utilizar en la evaluación de resultados. Sólo al especificar el estudio 1 a realizar sobre los objetivos del Plan se dice que “se pretende conocer en qué medida las mujeres andaluzas conocen las diferentes acciones puestas en marcha y quiénes han participado, con el objeto de poder comparar ambos colectivos, identificarlos y diferenciarlos” (p. 58). Sin embargo, en el estudio 2 se dice que “se intenta evaluar en qué medida los cambios detectados se deben a la participación de las mujeres en actividades” (p. 60). Y en el estudio 3 que se pretende analizar el “cambio observado como el efecto que es atribuido al propio Plan” (p. 62). Es decir, se sigue sin especificar cuáles son los aspectos o dimensiones sobre los que se van a observar dichos cambios.

Por **grado de cumplimiento de los objetivos (Canarias)** se entiende la recopilación de información sobre las actuaciones realizadas, ordenadas según las áreas en las que está dividido el plan. Es decir, se trata de una descripción ordenada de lo que se ha hecho. Si se realizan algunas valoraciones éstas suelen ser de tipo muy general; normalmente, cuando se habla del “grado de cumplimiento”, los estudios de evaluación se limitan a expresar, en porcentajes, las actuaciones realizadas sobre las previstas inicialmente. En el caso de la evaluación **castellano-manchega**, este tipo de “evaluación” se denomina “Memoria de resultados de ejecución del plan” o “Mapa de resultados”, términos más modestos que los anteriores, ya que una memoria o un mapa no tiene por qué implicar evaluación.

Al igual que en la evaluación del **II Plan andaluz**, también se realiza explícitamente evaluación del **diseño** en el **II Plan castellano-manchego** (sólo en estas dos evaluaciones se realiza como tal). Pero existen diferencias entre estas dos evaluaciones de diseño. Mientras en la andaluza se pretende realizar una evaluación más global, que estudia los objetivos en su conjunto (lógica de conjunto, jerarquía y compatibilidad entre ellos y adecuación a las intenciones iniciales y a las necesidades reales de las mujeres andaluzas, además de una evaluación de los indicadores

propuestos), en la castellano-manchega se analiza la concreción, coherencia, posibilidades de realización y puntos fuertes y débiles de cada uno de los objetivos y actuaciones por separado. En la evaluación castellano-manchega se hace un esfuerzo por especificar, además, los criterios utilizados; es decir, en una “tabla de correspondencias” (p. 15 del Informe de evaluación) se especifica para cada tipo de análisis (coherencia del objetivo, coherencia de la actuación, concreción del objetivo, concreción de la actuación y posibilidad de realización), cómo se definen las tres categorías para cada uno de esos análisis (coherencia, concreción o posibilidad de realización ALTA, MEDIA O BAJA)⁵. Los puntos fuertes y débiles también están especificados (aunque se hace “a posteriori”, cuando ya se tiene la información), ya que se identifican una serie de factores (entre paréntesis en la tabla 6.3. a) que determinan la bondad o fortaleza de las actuaciones que se analizan. Menos clara resulta lo que llaman la “Valoración cualitativa del desarrollo del Plan”. Aunque parece que inicialmente pretenden recoger las opiniones de las personas implicadas sobre el proceso, ejecución y desarrollo del Plan, los objetivos aumentan a lo largo del proceso y, finalmente, consignan que recogen información no sólo sobre el proceso de puesta en marcha desde el punto de vista de la organización, el nivel de implicación de las diferentes entidades y la dotación de recursos sino también sobre la valoración de los resultados obtenidos, el nivel de conocimiento e impacto en la sociedad y el diseño de los contenidos. También recogen sugerencias ya que, como se ha comentado anteriormente, en Castilla -La Mancha, el proceso de recopilación de información para la evaluación del II Plan y para la elaboración del III fue paralelo (es decir, se aprovecharon los mismos grupos de discusión y las mismas entrevistas).

Los criterios de evaluación que se utilizan en la “evaluación” del **I Plan castellano y leonés** tienen que ver con la evolución de las mujeres de la región en el período de vigencia del Plan, en una serie de dimensiones (que se relacionan en la tabla 6.3. a, pero que no refleja ninguna conexión con lo que los poderes públicos han hecho para promover la igualdad. La única información directamente relacionada con el Plan, es la de recoger algunos datos muy limitados sobre algunas actuaciones realizadas respecto a acciones culturales y de cooperación y a la promoción del asociacionismo y de la participación. Normalmente son datos en relación al gasto realizado (es decir, por ejemplo, “realización de una campaña de concienciación de empresarios con un costo de 10.000.000 ptas.” (p. 28 del “Resumen de la Evaluación”). Respecto a la promoción del asociacionismo los datos que se ofrecen son el monto de las subvenciones y datos que no tienen que ver directamente con el Plan como son la afiliación sindical de las mujeres y su participación en el Parlamento Autonómico y en los gobiernos locales. Los datos que se recogen respecto a la participación institucional se limitan a contar en el número de foros en los que ha participado la Secretaría Regional de la Mujer en cada provincia para la conmemoración del Día de la Mujer.

En la evaluación del **II Plan gallego**, el tipo de evaluación que se realiza es una evaluación de *outputs* (saber qué es lo que se ha hecho y qué es lo que ha producido los

⁵ Por ejemplo, respecto al tipo de análisis “CONCRECIÓN DEL OBJETIVO”, la categoría ALTA se define como “contenido preciso, claro, conciso y sin ambigüedades” y “en el propio contenido se incluye el modo de realización”; la categoría MEDIA se define como “contenido poco preciso, con cierta ambigüedad” y “no incluye el modo o ‘pistas’ para su realización”; y, por último, la categoría BAJA se define como “contenido no preciso, poco claro y/o con ambigüedades” y “no incluye el modo o ‘pistas’ para su realización”.

respectivos Planes), en la que se recoge más información que en la evaluación canaria del “grado de cumplimiento”, ya que se clasifican las actividades según diferentes dimensiones y se tienen en cuenta los colectivos destinatarios, además de los agentes responsables. Este sistema de recopilación de información podría resultar una buena base también para una cierta evaluación del diseño del Plan, permitiendo la reflexión y valoración sobre qué tipo de actividades se realizan en mayor o menor medida, quiénes son los agentes institucionales a los que les cuesta más participar y a quiénes menos y a qué sectores de la población se les está dando mayor o menor prioridad. Esto podría dar origen a una reflexión no sólo sobre si se está ejecutando lo que se tenía previsto sino, incluso, si lo que se tenía previsto y se realiza guarda una lógica con lo que se pretende, si se está abordando el problema público de la desigualdad de género de forma coherente con las premisas políticas y cuál es el concepto de igualdad que los poderes públicos están transmitiendo realmente con sus actuaciones. Sin embargo, este tipo de reflexión no parece hacerse en la evaluación gallega.

Finalmente, en la evaluación del **II Plan madrileño** tampoco aparecen explicitados los criterios de evaluación; de hecho, ni siquiera existen, ya que en dicha “evaluación” no se valora la información recopilada. La recopilación de información se realiza, en esta evaluación, a través de un sistema de indicadores elaborado a partir de los objetivos del Plan. En la **tabla 6.3. a** referente a los criterios de evaluación utilizados, se reflejan dichos indicadores ya que son los aspectos o dimensiones sobre los que se ha recopilado información, es decir, el QUE de dicho estudio. Es decir, aunque no existan criterios de valor específicos (ya que no se emiten juicios) se explicitan en este apartado las variables estudiadas, reflejadas a través de los indicadores utilizados. Como se puede observar, la mayoría de las variables tienen que ver con la situación social general de la mujer y no tanto con el Plan evaluado.

En general, uno de los indicadores de la poca importancia concedida a la definición previa de los **criterios de evaluación (qué se va a evaluar)** es cómo, en muchos de los informes de evaluación, en los capítulos correspondientes al diseño y metodología de la evaluación, no suele aparecer un apartado específico para ello. Ni en el informe canario, ni en el de Castilla y León ni en el de Madrid se hace ninguna referencia a los criterios. En el informe gallego se pasa de los puntos referidos al concepto y metodología de la evaluación y al punto de partida de la evaluación, a los puntos referidos a la recogida de información, los instrumentos, el trabajo de campo y el procesamiento de dicha información. En el caso del informe final andaluz, sin embargo, sí se explican los criterios generales que se han utilizado.

3.2. Criterios e indicadores en las evaluaciones mixtas

Como ya se ha comentado en el apartado de las evaluaciones externas, en casi todas las experiencias de evaluación analizadas, la fundamentación de la concreta elección y definición de los criterios e indicadores utilizados suele ser muy débil. Este es el caso de la evaluación del IM, en los que se ha tenido que recurrir a los instrumentos de recopilación analizados para poder extraer las variables estudiadas, ya que en los informes de evaluación no se habían especificado los criterios de evaluación. En las evaluaciones catalana y vasca, aunque son algo más claras respecto a los criterios utilizados, tampoco se hacen claramente explícitos en los informes de evaluación.

Ya se ha visto que en muchos de los informes de evaluación, en los capítulos correspondientes al diseño y metodología de la evaluación, no suele aparecer un apartado específico para dichos criterios. En el caso de la evaluación catalana existe un apartado en el que se relacionan los aspectos considerados en el análisis de las acciones a evaluar; en este apartado se explican dichos aspectos, aunque no se fundamentan de forma específica, ni se explicita por qué se consideran dichos aspectos y no otros. Incluso la evaluación del II PAPME vasco, que como ya hemos visto tiene unos propósitos más claros y explicitados que en las otras experiencias, del punto en el que se explican los objetivos se pasa a la metodología. Los criterios de evaluación se deben entresacar del punto en el que se especifican los objetivos de la evaluación⁶.

En la **tabla 6.3.b.** que aparece más adelante se exponen los criterios para cada una de las tres experiencias mixtas.

⁶ Documento sobre evaluación, presentado por Emakunde al Seminario sobre Evaluación de Políticas de Igualdad en la UIMP, Santander, Junio 1998. p. 16.

**TABLA 6.3. b: CRITERIOS Y PREGUNTAS DE EVALUACION
EVALUACION MIXTAS**

	Criterios o preguntas
<p>Instituto Mujer II (93-95)</p>	<p>Evaluación de la implementación: Se pretendía:</p> <ul style="list-style-type: none"> - conocer el grado de cumplimiento del Plan, es decir, qué actuaciones, de las 170 previstas, se han puesto en marcha y cuáles no - conocer cómo se ha ejecutado el Plan, es decir, qué actividades se han llevado a cabo para dar cumplimiento al mismo. <p>Se recoge información sobre datos de identificación del indicador o actividad (área, objetivo, actividad), temporalización (fechas de inicio y final, reales y previstas), grado de desarrollo, motivo de incumplimiento si procede, presupuesto invertido y descripción del indicador o actividad desarrollada.</p> <p>Evaluación cualitativa: ¿Cuál es la valoración sobre el PIOM de grupos representativos del colectivo de mujeres (percepción social que, de los efectos de las actuaciones, han tenido sus destinatarias más directas) y por las instituciones que han jugado un papel importante en su ejecución (percepción personal que del desarrollo y los efectos de las actuaciones han tenido los responsables de la coordinación desde los distintos ministerios)?</p> <p>Evaluación cuantitativa: ¿Cuáles son los efectos secundarios de la puesta en marcha del Plan, en aspectos o dimensiones que subyacen y justifican a los posibles cambios estructurales que hayan podido tener lugar? Son:</p> <ul style="list-style-type: none"> - Estereotipos sexuales: estereotipia de rasgos (ver si determinados rasgos son atribuidos preferentemente a las mujeres y no a los hombres) y estereotipia de roles (acuerdo en una serie de frases referidas a roles sexuales). - Valores: instrumentales, referidos a modos de conducta deseables y finales, referidos a estados finales de la existencia. - Actitudes hacia la igualdad de géneros: predisposición a que las mujeres cambien del rol tradicional de madre de familia y ama de casa. - Calidad de vida: grado de satisfacción de las personas con aspectos de su vida, como la educación, la familia, la pareja, las amistades, la salud, la situación económica y el tiempo libre. Comparación entre lo que se tiene y aquello a lo que se aspira y desea. - Usos del tiempo: en necesidades personales, trabajo doméstico, educación, trabajo remunerado y tiempo libre. Ponderando de manera diferencial el tiempo dedicado a cada uno de estos grupos de actividades, se establecen los siguientes indicadores de igualdad: igualdad social, calidad de vida, autonomía, trabajo doméstico y productividad.

<p>Cataluña II (94-96)</p>	<p>Evaluación del cumplimiento del Plan: ¿Qué se ha hecho?</p> <ul style="list-style-type: none"> - Grado de implementación de las acciones (se han hecho totalmente, parcialmente, se encuentran en ejecución, se ha hecho algún tipo de actuación pero se desconoce el grado de implementación, no era procedente, no se han realizado), según ámbito (área del Plan): sensibilización, legislativo, sociocultural, coeducación y formación reglada, formación ocupacional y de trabajo, sanitario, servicios sociales y comunitarios. - Grado de consecución de los objetivos (se han cubierto de forma amplia por las acciones implementadas, de forma parcial, se realiza algún tipo de actuación pero se desconoce el grado de consecución, no era procedente, no ha habido ningún tipo de cobertura) - Análisis del grado de implementación de las acciones según: <ul style="list-style-type: none"> Tipo de política: promoción (su propósito es cambiar la percepción social sobre el papel de la mujer); igualdad (pretende cambiar las pautas sociales de asignación de recursos y oportunidades entre hombres y mujeres); paridad (pretende cambiar las pautas de relación hombre-mujer, en el ámbito público y privado). Tipo de actuación: específica (originada por el II Plan); genérica modificada (con otro origen, pero reformulada con el II Plan) y genérica (con otro origen y sin adaptación al II Plan). Ámbito territorial: según provincias. <p>Evaluación de la participación de los diversos actores responsables de implementar las acciones del Plan (Departamentos de la Generalidad)</p> <p>Evaluación de las acciones dirigidas a los diversos colectivos destinatarios del Plan</p> <p>Evaluación de los recursos destinados por ámbito, actor y colectivo destinatario</p>
---------------------------------------	---

<p>País Vasco II (95-98)</p>	<p>Tipo de actuaciones que se realizan desde los tres niveles de la Administración Pública Vasca (Departamentos de Gobierno y Organismos Autónomos; Diputaciones Forales; Ayuntamientos y Mancomunidades).</p> <p>Destinatarios: colectivos destinatarios de las actuaciones realizadas.</p> <p>Grado de ejecución del II PAPME: ¿en qué medida están siendo o no ejecutadas las acciones propuestas desde el mismo?</p> <p>Valoración del conjunto de actuaciones para que orienten la ejecución del II PAPME en toda su integridad durante su período de vigencia.</p> <p>Se definen, para la evaluación del II PAPME, siete grandes categorías, en las que se clasifican todas acciones propuestas en el II PAPME según el tipo de intervención. La información sobre la temática tratada en cada caso, el colectivo destinatario y el modo concreto en que se ha desarrollado la actividad se recoge y se trata en la evaluación pero no son criterio para esta categorización:</p> <ol style="list-style-type: none"> 1. SENSIBILIZACION, INFORMACION, DIVULGACION: En esta categoría se incluyen todas las acciones cuyos objetivos son aumentar el grado de sensibilización, incrementar la información correcta o disminuir la información incorrecta, a través de actividades diversas dirigidas a colectivos específicos o a la población general. 2. ORIENTACION, ASESORAMIENTO, APOYO: Incluye acciones con las que se pretende proporcionar apoyo, orientación o asesoramiento de diversos tipos, a través de servicios, materiales, recomendaciones, etc. 3. FORMACION: acciones dirigidas a aumentar, mejorar o introducir cambios en la formación general o específica de colectivos diferentes (mujeres, profesionales, profesorado, empresariado, políticos/as). También incluye acciones que suponen introducir nuevos programas de formación, incluir módulos nuevos en los programas de formación ya establecidos, etc. 4. REALIZACION ESTUDIOS, INVESTIGACIONES, EVALUACION: acciones que pretenden impulsar y favorecer la creación de conocimiento en campos diversos. 5. AMPLIACION, ADAPTACION DE RECURSOS: acciones que suponen cambios (crear, proporcionar, aumentar, adaptar) en cuanto a los recursos económicos, humanos, materiales, sociocomunitarios, etc., existentes en la comunidad. 6. COORDINACION, COLABORACION, INTERCAMBIO: acciones que tienen por objetivo aumentar la coordinación, la colaboración y el intercambio de experiencias entre entidades y/o agentes sociales, a través de actividades como jornadas, congresos, reuniones, acuerdos, convenios, etc. 7. INTERVENCION SOBRE NORMAS JURIDICAS Y SOCIALES: acciones que pretenden modificar situaciones de base que generan o favorecen la discriminación. Se incluyen acciones que benefician, directa o indirectamente tanto a las mujeres como a la sociedad en general; o bien acciones dirigidas a colectivos específicos que por encontrarse en situaciones extremas (personas con cargas familiares, con discapacidades, con adicciones, cumpliendo condena, etc.) requiere de una actuación específica para poder beneficiarse de un plan de igualdad de oportunidades.
---	---

En el caso del **IM**, la única referencia a los criterios utilizados en la evaluación **cuantitativa** es la relacionada con los “efectos secundarios” del Plan en aspectos o dimensiones subyacentes a los cambios estructurales. En el resumen publicado en el que se presentan los resultados, en el apartado de la evaluación cuantitativa se realiza esa única referencia a los criterios en una única frase (es decir, la referencia a **QUE** se quiere evaluar) y sin embargo, se desarrolla mucho la parte relacionada con el **COMO**, referida al procedimiento de muestreo y a las técnicas de recogida de información utilizadas. En la sección de técnicas de recopilación de información, además de las escalas de información, se hace referencia a los aspectos medidos, por lo que se puede entresacar algo sobre lo que se ha querido evaluar (dimensiones y aspectos que se relacionan en la tabla 6.3. b). Ocurre algo similar en la evaluación **cuantitativa**, aunque aquí no se extienden tanto en el aparatado metodológico; tampoco en este caso se fundamenta por qué los criterios de valor utilizados son la percepción personal del desarrollo y efectos del PIOM, por parte de técnicos implicados en los diferentes ministerios (agentes activos) y la percepción social (sólo de los efectos) de las destinatarias (población diana). Por otra parte, se sobreentiende que las destinatarias del PIOM son exclusivamente las mujeres, no el resto de las instituciones públicas, no la sociedad en general. Es decir, parece que se propicia la antigua concepción de las políticas de igualdad como promoción de la mujer y no como fomento de la inclusión de las relaciones y la perspectiva de género en el resto de las políticas públicas. Respecto a la evaluación de la **implementación**, en realidad se trata de una muy sencilla evaluación de *outputs* en la que simplemente se relacionan las acciones realizadas según las áreas del plan. En este caso, a diferencia de algunas experiencias en CC.AA., ni siquiera se clasifican las acciones según otra tipología, según el colectivo destinatario, presupuesto, etc.

En la evaluación del **II Plan catalán**, como en la del II plan gallego, el tipo de evaluación que se realiza es una evaluación de *outputs* (saber qué es lo que se ha hecho y qué es lo que ha producido los respectivos Planes), en la que se recoge más información que en las evaluaciones del “grado de cumplimiento”, ya que se clasifican las actividades según diferentes dimensiones y se tienen en cuenta los colectivos destinatarios, además de los agentes responsables. La reflexión a partir de esta información parece que se hace algo más en la evaluación catalana de lo que se hace en la gallega por ejemplo, ya que las catalanas llegan a conclusiones sobre el tipo de política de se está llevando a cabo (si promoción, igualdad y paridad), detectando que las actuaciones que podrían inscribirse realmente en una política de fomento de una verdadera paridad entre hombres y mujeres son, todavía, muy pocas.

La evaluación del **II PAPME vasco** se beneficia enormemente de la experiencia acumulada con la evaluación del I. Tanto los criterios (tabla 6.3. b) como la metodología de esta segunda evaluación están más pulidos que en la primera⁷. De hecho, se revisan

⁷ Los criterios de evaluación del I PAPME eran algo más sofisticados, y eran los siguientes:

Actividades realizadas: Naturaleza y contenido de las actuaciones ejecutadas, ¿Qué acciones han sido ejecutadas? ¿Cómo han sido ejecutadas? ¿A qué tipo, clase o estrategia corresponden?

las tipologías utilizadas en la evaluación del I PAPME, ya que “la utilización de las mismas había resultado compleja, en parte, por contemplar de modo paralelo el tipo de intervención y la temática y, por otro, porque la clasificación de las actividades en las diferentes tipologías se hacía a partir de la información suministrada en relación con la actividad, no siempre suficientemente exhaustiva” (documento presentado en el Seminario de la UIMP. 1998. p.18). Como se puede observar en la tabla 6.3. b, la tipología utilizada en la segunda evaluación resulta más sencilla que la utilizada en la primera. Además, la experiencia de evaluación vasca es en la que se ha reflexionado más sobre los criterios de evaluación y esto se refleja, lógicamente, en una mayor claridad del conjunto del proceso evaluativo.

Para clasificar las actividades se tienen en cuenta dos criterios:

- 1) La forma de actuación: qué se pretende y a quién se dirige (se establecen 9 tipos de actividades)
- 2) El contenido concreto de la actividad en cuanto al tema que aborda (se establecen subtipos de actividades dentro de cada uno de los 9 tipos generales)
 1. CULTURA/OCIO: 1.1. Formación básica; 1.2. Formación en habilidades; 1.3. Ocio y deporte; 1.4. Cultura general
 2. INFORMACIÓN/ASESORAMIENTO GENERAL: 2.1. Derechos y recursos; 2.2. Empleo; 2.3. Sobre actividades institucionales; 2.4. Violencia y malos tratos; 2.5. Familia.
 3. SENSIBILIZACIÓN: 3.1. Corresponsabilidad; 3.2. Coeducación y estereotipos sexistas; 3.3. Recursos, derechos, igualdad de oportunidades; 3.4. Uso no sexista del lenguaje; 3.5. Malos tratos/acoso sexual; 3.6. Situación de la mujer en general; 3.7. Salud; 3.8. Sexualidad.
 4. CREACIÓN DE CONOCIMIENTO DE GÉNERO: 4.1. Estudios, investigaciones, informes, monografías; 4.2. Difusión, jornadas, intercambios; 4.3. Becas, subvenciones.
 5. FORMACIÓN DIRIGIDA A PROFESIONALES: 5.1. Formación de formadores/as; 5.2. Formación a profesionales (educación, salud, etc.); 5.3. Asesoramiento a profesionales
 6. ACCIÓN DIRECTA ANTE AL DISCRIMINACIÓN: 6.1. Acción directa individual (formación ocupacional; asesoramiento: jurídico, empleo); 6.2. Acción directa estructural (legislativa, fomento empleo, servicios socio-comunitarios; 6.3. Premios y acción positiva; 6.4. Asistencial, servicios para paliar la discriminación (servicios específicos, programas de salud, sexualidad).
 7. FORTALECIMIENTO DE ASOCIACIONES DE MUJERES: 7.1. Formación en asesoramiento; 7.2. Fomento de asociacionismo; 7.3. Subvenciones; 7.4. Diseño de políticas/ coordinación institucional; 7.5. Asesoramiento.
 8. INTRODUCCIÓN TEMÁTICA DE GÉNERO EN ADMINISTRACIÓN. PÚBLICA
 - 8.1. Creación y gestión de estructuras específicas; 8.2. Utilización no sexista del lenguaje; 8.3. Coordinación institucional en los diferentes niveles (autonómico, central e internacional); 8.4. Documentos, difusión; 8.5. Subvención institucional; 8.6. Asesoramiento a las instituciones; 8.7. Estudios/ recogida de información.
 9. COOPERACIÓN INTERNACIONAL: 9.1. ONGs; Administraciones + ONGs.

Destinatarios: ¿Quiénes han sido los colectivos destinatarios de estas acciones?

Participación de la Administración Pública Vasca: Incorporación de las políticas a su quehacer cotidiano; percepción de la propia Administración sobre su quehacer y cómo lo traduce en actividades.

3.3. Criterios e indicadores en las evaluaciones internas

Tampoco en las evaluaciones internas, los criterios de evaluación quedan especificados. En los dos ejemplos de evaluaciones internas, que se señalan en la tabla 6.3. c, se reflejan procesos internos que se centran en las tareas de seguimiento de los planes y en la obtención de datos necesarios para el funcionamiento cotidiano y una cierta reflexión posterior. En ambos casos - Aragón y La Rioja- los criterios se centran en las acciones o actividades realizadas y su “grado de ejecución”, entendido como la realización o no de las mismas en los tiempos planificados. En La Rioja, además, se consignan los gastos y los organismos que participan.

Efectivamente, y como ya se ha comentado anteriormente, por “grado de ejecución” de los Planes se entiende la recopilación de información sobre las actuaciones realizadas, ordenadas normalmente según las áreas en las que están divididos los planes. Es decir, se trata más bien de una descripción ordenada de lo que se ha hecho.

**TABLA 6.3. c: CRITERIOS Y PREGUNTAS DE EVALUACION
EVALUACIONES INTERNAS**

	Criterios o preguntas
Aragón I (94-96)	Grado de ejecución de las acciones del Plan y organismo que las ejecuta, según la siguiente clasificación: A) Acciones realizadas: en su totalidad, acciones habituales (que se integran en la dinámica normalizada del IAM y deberán continuar impulsándose) y acciones realizadas parcialmente. B) Acciones sin realizar
La Rioja II (96-99)	Para cada actuación del Plan, se establecen una serie de indicadores (<i>Listado de Indicadores</i>). De cada actuación se evalúan, anualmente: - Actividades realizadas (en relación a los indicadores propuestos) - Organismos que participan - Gasto ejecutado (por la entidad responsable y por otras) - Grado de ejecución : si la actividad ha finalizado en el año de referencia o si tiene continuación en años posteriores - Temporalización de la ejecución de las medidas del Plan: para saber en qué año se ejecutan y las previsiones para los siguientes.

En la evaluación del **II plan riojano** se han entresacado los criterios de evaluación de las fichas con las que se recoge la información. En este caso los indicadores sí tienen que ver más directamente con la evaluación del plan (no como en Madrid y Castilla y León que son más bien indicadores de la situación general de la mujer), como por

ejemplo, nº de mujeres informadas sobre diversos temas, nº de planes distribuidos y departamentos donde se ha hecho, nº de reuniones, nº de acciones informativas, nº de documentos revisados, etc. (para ver el listado completo, consultar documento “Listado de indicadores”).

4. PROCEDIMIENTOS Y METODOLOGÍA EMPLEADA

El último elemento clave de los procesos de evaluación a considerar es el de los procedimientos y metodología utilizada en dichos procesos para recopilar y analizar la información.

De forma similar a lo que ocurre respecto al apartado anterior (referido a los criterios de evaluación) los procedimientos utilizados no suelen estar profusamente fundamentados en los informes; es decir, difícilmente encontramos una justificación de por qué se elige dicho instrumento y no otro. Sí encontramos, en algunos casos, especificaciones muy detalladas sobre los instrumentos en sí: estas especificaciones son siempre de carácter técnico y versan sobre las características intrínsecas de dichos instrumentos, no tanto sobre su valía o adecuación para la situación concreta en la que se aplica.

En algunos de los casos estudiados, se entiende la metodología y los procedimientos empleados de una forma más amplia y se especifican en los informes los pasos dados y las estructuras de coordinación utilizadas para la recopilación de información, como parte de dichos procedimientos. Los casos en los que parece existir esta concepción metodológica más amplia (o al menos se refleja en los documentos e informes resultantes de los procesos de evaluación), son la evaluación del **I Plan canario**, y el **II castellano-manchego** en cierta medida, y del **II Plan catalán**, el **II PAPME vasco** y el **II Plan gallego**, en mayor medida. Como se puede observar en la tabla 6.4.a, en estos casos, cuando se describen los procedimientos, se explica de alguna manera y de forma más o menos sistematizada, a través de qué mecanismos se ha recopilado la información, además de explicar las técnicas concretas utilizadas. En el caso de Galicia, además, se crea una estructura *ad-hoc* (la Comisión Interna de Evaluación) que es la que guía o pretende guiar el proceso de recopilación (en mayor medida) y análisis (en menor medida) de la información.

En las tablas 6.4.a, 6.4.b y 6.4.c (Procedimientos y metodología para la recopilación y análisis de la información para las evaluaciones externas, mixtas e internas respectivamente) se señalan, para cada experiencia, los procedimientos utilizados para la recopilación y el análisis de la información. Cuando no se consigna, es que no se especifica nada respecto al análisis de la información en los documentos e informes de evaluación analizados. Así como en los análisis cuantitativos se suele especificar de forma profusa los procedimientos de muestreo y los de análisis estadístico, esto no es así para los análisis de la información extraída a través de técnicas de tipo cualitativo ni a través de fichas o registros de información; en este caso se suele especificar más sobre el instrumento y menos sobre los procedimientos posteriores para analizar la información recopilada a través de dichos instrumentos.

4.1. Procedimientos y metodología de las evaluaciones externas

En el “Avance de resultados” de la evaluación del **II Plan andaluz**, en el que se expone la evaluación del diseño y conceptualización del plan, no se hace en ningún momento referencia a los procedimientos con los que se ha realizado el análisis, sino que se exponen los resultados de la misma directamente. Para saber algo sobre estos procedimientos, hemos tenido que recurrir al proyecto inicialmente presentado al IAM. Tampoco aparece nada en la publicación de los resultados de la evaluación del **II Plan madrileño**, no haciéndose referencia, en ningún momento, a los procedimientos empleados para la recopilación y el análisis de la información. Esta es una tendencia generalizada en todas las experiencias de evaluación estudiadas.

TABLA 6.4. a: PROCEDIMIENTOS Y METODOLOGIA PARA LA RECOPIACION Y EL ANALISIS DE LA INFORMACION. EVALUACIONES EXTERNAS

	Procedimientos y técnicas de recopilación y análisis
Andalucía II (95-97)	<p>Evaluación de la conceptualización y el diseño:</p> <p>A) Recopilación documental: el propio Plan, listado de indicadores y otra documentación escrita como documentos en los que se explica el programa, actas de reuniones, informes.</p> <p>B) Contactos informales con miembros del IAM, Consejerías implicadas y organizaciones de mujeres.</p> <p>C) Entrevistas con los responsables de cada una de las áreas prioritarias para conocer los posibles obstáculos que prevén encontrarse y su opinión sobre el grado de concreción de los objetivos.</p> <p>D) Juicio de personas con experiencia en programas similares (no se especifica la técnica a través de la cual se va a extraer dicho juicio).</p> <p>Evaluación de la ejecución:</p> <p>A) Análisis de actividades: 1) <i>descripción</i> (características, cobertura, tasa de abandono, calidad); 2) <i>obtención de datos empíricos</i> que la avalen (a través de análisis de <i>documentos</i> -análisis de datos secundarios que han sido generados por los responsables de ejecutar la actividad con fines distintos a la evaluación externa- y <i>entrevistas</i>). Para homogeneizar estos datos se realiza una ficha con una parte común y una específica, en función del tipo de actividad de que se trate. 3) <i>Comparación con la realidad</i>, comparando si lo que estaba previsto se ha llevado a cabo de la manera, en el tiempo y con los recursos previstos.</p> <p>B) Aspectos de la organización: examen del entorno institucional del II Plan (si se tienen en cuenta los objetivos de las organizaciones), trámites de acceso a las actividades y tiempo que tarda la institución en dar respuesta a través de <i>documentos</i> y <i>entrevistas</i>.</p> <p>C) Estudio Piloto: con tres acciones del Plan, para detectar posibles errores o deficiencias en la investigación prevista para subsanarlas a priori; comprobar las dificultades de obtención de información de los organismos implicados y en cómo se van a gestionar los datos y qué resultados se van a obtener.</p> <p>D) Propuesta de indicadores</p> <p>E) Diseño de un Sistema de información</p> <p>Evaluación de la eficacia e impacto:</p> <p>Programa de estudios a través de encuestas:</p> <p>A) Encuesta sobre la situación social de la mujer (realizada en 1988)</p> <p>B) Estudio 1: Encuesta sobre los objetivos del II PIOM: 1200 encuestas dirigidas a mujeres residentes en Andalucía, muestreo estratificado con afijación proporcional, con selección de los individuos por cuotas y rutas aleatorias.</p> <p>C) Estudio 2: Encuesta sobre las actuaciones del II PIOM en dos <i>olas</i>, (1997 y 1998) manteniendo constante la muestra de 800 beneficiarias y usuarias de las actividades y servicios, mayores de 16 años (800 encuestas en cada ola). El tipo de diseño muestral dependerá de la estructura de población de usuarias/ beneficiarias que se detectará con el estudio 1.</p> <p>D) Estudio 3, que integra tres encuestas distintas:</p> <ul style="list-style-type: none"> - Encuesta sobre la situación social de la mujer en Andalucía 1998 - Objetivos del II PIOM - Actuaciones del II PIOM (2ª ola del estudio 2) <p>3.800 encuestas a mujeres de más de 16 años, residentes en Andalucía, en dos submuestras: 3.000 a mujeres en general y 800 a usuarias/ beneficiarias del Plan, la misma muestra que en el estudio 2.</p> <p>Muestra estratificada proporcional, incluyendo mujeres de 16 y 17 años.</p>

<p>Canarias I (95-96)</p>	<p>La metodología seguida fue la siguiente:</p> <ul style="list-style-type: none"> - Definición de objetivos - Localización de interlocutores en: ICM y Consejerías - Diseño de un programa de actuaciones: definición de un calendario; realización de fichas para la solicitud de información (se hace una genérica y una por Consejería. Más que ficha es un guión <i>a cumplimentar por cada Consejería</i>); inicio de contactos; recopilación de información, análisis y solicitud de nuevos datos; revisión de contactos, corrección de informaciones y búsqueda de una visión global por Consejería; análisis del cumplimiento de objetivos, por áreas; valoración. - Redacción del documento final
<p>Castilla-La Mancha II (95-99)</p>	<p>Sistema con dos ejes de recopilación de información:</p> <p>1. Recopilación de resultados de ejecución de las acciones del Plan, a través de:</p> <ul style="list-style-type: none"> - sistematización y ordenación de las memorias de resultados anuales de la DGM - Solicitudes de información y entrevistas a personal técnico y responsable de otros departamentos implicados en el desarrollo de las acciones. <p>Para esto se realiza un documento en el que se especifican las actividades, metas y objetivos en los que se suponía podían estar implicadas las diferentes unidades administrativas, con la finalidad de apoyar la recopilación. También un guión de entrevista y otras fichas para sistematizar la información recopilada.</p> <p>2. Opiniones sobre el diseño, el desarrollo y el impacto de este II Plan (de los responsables de departamentos implicados, asociaciones de mujeres y mujeres de corporaciones locales), a través de:</p> <ul style="list-style-type: none"> - entrevistas individuales a responsables y personal técnico de los centros que coordina la DGM y las Delegaciones Provinciales, y de otros depart. - sesiones de trabajo, en cada una de las provincias, con mujeres se colectivos y agentes sociales; una sesión de trabajo con mujeres que participan en las corporaciones locales de la región, mujeres alcaldesas y concejales.
<p>Castilla y León I (94-96)</p>	<p>Los datos disponibles para los análisis proceden de dos clases de fuentes:</p> <ul style="list-style-type: none"> - Fuentes secundarias: estadísticas que se buscan, reelaboran y analizan, en cualquier institución que ofrezca datos, que tengan que ver con el contenido del PIIOM y que se refieran a la Comunidad de Castilla y León. - Fuentes primarias (administrativas): La información facilitada por la propia Junta de Castilla y León, referida, generalmente, a sus propias actuaciones.

<p>Galicia II (95-97)</p>	<p>1. En primer lugar se creó una comisión interna de evaluación compuesta por las personas designadas por las Consejerías, así como personas responsables de otros organismos públicos de interés (RTVG, etc.)</p> <p>2. Se diseñó por parte del equipo evaluador los instrumentos o protocolos de recogida de información.</p> <p>3. Se adecuan o modifican estos instrumentos al objeto de adaptarlo a las necesidades específicas de la comisión.</p> <p>En un principio se redactó una ficha por cada tipo de actividad pero, debido a su complejidad, se condensó en una única ficha, para facilitar el manejo y tabulación de los datos, aunque en cada actividad se completaron solo los ítems que le correspondían o se adecuaban.</p> <p>En la ficha se contemplaba: área, programa y objetivos, actividad o acción, con una serie de parámetros: localización temporal y espacial, tipo de destinatarias, financiación, requisitos para su acceso o participación, y canales de difusión, y valoración de la acción.</p> <p>La recopilación de información se realizó por parte de las personas que formaban parte de la Comisión, designadas por cada una de las Consejerías.</p> <p>A medida que se reciben las fichas, se procede a analizar la documentación entregada completándola en su caso. Luego, se trasvasa la información en bruto con el fin de prepararla para su proceso informático.</p>
<p>Madrid II (93-95)</p>	<p>En la publicación de los resultados de la evaluación (“Las mujeres en la Comunidad de Madrid”) no se especifican en ningún momento los procedimientos utilizados para la obtención de los datos.</p> <p>En la fase de la elaboración del sistema de indicadores, se proponen los siguientes criterios para su selección: que sean pertinentes, sencillos, que permitan una evaluación continua y que sean comparables y específicos. Los procedimientos en esta fase son:</p> <p>1) Obtención, evaluación y, en su caso, adaptación de los indicadores ya existentes en el Estado y en instituciones internacionales.</p> <p>2) Diseño, en su caso, de aquellos indicadores para los no haya precedentes útiles, distinguiendo entre: a) aquellos que pueden ser alimentados con datos existentes en fuentes secundarias y b) aquellos que tendrán que ser alimentados con encuestas o análisis propios.</p> <p>La metodología que se propone para la segunda fase (Consultoría y asistencia para la evaluación del II Plan) debe ajustarse, según el pliego de condiciones técnicas, a los siguientes requisitos:</p> <ul style="list-style-type: none"> - comparabilidad entre los datos relativos a la CM y los correspondientes al Estado Español y a la Unión Europea - inclusión de todos los indicadores relativos a todos los objetivos y a todas las acciones del II Plan, susceptibles de evaluación - representatividad de los datos primarios relativos a las mujeres de la CM, en el caso de que se considere necesaria la obtención de dichos datos mediante estudios específicos, con un nivel de confianza del 95,5 % y con un margen de error inferior al 3%. <p>Para esto se utilizan:</p> <ol style="list-style-type: none"> 1. Datos de la propia Dirección General de la Mujer (Fuentes administrativas) 2. Fuentes primarias: realización de una encuesta 3. Fuentes secundarias: extraídas de diferentes fuentes, censos y estadísticas.

Como se puede observar en la tabla 6.5.a (Instrumentos de recopilación de información utilizados en las evaluaciones externas), los instrumentos más utilizados en esta segunda generación de evaluaciones son los diferentes **registros o fichas**, realizados *ad-hoc*, en los que se recogen informaciones sobre las actividades y actuaciones realizadas como parte del plan evaluado. En todas las evaluaciones estudiadas se utiliza este tipo de instrumento, aunque en algunos casos más sistematizado o estructurado que en otras. Por ejemplo, en Castilla- La Mancha y en Madrid no tenemos constancia de que existiera una ficha o registro concreto, sino que se sistematizaba la información aportada por las fuentes administrativas y encontrada en diversa documentación. Como se ha comentado antes, en este tipo de “sistemas de información”, tan importante como los instrumentos en sí, son los procedimientos y el sistema para la recopilación de información, ya que en la mayoría de los casos las fuentes de información se encuentran en departamentos diferentes a los propios organismos de igualdad. En los casos en los que se informa, explica y compromete a los otros organismos y departamentos, y son los propios organismos responsables los que recogen la información (aunque exista un proceso paralelo de seguimiento por parte del organismo de igualdad) existen mayores probabilidades de obtener una información de calidad (de mayor alcance y mayor fiabilidad). Esto ocurre, entre las evaluaciones externas, únicamente en el caso gallego. En el resto de los casos, la recopilación de información ha dependido sobre todo de las agencias externas que realizaban las evaluaciones (como es el caso de Andalucía, Canarias, las dos Castillas y Madrid).

Los **cuestionarios o encuestas** se utilizan en dos de las experiencias externas (Andalucía y Madrid). En los dos casos se realizan encuestas a la población femenina adulta en general. En el caso andaluz, al realizarse varias encuestas, algunas de ellas van dirigidas a una población más acotada, definida como “usuarias” del II Plan andaluz. Tanto para el II Plan andaluz como para el madrileño la encuesta versa sobre las actividades o actuaciones del plan y, en menor medida, sobre sus objetivos.

Las **entrevistas** personales se realizan, formalmente, en otros dos casos (Andalucía y Castilla- La Mancha). Los **grupos de discusión**, o más bien “sesiones de trabajo” (entrevistas colectivas) se hacen en Castilla- La Mancha. La utilización de **fuentes secundarias** como censos, estadísticas, etc. se da en las evaluaciones de Madrid y Castilla y León. Finalmente, en otros dos casos, se utilizan **otros instrumentos o procedimientos** como es una recopilación documental más exhaustiva, y los contactos informales mencionados (Andalucía y Castilla- La Mancha), además del juicio de expertos en la evaluación del II Plan andaluz y, en la misma experiencia, un estudio piloto previo respecto a tres acciones con el objeto de detectar las dificultades de obtención de información previamente a diseñar el sistema de información definitivo.

TABLA 6.5. a: INSTRUMENTOS DE RECOPIACION DE INFORMACION UTILIZADOS EVALUACIONES EXTERNAS

	Registros, fichas Sist. infor.	Cuestionarios, encuestas	Entrevistas personales	Grupos de discusión, sesiones de trabajo	Fuentes secundarias (censos, etc.)	Otros
Andalucía II (95-97)	X	X	X			Recopilación documental, contactos informales, juicios de expertos, estudio piloto
Canarias I (95-96)	X					
Castilla-La Mancha II (95-99)	X (sistemati- zación)		X	X		Recopilación documental, contactos informales
Castilla y León I (94-96)	X (fuentes admtivas.)				X	
Galicia II (95-97)	X					
Madrid II (93-95)	X (fuentes admtivas.)	X			X	

4.2. Procedimientos y metodología de las evaluaciones mixtas

En la tabla 6.4. b se señalan, para cada una de las tres experiencias mixtas, los procedimientos utilizados para la recopilación y el análisis de la información. En el caso de la evaluación cuantitativa realizada por el IM se especifican de forma profusa los procedimientos de muestreo y los de análisis estadístico, aunque esto no es así para los análisis de la información extraída a través de técnicas de tipo cualitativo ni a través de fichas o registros de información; en lo cualitativo, se suele especificar más sobre el instrumento y menos sobre los procedimientos posteriores para analizar la información recopilada a través de dichos instrumentos. Esta es, de todas formas, una tendencia generalizada en todas las experiencias de evaluación estudiadas.

TABLA 6.4. b: PROCEDIMIENTOS Y METODOLOGIA PARA LA RECOPIACION Y EL ANALISIS DE LA INFORMACION EVALUACIONES MIXTAS

	Procedimientos y técnicas de recopilación y análisis
Instituto Mujer II (93-95)	<p>Evaluación de la implementación: Se diseña una ficha para cada indicador (se hizo un listado de indicadores por cada actuación), que se cumplimenta al finalizar cada año de vigencia del Plan. Dichas fichas se cumplimentan por parte de las Unidades del IM responsables de la ejecución de las diferentes actuaciones del Plan. Vaciado de fichas y tratamiento de la información (Área de Estudios)</p> <p>Evaluación cualitativa: A) Entrevistas en profundidad: 11 entrevistas a responsables de la puesta en marcha de las políticas de igualdad en los Ministerios de Administraciones Públicas, Justicia e Interior, Trabajo, Educación y Ciencia, Industria y Sanidad. B) Grupos de discusión: 8 grupos constituidos por representantes de colectivos de mujeres. Participaron representantes de: institutos y seminarios universitarios de estudios de las mujeres; mujeres participantes en los cursos/programas formativos de alfabetización y acceso a la cultura; mujeres participantes en cursos promovidos desde la iniciativa NOW, para su promoción e inserción en el mundo laboral; mujeres pertenecientes a asociaciones de carácter social, educativo, político, feminista...; mujeres que ocupaban puestos de dirección o predirección, tanto en las administraciones o empresas públicas, como en la empresa privada; mujeres que hubieran tenido un hijo/a en el último año, habiendo sido asistidas por la sanidad pública; mujeres pertenecientes o colaboradoras de colectivos de mujeres marginadas.</p> <p>Evaluación cuantitativa: Diseño cuasi-experimental pre y post con grupo de control (el diseño tiene una medida pre y otra post con un grupo experimental -mujeres- y un grupo de control -hombres). Se realizan dos recogidas de datos, mediante la aplicación de cuestionarios <i>ad-hoc</i>, a una muestra representativa de la población española mayor de edad, en dos momentos diferentes (1993-1996). La selección de las 1.255 personas que compusieron sendas muestras, pertenecientes a todo el territorio español, se llevó a cabo mediante un muestreo de conglomerados polietápico, con estratificación de las unidades de primera etapa (secciones censales). Las unidades de la segunda etapa fueron las viviendas; dentro de las secciones, se seleccionaron mediante muestreo sistemático con arranque aleatorio. Las unidades muestrales de la tercera etapa (personas) se eligieron mediante un sistema de cuotas por sexo, edad, ocupación, según la composición de la sección respecto a estas variables. Tratamiento estadístico: razón diagnóstica (esterotipia de rasgos y valores); alfa de Cronbach, análisis de componentes principales y análisis de covarianza (estereotipia de roles y actitudes hacia la igualdad de géneros); estadísticas descriptivas -media y desviación típica- y técnicas multivariantes -regresión logística-(calidad de vida); medias, desviaciones típicas y %, análisis de varianza y covarianza (usos del tiempo); análisis discriminante e indicadores de igualdad.</p>

<p>Cataluña II (94-96)</p>	<p>El proceso de evaluación se estructura en las etapas siguientes:</p> <ol style="list-style-type: none"> 1. Determinación de los objetivos y metodología de evaluación. 2. Identificación de las acciones a evaluar y recopilación y sistematización de la información disponible por cada acción. 3. Análisis de la información recogida y elaboración de un primer documento sobre el grado de cumplimiento del II Plan por ámbitos. Presentación de las conclusiones obtenidas al ICD. 4. Búsqueda de información adicional y actualización de la base de datos a partir de la nueva información aportada por los departamentos sobre la actividad desarrollada en el 2º semestre de 1996. 5. Elaboración de un segundo documento, avance del informe de evaluación tanto globalmente como por ámbitos. Presentación al ICD. 6. Transmisión a los departamentos, para su contraste y validación, de la información sobre la que el equipo evaluador elabora el avance de informe. Esta información se presenta en forma de <i>Informe de implementación</i> específico para cada departamento. 7. Incorporación de la información adicional y de las observaciones que formulan los departamentos en el Informe de Evaluación que el ICD, en nombre del Gobierno de la Generalidad, presenta al Parlamento. <p>También se contrasta con el <i>Consell Nacional de Dones</i> de Cataluña.</p> <p>Fuentes de información: El Área de Seguimiento del Plan del ICD realiza desde 1994 una tarea de seguimiento de la puesta en práctica del plan, a través de una ficha para cada una de las acciones previstas. Con esta información se realizan los Informes de situación sobre el grado de ejecución de las acciones. Esto hace que el equipo evaluador pueda concentrarse en las tareas de análisis y evaluación propiamente dicha.</p> <p>También el equipo evaluador ha realizado numerosas consultas a la jefa del Área de Seguimiento del ICD.</p>
---------------------------------------	--

<p>País Vasco II (95-98)</p>	<p>En una primera fase, paralelamente a la elaboración del II PAPME, se definieron indicadores para cada una de las acciones propuestas (396). Se definen siete grandes categorías en las que se clasifican todas las acciones propuestas en el II PAPME, según el tipo de intervención, y se realiza un cuestionario diferenciado para cada uno de los 7 tipos de categoría de acciones (en la evaluación del I PAPME se había utilizado un único cuestionario que resultó complejo y poco sensible a la diversidad de las actuaciones propuestas)</p> <p>Los siete cuestionarios tienen dos partes:</p> <ol style="list-style-type: none"> 1. Una genérica, común para todos, compuesta por 11 items: descripción de la actividad realizada; acción del Plan con que se corresponde la actividad; grado de correspondencia entre la actividad realizada y la acción del Plan marcada (escala de 1 a 5); objetivos que se ha pretendido lograr con dicha actividad; fecha de inicio de la actividad; organismo que ha realizado la actividad; organismos colaboradores; tipo de colaboración; lugar/es donde se ha llevado a cabo la actividad; inversión económica; colectivos receptores de la actividad realizada. 2. Una segunda parte, con preguntas específicas para cada categoría y que posibilitan analizar la actuación con mayor profundidad y de forma más cualitativa (temáticas o problemáticas abordadas; tipo de actividad realizada - jornada, conferencia, cursos, materiales,...-; horas dedicadas a las diversas actividades; número de personas, edad y sexo, etc. <p>Tras la experiencia de la evaluación del I PAPME se incorpora el objetivo de aumentar el grado de implicación y participación de los organismos públicos en la evaluación. Para esto, se delimitó con mayor exactitud las Áreas y Subáreas de actuación, objetivos y acciones en cuya ejecución está implicado cada uno de los organismos de la Administración Pública Vasca y se diseñaron sendos documentos diferenciados, para cada uno de los 21 Organismos de la Administración general (Departamentos de Gobierno, Organismos Autónomos, Sociedades Públicas), uno para Diputaciones Forales y uno para Ayuntamientos. Estos instrumentos se estructuran en tres partes:</p> <ol style="list-style-type: none"> 1. Las acciones e indicadores concernientes a cada organismo 2. Las acciones e indicadores agrupados según las siete categorías 3. Los instrumentos de evaluación correspondientes <p>Se organizaron reuniones para entregar y explicar los objetivos y contenidos de los materiales elaborados, en forma de borrador y para que pudieran proponer los cambios pertinentes. Cada Departamento nombró una persona responsable de la coordinación entre <i>Emakunde</i>, los departamentos de Gobierno y sus respectivos Organismos Autónomos, y se realizaron posteriores reuniones individualizadas. En el caso de las Diputaciones Forales, la estrategia fue similar (reuniones individualizadas con algunos responsables políticos y posteriormente informativas con personal técnico). Con los Ayuntamientos, se realizaron reuniones en los de más de 2.000 habitantes. En el resto se les envió el documento por correo, junto con un escrito explicativo del contenido, y objetivos del mismo.</p>
---	---

En los casos del **II Plan catalán** y el **II PAPME vasco**, se entiende la metodología y los procedimientos empleados de una forma más amplia y se especifican en los informes los pasos dados y las estructuras de coordinación utilizadas para la recopilación de información, como parte de dichos procedimientos. Además, en estos dos casos (y de forma similar a lo que ocurre en el caso gallego) se crean ciertas estructuras de documentos y de reuniones específicas con los diferentes departamentos gubernamentales, que son las que guían, en cierto modo, el proceso de recopilación (en mayor medida) y análisis (en menor medida) de la información.

Como se puede observar en la tabla 6.5. b, y al igual que en las evaluaciones externas, los instrumentos más utilizados son los **registros o fichas**. De hecho, estos instrumentos son la base de las evaluaciones catalana y vasca y de la evaluación de la implementación realizada por el IM. El grado de complejidad, cantidad y calidad de la información recopilada varía, siendo el sistema más sofisticado y elaborado y el que recopila una mayor cantidad de información ajustada a las necesidades del proceso de evaluación, el del País Vasco⁸. El siguiente en complejidad y alcance sería el catalán y, muy por debajo, el utilizado por el IM. En el caso de la evaluación del II PAPME vasco y también del II plan catalán se informa, explica y compromete a los otros organismos y departamentos, y son los propios organismos responsables los que recogen la información (aunque exista un proceso paralelo de seguimiento por parte del organismo de igualdad), cuestión que, como se ha visto en el caso de las evaluaciones externas, ofrece una información más fiable. No es así en el caso del IM en el que se pide información a los diferentes ministerios pero ésta es procesada y sistematizada por el propio IM.

En los casos catalán y vasco no se utilizan prácticamente otros instrumentos de recopilación de información, aunque se contrasta la información con entrevistas informales y visitas a los responsables de otras unidades gubernamentales. En el caso del IM sí se utilizan otros instrumentos de recopilación y análisis de la información. La encuesta que se realiza a nivel nacional se aplica tanto en mujeres (grupo experimental) como en hombres (grupo de control) y ésta versa sobre una serie de variables relacionadas con actitudes, valores, estereotipos y usos del tiempo, y no tanto sobre actividades o servicios del plan.

La evaluación del II Plan nacional es la única que incluye formalmente grupos de discusión. No se utilizan fuentes secundarias en ninguna de las tres experiencias mixtas.

TABLA 6.5. b: INSTRUMENTOS DE RECOPIACION DE INFORMACION UTILIZADOS

⁸ En el caso de la evaluación del II PAPME, a los instrumentos utilizados para recopilar la información se les denomina en alguna ocasión cuestionarios, ya que en realidad son preguntas que se realizan a los diferentes organismos y departamentos de la Administraciones Públicas Vascas. Sin embargo, aquí hemos preferido considerarlo como registros de información que deben “rellenar” los otros departamentos, más que como cuestionarios o encuestas, ya que se asemejan más a aquéllos que a éstos.

EVALUACIONES MIXTAS

	Registros, fichas Sistemas de información	Cuestiona- rios, encuestas	Entrevistas personales	Grupos de discusión, sesiones de trabajo	Fuentes secundarias (censos, etc.)	Otros
Instituto Mujer II (93-95)	X	X	X	X		
Cataluña II (94-96)	X					Recopilación documental, contactos informales
País Vasco II (95-98)	X (*)					Visitas a las administracio- nes públicas vascas

(*) En *Emakunde* a estos instrumentos se les denomina en alguna ocasión cuestionarios, ya que en realidad son preguntas que se realizan a los diferentes organismos y departamentos de la Administraciones Públicas Vascas. Sin embargo, aquí los consideramos registros de información que deben “rellenar” los otros departamentos, más que como cuestionarios o encuestas, ya que, en una consideración general, se asemejan más a aquellos que a éstos.

4.3. Procedimientos y metodología en las evaluaciones internas

En la tabla 6.4.c se reflejan los procedimientos utilizados para la recopilación y el análisis de la información en las dos experiencias de evaluación clasificadas como internas. En ambos casos, la evaluación se ha llevado exclusivamente a través de registros o fichas realizadas “ad-hoc”, y con una sistema que implicaba un fuerte protagonismo del organismo responsable en la recopilación, el procesamiento y la sistematización de la información.

TABLA 6.4. c.: PROCEDIMIENTOS Y METODOLOGIA PARA LA RECOPIACION Y EL ANALISIS DE LA INFORMACION EVALUACIONES INTERNAS

	Procedimientos y técnicas de recopilación y análisis
Aragón I (94-96)	Ficha para cada acción, en la que se recogía datos de identificación de la de la acción, organismos y entidades implicadas, tipo de participación, temporalización, ámbito territorial de actuación, tipo de usuarias/as y presupuestos. La información se recogió por parte del propio IAM.
La Rioja II (96-99)	Se elaboran los siguientes documentos e instrumentos: <ul style="list-style-type: none"> - Listado de indicadores de evaluación: para cada área, objetivo y actuación del Plan Integral. - Ficha de temporalización de la ejecución de las medidas del Plan. En ella se reseña el organismo implicado el área, objetivo y medida evaluada, cuando fue iniciada, cuando finalizada y si hay gasto previsto para los siguientes ejercicios. - Ficha de evaluación de medidas y acciones, una para cada medida (reflejando objetivo y área, organismo implicado), con espacio para especificar los indicadores correspondientes para dicha medida. En esta ficha se consigna el año, las actividades en relación a la medida, los organismos participantes, el gasto ejecutado y el grado de ejecución. - Criterios a seguir para cumplimentación de los documentos de seguimiento y evaluación de los planes integrales de personas con discapacidad y de mujer (se elabora el mismo sistema para ambos planes).

CAPÍTULO 7

LA UTILIZACIÓN DE LAS EVALUACIONES: PERTINENCIA, COMUNICACIÓN E IMPACTO

INTRODUCCIÓN

En los dos capítulos anteriores se han descrito las experiencias de evaluación metaevaluadas, respecto a dos dimensiones: la primera, referida al diseño y conceptualización de los procesos de evaluación, trató el planteamiento de dichos procesos y las estrategias evaluativas seguidas para la evaluación de los planes; la segunda dimensión, referida al proceso de puesta en práctica de la evaluación, señaló los elementos clave de los procesos evaluativos metaevaluados (agentes, momentos, criterios y metodología seguida). Este capítulo discute sobre la tercera dimensión que, referida a los resultados esperables de las evaluaciones, trata el tema de la **utilización** de las mismas.

Por lo tanto, se trata de recabar información sobre la utilización y sobre la utilidad de los procesos de evaluación para las políticas de igualdad de género. En este capítulo se analiza, en primer lugar, la **pertinencia y utilidad de la información producida**, tanto para el plan evaluado, como para el organismo de igualdad protagonista de la evaluación, y para la situación evaluativa concreta. Para esto se tuvo en cuenta el alcance de la información (si se ha recopilado, tratado y analizado la información suficiente) así como su globalidad (si se contempla globalmente el plan o sólo aspectos concretos del mismo). En evaluación, además de la recopilación y el análisis sistemático de información, existen dos niveles más de análisis que son, por un lado, la emisión de juicios de valor que es lo que le otorga su carácter valorativo y el establecimiento de recomendaciones que son una concreción de su carácter aplicado y dirigido a la acción. Por lo tanto, la pertinencia y utilidad de la información producida por una evaluación tiene también con el grado de elaboración o “traducción” práctica de dicha información, es decir, con el grado de análisis y elaboración de conclusiones y la existencia y pertinencia de las recomendaciones. En definitiva, por muy buena que sea una información recogida y analizada, si no se ha dado un paso más en su valoración global y en su “traducción” en recomendaciones concretas para la política evaluada, su pertinencia y utilidad será mucho menor.

En segundo lugar, se analizan los procesos de **comunicación de los resultados** de las evaluaciones, como otro elemento clave para la utilidad y la utilización de las mismas. El principal elemento de comunicación de una evaluación es su informe. De esta manera, se estudiaron los informes de evaluación producidos por cada experiencia. Por un lado, se analizaron sus contenidos, poniendo especial énfasis en el grado en el que reflejaban los propios procesos de evaluación llevados a cabo además de los resultados de los mismos. Por otro, cuán completos eran dichos informes, partiendo de

la base de que un buen informe de evaluación debe describir, además de los resultados, los propósitos del estudio, el contexto del plan evaluado, la metodología de evaluación, la estrategia evaluativa general, el modo en el que se han establecido los criterios de evaluación, los propios criterios de evaluación, las estructuras y foros que se han utilizado para la recopilación y cómo se han establecido las valoraciones y recomendaciones.

Por último, se trata en este capítulo del **impacto que los estudios de evaluación** tienen tanto en los organismos de igualdad y en su función de evaluación como en la propia política. El concepto de impacto de las evaluaciones está inextricablemente unido al de uso; el impacto de las evaluaciones depende del uso que se haga de las mismas. En este apartado se realiza una valoración de los cambios que se produjeron por la propia existencia de los procesos de evaluación estudiados, siendo conscientes de las inevitables limitaciones encontradas para recopilar información de suficiente alcance y calidad sobre este extremo. En general, además de analizar los posibles usos de los resultados de las evaluaciones, también se estudiaron la potencial utilización de los propios procesos de evaluación, pudiendo servir éstos para introducir una cultura y una metodología de evaluación en los organismos de igualdad y para generar una mayor comprensión de las políticas de igualdad y de cómo abordar el problema de la desigualdad de género.

También en este capítulo se utiliza la clasificación en evaluaciones externas, mixtas e internas para la exposición de los datos.

1. PERTINENCIA Y UTILIDAD DE LA INFORMACIÓN PRODUCIDA EN LAS EVALUACIONES EXTERNAS

En este apartado se realizan dos tipos de análisis: en primer lugar, la **pertinencia de los contenidos del estudio** para el plan evaluado, para el organismo de igualdad protagonista y para la situación evaluativa concreta. Para esto se tiene en cuenta el alcance de la información (es decir, si se ha recopilado, tratado y analizado la información suficiente) y la globalidad de la misma (es decir, si permite hacer una valoración global del plan o sólo de aspectos concretos sobre el mismo). También se tienen en cuenta las opiniones vertidas por las responsables políticas entrevistadas sobre este aspecto.

En segundo lugar, se estudia el **grado de análisis y elaboración de conclusiones y la existencia y la calidad de recomendaciones en los informes de evaluación**, como resultado de la misma: en definitiva, se analiza el **grado de “traducción práctica” y aplicabilidad de los procesos de evaluación**, como condición facilitadora de su utilidad.

1.1. Pertinencia y utilidad de la información producida en las evaluaciones externas

El planteamiento de la evaluación del **II PIOMA andaluz** es el más completo respecto a los contenidos, es decir respecto a lo que se evalúa, tanto en alcance como en globalidad. Dicho estudio de evaluación pretende cubrir todo el espectro de fases del plan, planteando una evaluación de la conceptualización y el diseño del plan (es la única experiencia junto con Castilla- La Mancha que se plantea formalmente este tipo de evaluación), una evaluación de la ejecución y una evaluación de la eficacia e impacto. Esta visión global del Plan, planteada por el estudio de evaluación, es en principio más útil que una visión parcializada del mismo, ya que cubre todos los aspectos posibles: cuestiones de diseño, de funcionamiento o ejecución y de resultados; por lo tanto, tiene mayores probabilidades de ofrecer información pertinente para diferentes necesidades de información y para diferentes decisiones. Respecto al diseño, se estudian los propios objetivos del plan (coherencia, racionalidad, orden, jerarquía, ausencia de solapamientos), la congruencia respecto a los indicadores planteados y la validez aparente, fiabilidad y globalidad de éstos y ajuste del Plan a las necesidades reales. Este tipo de evaluación permite, al menos potencialmente, una reflexión seria sobre la definición del problema de la igualdad¹ y sobre cuál debe ser el papel y los objetivos de

¹ Siendo muy interesante el análisis que se realiza sobre qué tipo de igualdad se pretende con el plan (no sólo igualdad formal sino de hecho; algo que va más allá de la igualdad de oportunidades, pretendiendo

la evaluación en los planes de igualdad. Respecto a la ejecución, se plantean, además de la típica evaluación de “productos” de prácticamente todas las experiencias de evaluación estudiadas, una verdadera evaluación de proceso, teniendo en cuenta el entorno institucional del Plan (si se tienen en cuenta los objetivos de las organizaciones involucradas), los trámites de acceso a las actividades y el tiempo que tarda la institución en dar respuesta. Es decir, aunque no sea completa, es en la única experiencia estudiada que se evalúa verdaderamente la fase de la implementación o ejecución (junto con Castilla-La Mancha, que lo hace sólo tangencialmente y a través de las opiniones de personas implicadas en el Plan), estableciendo criterios de evaluación que tienen que ver con el funcionamiento y puesta en marcha del plan, y no sólo con lo que se realiza o produce de forma finalista. Finalmente, y en lo que se refiere a la evaluación de la eficacia e impacto, es la única experiencia analizada que, en el proyecto inicial intenta realmente valorar los “efectos netos” del Plan y propone los instrumentos metodológicos para ello, haciendo una serie de encuestas en series temporales distintas (tres estudios integrados dirigidos a población general y a usuarias/beneficiarias, en diferentes tiempos y “olas”, más el estudio de la situación social de la mujer andaluza realizado en 1988²) que permiten un cierto control de variables para aislar dichos “efectos netos”. Esto marca una diferencia con el resto de experiencias que pretenden una evaluación de impacto o efectos (II Plan nacional, I castellano leonés y II madrileño), en las que, aunque existe una cierta medición temporal o de evolución (mediciones de antes después o indicadores extraídos de datos en tiempos distintos), los datos se recogen de la población general y no existe ningún control de variables que asegure la medición de los efectos netos del programa. Es decir, se mide la evolución de la situación social de la mujer en determinadas variables en general, no la parte de evolución de los datos directamente achacable a los planes. Sin embargo, en el informe final de evaluación publicado, la evaluación de impacto queda

además de la igualdad ante la ley, la igualdad de recursos, la igualdad de bienestar y la igualdad política), el análisis que se realiza sobre el poco poder discriminante del plan por tratar colectivamente a las mujeres, y la idea de que el Plan debería haberse planteado también la igualdad no sólo respecto a los hombres, sino *entre las mujeres* (p. 21 del “Avance de informe”), nos parece, cuando menos, algo ingenua y creemos que denota, en cierto modo, la falta de conocimiento y especialización del equipo evaluador en temas de género. Efectivamente, el *leit-motiv* de las políticas de igualdad de género no es la igualdad de oportunidades entre todos los ciudadanos, objetivo muy loable en una sociedad democrática, pero que excede con mucho los objetivos de estas políticas, sino de un determinado tipo de desigualdad que es la que se explica por el género, a lo largo y ancho de todas las situaciones y clases sociales. El objetivo de la igualdad de género tiene que ver con conseguir que la variable sexo no explique tantas diferencias significativas en toda una multiplicidad de sectores y situaciones sociales. Creemos que la igualdad *entre* las mujeres tampoco tendría mucho sentido planteársela en aislado, es decir, sin plantearse paralelamente la igualdad *entre* los hombres. Por supuesto esto no significa que a la hora de formular las políticas y programas no se tengan en cuenta las situaciones sociales específicas de la población a la que van dirigidas (insistimos de nuevo en que las destinatarias de los planes de igualdad no deben ser exclusivamente las mujeres; pueden y deben ser también los hombres), no se estudien y asuman las interacciones e interdependencia de determinadas desigualdades y problemas sociales (sin caer en la trampa de estereotipar algunos problemas como, por ejemplo, equiparar la violencia contra las mujeres exclusivamente a situaciones de marginación, obviando el hecho de que la violencia contra las mujeres se da en todos los estratos sociales), e incluso que no se pueda y deba hacer un estudio de las desigualdades sociales desde la perspectiva de género, como ya se hace en torno a los estudios feministas o *women studies* (feminización de la pobreza; la doble marginación de género y raza, etc.). Pero de ahí a plantearse que las políticas de igualdad de género, por el hecho de plantearse las desigualdades debidas al género deben necesariamente también plantearse cualquier otro tipo de desigualdad entre las mujeres, nos parece errar de alguna forma en el planteamiento y distraer del objetivo fundamental para el que se concibieron.

² Véase tabla 3.9.

reducida al análisis del grado de satisfacción de las usuarias, dejando de lado muchas de las cuestiones inicialmente planteadas en el proyecto de evaluación.

Respecto a la “traducción práctica” de la evaluación del II Plan andaluz es poca en los primeros estadios del proceso de evaluación. En el “Avance de informe de evaluación” presentado por el IESA-A en julio de 1997, no aparecen ningunas conclusiones o recomendaciones concretas respecto a la evaluación del diseño que allí se presenta; y realmente se podría haber hecho. Los otros dos tipos de evaluación - implementación y resultados e impacto- se presentan todavía en proyecto, por lo que es normal que no existan recomendaciones sobre las mismas, aunque en el proyecto de evaluación de la implementación se analiza, a través del estudio piloto de dos acciones, la necesidad de un sistema de información; sin embargo, creemos que la realización de ese estudio piloto podría haber generado una más clara traducción práctica que hubiera revertido, además, en una mayor claridad y transparencia del proceso de evaluación.

En el informe final de evaluación publicado, en el que se hace una síntesis de todo el proceso de evaluación, se exponen los resultados de los diferentes estudios, organizados según las áreas del plan. Después de la exposición de los resultados en cada una de las áreas, se señalan unas conclusiones a modo de resumen. Este apartado resume los resultados y hace algunas valoraciones, pero en ningún caso se establecen recomendaciones.

La valoración general del proceso de evaluación que se realiza por parte de las responsables internas del IAM no es muy positiva. Para ellas, la evaluación se ha quedado en una fase esencialmente descriptiva y opinan que no se ha explotado casi el cuestionario realizado. Esto en parte lo achacan a los cambios que se han dado en el equipo externo a lo largo del proceso.

La evaluación del **I Plan Canario** es muy limitada en cuanto a los contenidos evaluados y a su carácter global, es decir, su alcance es muy pequeño. En este caso se realiza externamente ya que, en aquel momento, la estructura del Instituto Canario de la Mujer era muy escasa, prácticamente inexistente. Ya se ha visto que la evaluación se limita a una descripción de las actividades realizadas al amparo de este I Plan. Así también, la evaluación sirve como una forma de descripción de lo realizado hasta el momento, aunque la valoración del grado de utilidad por parte de las responsables del ICM no es positiva en este caso (desde el propio ICM se comenta que el resultado de este estudio ha sido “francamente pobre”). Es posible que al ser una evaluación externa, se tuvieran grandes expectativas sobre la misma que posteriormente no se vieron satisfechas.

El informe de evaluación (autodenominado “Auditoría”) pretende una cierta “traducción práctica”, ya que existe un apartado de conclusiones en el que se ofrecen, además, algunas sugerencias. Además, se ofrecen conclusiones generales y por cada una de las áreas. Sin embargo, tanto las conclusiones como las sugerencias son muy generales y escasas (en realidad se circunscriben a sugerir seguir y aumentar la política interdepartamental, los flujos de comunicación con otros departamentos e incrementar la difusión del Plan y del ICM). Es decir, además de su generalidad, las dimensiones son

muy escasas ya que el apartado de conclusiones ocupa un 0,03% (8 de las 265 páginas) de la extensión del informe final.

Aunque la evaluación del **II Plan castellano-manchego** es algo más completa, desde el punto de vista de los contenidos, tanto en alcance como, sobre todo, en el carácter global que se pretende, ya que se evalúa diseño del plan, el grado de ejecución (desde el punto de vista utilizado en otras experiencias de evaluación de *outputs*), y la valoración que del Plan hacen responsables y colectivos de mujeres respecto al diseño, al desarrollo y a los resultados e impacto del plan. Sin embargo, el grado de “traducción práctica” no es muy alto, ya que no se realizan conclusiones generales que integren los tres tipos de evaluación (diseño, resultados y valoración), ni se establece recomendación alguna. Únicamente, se recoge un apartado, como parte del tercer tipo de evaluación (valoración cualitativa), titulado “propuestas para el diseño y desarrollo de futuras políticas de igualdad de oportunidades” y en el que se reflejan las propuestas de actuación para el III Plan de Igualdad que fueron surgiendo en las sesiones de evaluación (pp. 81-84 del “Informe previo de evaluación del II Plan de Igualdad de Oportunidades ‘Las Mujeres de Castilla-La Mancha hacia el siglo XXI’”); pero, en este caso, se recogen las propuestas aportadas en los grupos de discusión y entrevistas en profundidad como un dato más, no como el cuarto nivel de análisis de la evaluación surgido del análisis y enjuiciamiento global de toda la información recopilada a lo largo del proceso evaluativo.

La evaluación del **I Plan castellano leonés**, como ya se ha visto con anterioridad, estudia variables relacionadas indirectamente con el plan que evalúa y que tienen más que ver con un estudio de la situación social de la mujer en la región, en una serie de aspectos. Temas que salen en el discurso de su responsable política y que parecen especialmente pertinentes para el aprendizaje sobre la propia intervención y para el propio rendimiento de cuentas, como es la coordinación intragubernamental (entre las diferentes Consejerías que deberían estar implicadas) y la coordinación intergubernamental (coordinación con las administraciones locales), no se tocan de ninguna manera en el estudio de evaluación. De hecho, ni siquiera se recoge ninguna información que ofrezca una visión (global o parcial) de lo que se ha hecho bajo el paraguas del Plan que supuestamente se está evaluando (lo que normalmente se denomina en las otras experiencias el “grado de cumplimiento”).

El estudio, además, está limitado, en su práctica totalidad, al análisis de datos secundarios, aunque también se contemplan también “fuentes primarias administrativas” que consisten en la información facilitada por la propia Junta de Castilla y León, referida a sus propias actuaciones, pero éstas son muy limitadas (acciones culturales, asociaciones subvencionadas y actos en los que está presente la Secretaría Regional de la Mujer). Es decir, la información es escasamente pertinente para la evaluación del Plan, aunque haga un análisis (limitado a la información disponible en las fuentes secundarias y administrativas) de algunos cambios acaecidos en la situación de las mujeres de Castilla y León durante el período de vigencia del Plan. La “traducción práctica” es asimismo nula, ya que en el informe de evaluación (“Resumen de la

Evaluación del Plan Integral de Igualdad de Oportunidades para la Mujer en Castilla y León”) no existe ningún apartado en el se contemplen conclusiones generales ni recomendaciones para la mejora del Plan, el control de responsabilidades o la formulación de políticas futuras.

La evaluación del **II Plan gallego** establece unos criterios muy completos para una buena evaluación de los *outputs* del mismo: cumplimiento de acciones según áreas, carácter de las actuaciones, contenido, ámbitos de impacto, destinatarios, participación institucional, ámbito geográfico, financiación. A esto añade datos de evolución de las actividades por años y sobre la difusión e información de actividades, que permite una cierta visión de proceso y alguna información sobre las valoraciones que hacen los responsables de las actuaciones sobre las mismas. Las dificultades en la obtención de información en este proceso evaluativo, a pesar de la creación de una comisión de evaluación *ad-hoc*, son bastante grandes y, por lo tanto, la cantidad y calidad de información recopilada no es tan alta como en otros casos (por ejemplo Cataluña y País Vasco). Sin embargo, se podría decir que la pertinencia es alta.

En esta evaluación, existe un cierto grado de “traducción práctica”, ya que existe un apartado de “Síntesis y conclusiones” en el informe final, aunque éste sólo represente el 4% del mismo (5 de 136 págs.). En este apartado se realizan algunas conclusiones sobre el diseño del plan, relacionadas con la importancia relativa de los objetivos y áreas del plan y las prioridades del plan, y sobre los resultados del mismo, es decir sobre las actividades resultantes: el incremento de actividades del Plan, sus destinatarios, su distribución por niveles gubernamentales, la financiación y los canales de difusión de las mismas. Las conclusiones son generales y no se plantea ningún apartado de recomendaciones, aunque en la redacción de las conclusiones se deja entrever una única recomendación, no relacionada con el plan sino con la evaluación del siguiente plan, en la que se recomienda se realice un esfuerzo de colaboración para lograr conocer el coste medio de las actividades puestas en funcionamiento para el cumplimiento de los objetivos del III Plan. Esta única recomendación es muy general ya que tampoco se sugiere cómo hacerlo.

Respecto a la evaluación del **II Plan madrileño**, y al igual que ocurre con la del castellano-leonés, se puede decir que se centra en variables relacionadas con el Plan de forma indirecta, aunque en este caso la información es más completa ya que, además del análisis secundario se hace también un análisis primario a través de una encuesta y un estudio más sistemático de las fuentes administrativas de la propia Dirección General de la Mujer. Es decir, el alcance de la información recopilada es alto (la cantidad y calidad de la misma es alta), pero la pertinencia de la misma, aunque alta para un estudio sobre la situación de la mujer madrileña aplicado a las políticas de igualdad, es muy baja para la evaluación del Plan. El hecho de centrarse exclusivamente en los cambios generales producidos en las mujeres madrileñas durante el período de vigencia del Plan es, además, contrario a lo que se requería inicialmente en el pliego de condiciones para el concurso público para la “Consultoría y Asistencia para la evaluación del II Plan de Igualdad de Oportunidades para las Mujeres de la Comunidad de Madrid” (Resolución

de 20 de junio de 1996 de la Secretaría General Técnica de la Consejería de Sanidad y Servicios Sociales. BOCM 28/6/96), en el que se especificaba que la empresa adjudicataria debía presentar, entre otras cosas, “las políticas alternativas que cabe llevar a cabo en los diferentes ámbitos de actuación a la vista de los resultados obtenidos mediante la aplicación del II Plan de Igualdad de Oportunidades”. En ninguna parte del estudio aparecen ni análisis globales de los diferentes datos recopilados ni valoraciones sobre la información recogida y analizada que permitan esta “traducción” hacia posibilidades futuras.

Por lo tanto, la “traducción práctica” en este caso no sólo es inexistente, sino que en la publicación de los resultados de la evaluación del II Plan que hace la Dirección General de la Mujer se hace alarde de ello. Así, en la presentación de dicha publicación se dice: “Se ha procurado, en fin, que la lectora o el lector sean quienes saquen sus propias conclusiones. Para ello, este libro ofrece la información necesaria para conocer los cambios que se están produciendo en la vida de las mujeres, procurando excluir los juicios de valor” (énfasis mío. p.3 en “Las Mujeres en la Comunidad de Madrid. Datos para abrir caminos hacia la igualdad”).

1.2. Pertinencia y utilidad de la información producida en las evaluaciones mixtas

Como ya se ha visto con anterioridad, la evaluación del **II PIOM estatal** se divide a su vez en tres estudios con poca conexión entre sí. En este caso, el alcance de la *evaluación de la implementación* no es muy limitado, ya que sólo se contemplan, y de forma individual para cada acción, su descripción, temporalización, grado de desarrollo y presupuesto invertido. Es decir, no se recoge agentes, ni destinatarios, ni ámbito geográfico, ni se realiza ningún tipo de clasificación de las actuaciones. El análisis, por lo tanto, es muy simple, limitándose a describir las actuaciones realizadas en cada una de las áreas. Esto supone un cierto control, ya que se conocen las actuaciones realizadas y las que no se han realizado, pero no aporta información ni sobre el diseño del plan, ni sobre el proceso de puesta en práctica del plan en su conjunto ni de las actuaciones en particular, ni tampoco aporta ninguna información sobre la calidad de las actuaciones. También la *evaluación cualitativa* recoge un espectro acotado de información, ya que se circunscribe a la opinión y valoración de personas y colectivos en torno al PIOM (colectivos de mujeres y responsables de ministerios); sin embargo, el alcance de la información respecto a esta dimensión -opinión sobre el PIOM- es importante ya que se realizan 11 entrevistas en profundidad y 8 grupos de discusión y con una cuidada selección. La *evaluación cuantitativa*, aunque muy cuidada en sus procedimientos técnicos y metodológicos, también se ocupa de aspectos o dimensiones que, aunque muy significativos respecto al posible avance en la igualdad de género (estereotipos sexuales, valores, actitudes, calidad de vida y usos del tiempo), representan sólo una parte acotada de los posibles efectos que pretendía el II PIOM. Es decir, el II PIOM pretendía muchas más cosas, además del cambio de actitudes ante la igualdad; y aunque la justificación de este estudio se halla en la idea de que los aspectos o dimensiones

estudiadas son las que subyacen y justifican los posibles cambios estructurales que hayan podido tener lugar, creemos que obviar la evaluación de dichos cambios (que son muchos y muy variados), contribuye a tener una idea muy parcial de lo que ha sido el II PIOM.

Respecto a la “traducción práctica” de la evaluación del II PIOM estatal, ésta es prácticamente inexistente. En la “Síntesis de resultados de la evaluación del II PIOM”, no se realiza ningún análisis global de los tres tipos de evaluación realizados, y ni siquiera de forma individualizada (cada tipo de evaluación por separado), existen conclusiones generales. Lógicamente, tampoco aparece ninguna recomendación. Esto no es exclusivo de la “síntesis” publicada. Tampoco en los informes originales aparece. Es decir, los informes se limitan a exponer los resultados obtenidos y sus interpretaciones, evitando realizar valoraciones y recomendaciones sobre los mismos.

La evaluación del **II Plan catalán** se centra, sobre todo, en una evaluación de productos o *outputs*, aunque el alcance de la información es amplio; incluso, como se ha visto antes, se establecen categorías sobre el tipo de política aplicado, tipo de actuación, etc., que permite también tener una visión global del plan y una cierta evaluación del diseño. La cantidad y calidad de la información recopilada es, además, amplia, por lo que podríamos decir que la pertinencia de la información es alta. Así, mucha de la información recopilada puede ser útil para la estructuración de planes posteriores.

Respecto a la “traducción práctica”, aunque en el informe final no existe ningún apartado de conclusiones globales ni de recomendaciones, en la exposición de los resultados, que se establece por ámbitos, sí existe un último punto de valoración global sobre el cumplimiento del plan en cada uno de esos siete ámbitos (sensibilización, legislativo, sociocultural, coeducación, formación ocupacional, profesional y trabajo, sanitario y de los servicios sociales). La valoración global que se realiza en cada uno de estos apartados contiene, además de las valoraciones generales, algunas sugerencias o recomendaciones para la acción. Así, por ejemplo, respecto al ámbito de sensibilización, una vez hechas las valoraciones (grado de implementación alto, descompensación de esfuerzos en unas pocas acciones, predominio de actuaciones específicas y de igualdad, mayoría de destinatarios en la propia administración, reducido número de Departamentos que han participado, estructuración del ámbito desigual y heterogénea), se establecen sugerencias bien ponderadas y establecidas en base a las valoraciones realizadas (dada la potencial amplitud que puede tener el propósito de sensibilización, se sugiere una reflexión más amplia y profunda sobre las temáticas a abordar y los destinatarios prioritarios, además del tipo de instrumentos a utilizar - política de comunicación de la Generalidad, la formación de colectivos en el ámbito público y las campañas específicas-, estableciendo recomendaciones concretas sobre qué combinación de instrumentos utilizar) pp. 60-79 del “Informe d’avaluació del II Pla d’actuació del Govern per a la igualtat d’oportunitats per a les dones”.

Finalmente, la evaluación del **II Plan vasco** representa una excelente y exhaustiva recopilación y análisis de información sobre los *outputs* del Plan. El esfuerzo de categorización de las actividades y los destinatarios ofrece una buena base para una

cierta evaluación del diseño del plan y para evaluar el grado de ejecución del Plan en su conjunto, y la recopilación sistemática y continuada de información, además de la realización de memorias anuales, permite también una evaluación del proceso. Además del esfuerzo y reflexión sobre los criterios de evaluación (qué se quiere evaluar) han elaborado y depurado mucho los instrumentos de recopilación de información, a lo largo de su experiencia evaluativa. Y tanto el alcance de la información como la calidad de la misma es alta y ha ido mejorando en los últimos años. El grado de pertinencia, por lo tanto, es alto.

Respecto a la “traducción práctica”, las memorias anuales de actuación de los poderes públicos que se presentan al Parlamento Vasco, contienen una sección con conclusiones en la que, se concluye y valora no sólo los resultados del proceso de evaluación sino también sobre la propia metodología de evaluación. Este es el único caso, entre las once experiencias analizadas, en el que ocurre esta atención al propio proceso evaluativo con fines de mejora de procesos futuros. En el punto en el que se hacen las conclusiones sobre los resultados, se hace un balance de toda la información detalladamente presentada en las secciones anteriores. Este balance ayuda mucho a tener una visión global e integral del proceso de evaluación, cuestión de la que carecen la mayoría de las restantes experiencias analizadas.

1.3. Pertinencia y utilidad de la información producida en las evaluaciones internas

En el caso de la experiencia de evaluación del **I Plan aragonés**, el proceso de evaluación es también muy limitado respecto a los contenidos, tanto en el alcance de la información recopilada como en la globalidad de la información. A pesar de esto, la evaluación sirve, en aquel momento concreto, para saber qué se había hecho hasta entonces, con una intención claramente descriptiva, y ante una situación de cambio de responsable y de vacío de poder durante un período relativamente largo (5 meses). Es decir, aunque fuera una evaluación muy simple y meramente descriptiva, parece que resultó útil para que las nuevas personas responsables se situaran en el contexto, saber qué se había hecho y qué no, y tomar decisiones concretas sobre qué hacer en el último período de vigencia del Plan. Es decir, en cierto modo, la evaluación, y a pesar de sus limitaciones, resulta pertinente para la situación evaluativa concreta de cambio de responsables y ausencia de información y control sobre el Plan.

Sin embargo, la “traducción práctica” del proceso evaluativo es muy baja, ya que el “Informe sobre el grado de cumplimiento del Plan” no contiene ninguna conclusión o recomendación clara: ni siquiera realiza juicios de valor, sino simplemente se limita a describir y explicar las actuaciones realizadas.

La información que se pretende recopilar para la evaluación del **II Plan riojano** se limita a una serie de indicadores, que permiten su seguimiento (actividades realizadas,

organismos participantes, gasto ejecutado, grado de ejecución y temporalización). Sin embargo, y como la recopilación de esta información se plantea desde el inicio del Plan, representa una buena base para realizar una verdadera evaluación de proceso y posterior evaluación de *outputs*. Aunque el alcance de la información y el grado de exhaustividad de la misma no es muy alto, la pertinencia de esta evaluación es alta para el seguimiento del mismo y para la recopilación de una información que resulta imprescindible para posteriores evaluaciones.

TABLA 7.1.
PERTINENCIA Y UTILIDAD DE LA INFORMACIÓN PRODUCIDA
SEGUNDA GENERACIÓN DE EVALUACIONES
 Cuadro resumen

	Globalidad de la información	Alcance de la información	Grado de elaboración práctica
ANDALUCÍA II (95-97)	Amplia	Amplio	Poco
CANARIAS I (95-96)	Limitada	Limitada	Algo
CASTILLA-LA MANCHA II (95-99)	Relativamente amplia	Relativamente amplio	Algo
CASTILLA Y LEÓN I (94-96)	Información no pertinente a la evaluación	Relativamente amplio	Nulo
GALICIA II (95-97)	Amplia (<i>outputs</i>)	Relativamente amplio	Algo
MADRID II (93-95)	Información no pertinente a la evaluación	Amplio	Nulo
INSTITUTO DE LA MUJER II (93-95)	Información parcial para cada tipo de evaluación	Amplio	Prácticamente inexistente
CATALUÑA II (94-96)	Amplia (<i>outputs</i>)	Relativamente amplio	Algo
PAÍS VASCO II (95-98)	Muy amplia (<i>outputs</i>)	Amplio	Bastante
ARAGÓN¹ I (94-96)	Limitada	Limitada	Nula

¹ Aunque de La Rioja disponíamos de información general, no pudimos disponer de ningún informe anual. Al no tener una base documental como en las restantes experiencias, no la incluimos en la tabla.

2. COMUNICACIÓN DE LOS RESULTADOS Y DESTINATARIOS DE LOS MISMOS

En este segundo apartado se analiza el proceso de comunicación de la evaluación como otro elemento clave para la utilidad y utilización de la misma. Para esto se estudia el(los) **informe(s) de evaluación**, en qué consisten, lo que reflejan del proceso de estudio llevado a cabo y cuán completos son³.

Los otros aspectos que se tienen en cuenta en este apartado son la **difusión** que se ha hecho de los informes de evaluación, quiénes han sido los **destinatarios** de dicha difusión y su **calendario**.

2.1. Comunicación de los resultados y destinatarios de los mismos en las evaluaciones externas

En la tabla 7.2.a se reflejan los esquemas de los informes de cada una de las seis experiencias externas de evaluación. En el caso de Andalucía, se presentan dos, uno correspondiente al informe intermedio y otro al informe final.

³ Como se ha explicado en el capítulo 4, en la parte correspondiente a los criterios para el análisis, se parte de la base de que un buen informe de evaluación debe explicitar los propósitos del estudio, describir el contexto del plan evaluado y la situación evaluativa concreta, relatar la metodología de evaluación (en la que se incluyan, además de las especificaciones técnicas de los instrumentos y técnicas utilizadas para la recopilación y el análisis de la información), una explicación de la estrategia evaluativa general, el modo en el que se han establecido los criterios de evaluación, los propios criterios de evaluación, las estructuras y foros que se han utilizado para la recopilación y cómo se han establecido las valoraciones y recomendaciones).

**TABLA 7.2. a: CONTENIDO DE LOS INFORMES DE EVALUACIÓN
EVALUACIONES EXTERNAS**

CONTENIDOS DE LOS INFORMES	
Andalucía II (95-97)	<p>“II PLAN ANDALUZ PARA LA IGUALDAD DE LAS MUJERES. AVANCE DEL INFORME DE EVALUACIÓN” (68 págs.) IESA-ANDALUCÍA. Documento. Julio 1997</p> <p>Índice Introducción (2 págs.) 1. Evolución de la situación social de la mujer (9 págs.) 2. Planes de igualdad de oportunidades de las mujeres en Andalucía (4 pp) 3. Modelo de evaluación (3 págs.) 4. Evaluación del diseño del II Plan de Igualdad (12 págs.) 4.1. Análisis de los objetivos y de los indicadores del II Plan 4.2. Análisis del objetivo final: El Plan y la igualdad de oportunidades 4.3. Análisis de la coherencia formal de los objetivos intermedios 4.4. Evaluación de los indicadores del II Plan 5. Proyecto de evaluación de la implementación (25 págs.) 5.1. Análisis de las actividades; 5.2. Aspectos de la organización; 5.3. Estudio piloto; 5.4. Propuesta de indicadores para la evaluación del II Plan; 5.5. Necesidad de un sistema de información 6. Evaluación de los resultados e impacto del Plan: estudios realizados a través de encuestas (11 págs.) 6.1. Encuesta sobre objetivos del Plan (Estudio 1) 6.2. Estudio sobre las actuaciones del Plan (Estudio 2) 6.3. Estudio sobre la situación social de la mujer en Andalucía en 1998 (Estudio 3) Bibliografía (2 págs.)</p> <p>“INFORME DE EVALUACIÓN. II PLAN ANDALUZ PARA LA IGUALDAD DE LAS MUJERES. 1995/1998”. (123 págs.) Instituto Andaluz de la Mujer. Publicación. 1999</p> <p>Presentación Índice Introducción: La igualdad de oportunidades y el II Plan (4 págs.) 1. La naturaleza de la Evaluación del Plan (8 págs.) 2. Desarrollo Democrático (19 págs.) 3. Desarrollo Educativo y Cultural (11 págs.) 4. Desarrollo del Bienestar Social (17 págs.) 5. Desarrollo Económico (17 págs.) 6. Perfil sociológico de las usuarias/beneficiarias del II Plan (15 págs.) 7. Satisfacción de las usuarias del II Plan (13 págs.) Anexo: II Plan Andaluz para la Igualdad de las Mujeres (21 págs.)</p>

<p>Canarias I (95-96)</p>	<p>“AUDITORIA. I PLAN DE IGUALDAD DE OPORTUNIDADES PARA LAS MUJERES CANARIAS 1995-1996” (265 págs.) Asesores de Imagen Corporativa S.A. Valencia, 20 Oct. 1996. Documento.</p> <p>Índice</p> <p>I. Introducción (4 págs.)</p> <p>II. Objeto del trabajo (14 págs.) Metodología (marco de actuación, análisis de actuaciones del ICM y análisis de actuaciones de Consejerías) Consideraciones generales</p> <p>III. Informe de auditoría (197 págs.) (1. Cultura y sociedad; 2. Educación; 3. Salud; 4. Justicia; 5. Empleo, formación, relaciones laborales; 6. Servicios Sociales; 7. Cooperación)</p> <p>IV. Conclusiones (9 págs.)</p> <p>V. ANEXOS (40 págs.) Anexo I: Fichas de solicitud de datos Anexo II: Relación de personas de contacto</p>
<p>Castilla-La Mancha II (95-99)</p>	<p>“ INFORME PREVIO DE EVALUACIÓN DEL II PLAN DE IGUALDAD DE OPORTUNIDADES ‘LAS MUJERES DE CASTILLA-LA MANCHA HACIA EL SIGLO XXI’” (84 págs.) Fundación Mujeres. Borrador. Documento interno.</p> <p>1. Estructura del informe (2 págs.)</p> <p>2. Estructura del II Plan de Igualdad de Oportunidades (2 págs.)</p> <p>3. Evaluación del diseño del Plan (21 págs.) A) Notas metodológicas B) Coherencia, concreción y posibilidades de realización del Plan C) Puntos fuertes y débiles D) Propuestas de mejora del Plan por parte de las entidades implicadas E) Conclusiones. Anexo de tablas</p> <p>4. Memoria de resultados de ejecución del Plan (44 págs.) (organizado por metas y objetivos del Plan)</p> <p>5. Informe de Valoración Cualitativa (15 págs.)</p> <p>5.1. Diseño del Plan</p> <p>5.2. Desarrollo del Plan</p> <p>5.3. Valoración de los resultados</p> <p>5.4. Impacto y nivel de conocimiento del Plan en la sociedad</p> <p>5.5. Propuestas para el diseño y desarrollo de futuras políticas de igualdad de oportunidades</p>

<p>Castilla y León I (94-96)</p>	<p>“RESUMEN DE LA EVALUACIÓN DEL PLAN INTEGRAL DE IGUALDAD DE OPORTUNIDADES PARA LA MUJER EN CASTILLA Y LEÓN 1994-1996” (78 págs.) Secretaría Regional de la Mujer. Documento interno. Sin fecha</p> <p>Índice Introducción: Diseño del estudio (5 págs.) (Diseño del estudio; Las fuentes; La presentación de la información; Contenido) Los cambios acaecidos en la situación de las mujeres de Castilla y León, durante el período de vigencia del PIIOM (26 págs.)</p> <ol style="list-style-type: none"> 1. Presentación de resultados 2. Los cambios acaecidos en la disponibilidad de empleo 3. Los cambios acaecidos en la demanda y en las condiciones de empleo 4. La facilitación del empleo para las mujeres en Castilla y León 5. La evolución de la salud de las mujeres de Castilla y León 6. Las anomias padecidas por las mujeres de Castilla y León 7. Las acciones culturales y de cooperación promovidas en el marco del PIIOM 8. La promoción del asociacionismo y de la participación en el marco del PIIOM 9. La participación institucional de la Junta de Castilla y León en actuaciones promovidas por organizaciones sin ánimo de lucro 10. La oferta y la demanda de educación entre las mujeres de C y L. <p>TABLAS (44 págs.) No existe un índice de tablas.</p>
<p>Galicia II (95-97)</p>	<p>“ INFORME DE LA EVALUACIÓN DEL II PLAN DE IGUALDAD (1995-1997)” (136 págs.) Amparo Almarcha Barbado. A Coruña, 25/2/1998</p> <p>CAP. 1: Introducción y metodología (15 págs.) 1. Concepto y metodología de la evaluación 2. Punto de partida de la evaluación del II PIOM 3. Recogida de información para la evaluación del II PIOM y sus implicaciones 4. Los instrumentos de recogida de información 5. El trabajo de campo 6. Dificultades en la recogida de información 7. La calidad y completitud de los datos 8. El procesamiento de la información</p> <p>CAP. 2: Descripción general de los resultados del estudio, por áreas del II PIOM (18 págs.)</p> <p>CAP. 3: Análisis de resultados (54 págs.) 1. Introducción 2. Evolución de las actividades por años 3. Clasificación de las actividades del PIOM 4. Destinatarios de las actividades del PIOM 5. Participación institucional 6. Ámbito geográfico 7. Organización de las actividades 8. Financiación de las actividades 9. Otros aspectos (canales de difusión, materiales, directivas europeas, convenios)</p> <p>CAP. 4: Análisis de las actividades de la Red de Información y Asesoramiento a la Mujer (RIAM) (23 págs.)</p> <p>CAP.5: Valoración de las actividades por parte de las personas responsables de los conciertos con las distintas instituciones (16 págs.)</p> <p>SÍNTESIS Y CONCLUSIONES (5 págs.)</p>

<p>Madrid II (93-95)</p>	<p>LAS MUJERES EN LA COMUNIDAD DE MADRID. DATOS PARA ABRIR CAMINOS HACIA LA IGUALDAD Publicación. Dirección General Mujer (Comunidad de Madrid). 1997. 145 pp.</p> <p>Presentación. 2 págs. Por Rosa Posada. Consejera de Sanidad y Servicios Sociales</p> <p>I. Características demográficas. 12 págs. II. Integración de las mujeres en el mercado laboral. 24 págs. III. La evolución de la calidad de vida y el empleo del tiempo de las mujeres. 32 págs. IV. La atención a la salud integral de las mujeres. 12 págs. V. La participación política y social de las mujeres. 10 págs. VI. La violencia que afecta a las mujeres. 24 págs. VII. La escolarización y los niveles de instrucción reglada. 16 págs. VIII. El analfabetismo y la formación de adultos. 8 págs. IX. Imagen de la DGM. 3 págs.</p>
-------------------------------------	--

En la evaluación del **II Plan andaluz** se entrega un avance del informe de evaluación en julio de 1997, en el que se refleja de forma bastante completa, el proceso de evaluación realizado hasta esa fecha. Existen unos primeros puntos introductorios sobre la situación social de la mujer andaluza y sobre los planes de igualdad andaluces, además de dedicar un punto específico al modelo de evaluación utilizado. El resto del informe está dedicado a reflejar una parte de la evaluación que ya se había realizado entonces (evaluación del diseño del Plan) y a comentar los proyectos y los avances realizados en cuanto a la evaluación de la implementación y a la evaluación de resultados a través de las encuestas. Es el único informe que tiene un cierto carácter académico, dedicando un apartado final a bibliografía. La difusión de este primer informe, debido a su carácter intermedio no final, es limitada y tiene una distribución únicamente interna para el Instituto Andaluz de la Mujer. La entrega de este primer informe se realiza tres meses después de lo estipulado en el convenio inicial (abril 1997).

En Andalucía, se publica un informe final en 1999, con bastante retraso sobre el calendario inicial propuesto, pero que tiene una difusión amplia. El informe es un buen resumen del proceso de evaluación. Además de los resultados del estudio, se expone una introducción sobre lo que significa el plan y un primer capítulo sobre la naturaleza de la evaluación, explicando el contexto de la misma y en qué consiste. Este es uno de los pocos informes estudiados que es explícito en esta materia. Como se ha comentado en el punto anterior, los resultados se exponen organizándose según las áreas del Plan; es decir, a pesar de que esta evaluación ha tenido diferentes “estudios”, los resultados no se organizan según fuentes o instrumentos sino que parece que se ha hecho un cierto análisis global de todo lo estudiado (a diferencia, por ejemplo, de la evaluación que realiza el IM y que veremos más adelante). La única excepción a esto es lo referente a la parte del estudio llamada “evaluación del impacto” que se reduce, al final del proceso, al análisis del grado de satisfacción de las usuarias con el Plan y que se refleja en un punto separado del informe.

El informe de evaluación del **I Plan canario** es bastante completo, si se tiene en cuenta las limitaciones del propio estudio de evaluación. Existen unas páginas dedicadas a la metodología y en las que se mencionan las dificultades existentes a la hora de recopilar la información, además de un apartado de conclusiones y unos anexos con las fichas de solicitud de datos que se utilizaron y un listado de las personas con las que se contactó para la recopilación de información. El resto del informe se dedica a exponer los datos recopilados, por áreas, y como hemos visto anteriormente, en muchos casos se limita a relacionar la información encontrada, probablemente en el mismo formato en el que ha sido ofrecida por los correspondientes organismos (bien sea el propio Instituto Canario de la Mujer, bien algún otro organismo o Consejería), ya que dicha información tiene un formato distinto para cada área (en algunos casos se reflejan los objetivos de las actividades y en otros no, en otros se habla de desarrollo, en unos pocos casos se exponen pequeñas valoraciones de cada una de las actuaciones, etc.). En general, la informe se limita prácticamente a señalar algún tipo de información sobre las actuaciones o actividades realizadas por cada uno de los objetivos y acciones del Plan.

En lo que se refiere a la difusión del informe canario, sabemos que éste se presentó al Consejo Rector del Instituto Canario de la Mujer en el mes de noviembre y, a los pocos días, se realizaron unas jornadas específicas de presentación del estudio de evaluación, invitando a colectivos y asociaciones y a mujeres de distintas instituciones. Es decir, se utiliza el informe de lo que se autodenomina “auditoría” del Plan para una presentación muy amplia de sus resultados. El calendario, además, se respeta de forma absoluta ya que el informe se entrega a finales de Octubre y se presenta el 9 de Noviembre de 1996 al Consejo Rector. Las Jornadas de presentación también se celebran durante el mes de noviembre, justamente antes de que empiece a ponerse en marcha el proceso de elaboración del II Plan.

La evaluación del **II Plan castellano manchego** se realiza de forma prácticamente paralela al proceso de elaboración del III Plan, por lo que, debido a la premura de tiempo y a la necesidad de dedicar todos los esfuerzos al nuevo Plan, el Informe final de evaluación del II Plan se entrega posteriormente a la aprobación del III Plan. Por lo que en este caso nos encontramos ante un desajuste claro de calendarios y tiempos. El Informe presentado a la Dirección General de la Mujer de Castilla-La Mancha por parte del equipo externo que realiza el estudio de evaluación refleja, de forma bastante completa, el proceso llevado a cabo. Además, este informe comienza con un punto en el que se expone cuál es la estructura del mismo y que refleja, de forma general, la metodología empleada, aunque le falta, al igual que en la mayoría de las experiencias analizadas, un primer punto en el que se exponga de forma clara la fundamentación y el propósito del estudio de evaluación, la delimitación del objeto a evaluar y las razones de los tipos de evaluación escogidos.

En ese primer punto introductorio del informe, se señala que la información que se ofrece se recopila a través de un sistema de recopilación diseñado para el estudio y que contemplaba básicamente dos ejes de recogida de información. Por un lado, la recopilación de resultados de ejecución de las acciones del Plan, a través de la

sistematización y ordenación de las memorias de resultados anuales de la Dirección General de la Mujer y de solicitudes de información y entrevistas a personal técnico y responsable de otros departamentos de la Junta implicados en el desarrollo de las acciones. Por otro, se recogieron las opiniones sobre el diseño, el desarrollo y el impacto del Plan, de los diferentes departamentos administrativos de la Junta implicados en la puesta en marcha y la ejecución de las acciones, y de las representantes de las asociaciones de mujeres y mujeres de las corporaciones locales a través de: entrevistas personalizadas (responsables y personal técnico), sesiones de trabajo sobre la evaluación de este II Plan con mujeres representantes de las organizaciones de mujeres y agentes sociales y una sesión de trabajo con mujeres alcaldesas y concejales. Toda la información recopilada a través de este proceso se sistematiza y estructura en el informe a través de tres apartados: 1) Evaluación del diseño del Plan, en el que se recogen una serie de reflexiones sobre las bondades y dificultades, fortalezas y debilidades que han podido surgir en el desarrollo del Plan, como consecuencia de su diseño. 2) Memoria de resultados de ejecución del Plan, donde se ofrece un mapa de los resultados obtenidos en función de las metas y los objetivos del II Plan, a través de la actuación de las diferentes Consejerías implicadas en su desarrollo. 3) Informe de valoración cualitativa, en el que se recogen las opiniones expresadas por mujeres representantes de asociaciones, de las corporaciones locales, y responsables y equipos técnicos. En cada apartado, además de la exposición de la información recopilada y analizada, se exponen cuestiones metodológicas y unas conclusiones.

El informe de evaluación del **I Plan castellano-leonés** refleja los resultados del estudio realizado, que como ya se ha visto, consiste en la comparación de los datos obtenidos a través de una serie de indicadores al principio y al final del período de vigencia del Plan. Se realiza internamente un resumen del informe de evaluación que expone una pequeña introducción sobre el diseño del estudio en el que se señalan, de forma muy general, las comparaciones realizadas, las fuentes utilizadas⁴ y cómo se presenta la información. Los cambios que se miden en la situación de las mujeres de Castilla y León, durante el período de vigencia del PIIOM se refieren al empleo (disponibilidad, demanda, condiciones y facilitación), la salud, las anomias y la oferta y demanda de educación. También se exponen datos relacionados con las acciones culturales y de cooperación promovidas por el PIIOM, la promoción del asociacionismo y de la participación y de la participación institucional de la Junta de Castilla y León en actuaciones promovidas por organizaciones sin ánimo de lucro.

Respecto a la difusión del informe, sabemos que los resultados del estudio se presentan al Consejo Regional de la Mujer en Marzo de 1997, prácticamente al mismo tiempo que se hace entrega del informe final por parte de la empresa externa que lo realiza. El resumen se realiza por una cuestión de “coherencia” en palabras de su responsable, ya que el informe final son cinco tomos y ni siquiera se podía plantear la posibilidad de hacer copias para todo el mundo del Consejo y además porque se

⁴ Lo que aparece en el resumen sobre las fuentes secundarias es lo siguiente: “Las estadísticas que se han buscado, reelaborado y analizado, en cualquier institución que ofrezca datos, que tengan que ver con el contenido del PIIOM y que se refieran a la Comunidad de Castilla y León” (Resumen de la evaluación pág. 6)

consideraba que muchos de los gráficos tenían un “contenido demasiado técnico y desde el punto de vista de la divulgación es poco eficaz”; sin embargo, se insiste en la disponibilidad del informe en su totalidad para quien quiera ir a las oficinas de la Secretaría Regional de la Mujer para consultarlo. También se insiste en que en el Consejo Regional se sabe de la existencia y disponibilidad del informe. El resumen de la evaluación que se realiza por parte de la Secretaría Regional de la Mujer se difunde por parte de dicha Secretaría, incluida su entrega a los asistentes al Seminario que se realiza en la Universidad Internacional Menéndez Pelayo de Santander en el verano de 1998 sobre “Evaluación de las políticas de igualdad entre mujeres y hombres”, y en el que una persona representante de la Secretaría Regional explicaba, en una mesa redonda, la experiencia de evaluación del II Plan Integral castellano leonés.

El informe de la evaluación del **II Plan gallego** recoge y sistematiza bien la información recopilada, aunque reconoce las dificultades y limitaciones de dicha información. En el informe existe un primer capítulo de introducción y sobre la metodología del estudio realizado, que incluye consideraciones generales sobre el concepto y la metodología de la evaluación, además de las dificultades y limitaciones mencionadas. Asimismo se refleja bien cuál era la finalidad del estudio y el tipo de evaluación (de proceso y de resultados) que se pretendía. Además explica el sistema de recopilación de información, los instrumentos utilizados, el proceso del trabajo de campo y el procesamiento de la información. Es decir, es un informe bastante completo en lo que a explicitación de la metodología y procedimientos se refiere, señalando incluso algo sobre cómo se procesa la información, que en la mayoría de los casos no se reseña. Sin embargo, y al igual que en el resto de las experiencias analizadas, no se explicitan los criterios de evaluación: se habla a grandes rasgos de una evaluación de proceso y de resultados y luego se señalan los contenidos de la ficha de recopilación de información. A través de la explicación del instrumento se extraen las variables y aspectos estudiados, pero éstos no están explícitamente formulados como criterios de evaluación, por lo que en muchos casos no están suficientemente fundamentados ni se explicita cuál es la base lógica para la valoración posterior (los criterios de valor).

Además de este capítulo metodológico y otros cuatro referidos a los resultados de la evaluación, existe un último capítulo sobre síntesis y conclusiones que, aunque no extenso (cinco folios, es decir, un 4% del total del informe), recoge conclusiones generales sobre la importancia relativa concedida a las distintas áreas del Plan (objetivos, prioridades y áreas), sobre las actividades y su incidencia (incremento de actividades por años, destinatarios, distribución geográfica, financiación, canales de difusión de las actividades) y un punto específico sobre actividades y prioridades de la Red de Información y Asesoramiento a la mujer (RIAM).

La difusión del informe gallego ha sido muy limitada, aunque no era la intención inicialmente. Según la responsable de la Agencia en verano de 1997, “la evaluación del plan, el plan evaluado, yo mi idea es que se publique (...), todavía no sabemos que posibilidades va a tener, pero desde luego, hacerlo público (...) por supuesto que sí, nosotros quisiéramos hacerlo público y ya sabemos que va a haber muchas cosas en las que no vamos a estar especialmente contentas...”; incluso se comenta que “es que además aunque no quisiéramos, nos lo iban a reclamar desde el Parlamento y tendrían

todo el derecho”. La cuestión es que se da un cambio de responsable política y quien recibe el informe final de evaluación no es la misma persona que había encargado la evaluación y con quien había trabajado el equipo de evaluación durante el proceso. Según la información que nos consta la difusión del informe de evaluación ha sido muy limitada (ni siquiera tenemos constancia de que se presentara el informe de evaluación al Parlamento gallego) y éste ha sido uno de los informes a los que ha sido más difícil acceder a efectos de la realización de esta investigación, debido a que la Agencia de Promoción de Igualdad no nos lo ha facilitado.

El informe de la evaluación del **II Plan madrileño** se publica y se presenta en un acto público en el año 1998. La publicación presenta un resumen de los resultados pero no hace ninguna mención ni a la fundamentación o propósitos de la evaluación, a los criterios utilizados ni a los procedimientos o metodología utilizados para la obtención de dichos datos. Ni siquiera se presenta una valoración de dichos datos (ya se ha comentado que se dice explícitamente que no se realiza dicha valoración para que sea el propio lector el que la realice) ni conclusiones o recomendaciones. Es decir, la publicación del informe refleja sólo parte del estudio realizado y en nada de asemeja a un informe final de un estudio de evaluación. De hecho, en la publicación no aparece en ningún momento en el título la palabra “evaluación”, titulándose “Las mujeres en la Comunidad de Madrid”. Sólo en la presentación, firmada por la entonces Consejera de Sanidad y Servicios Sociales, Rosa Posada, se hace referencia a que los resultados que se presentan se refieren al estudio de evaluación del II Plan.

Esta publicación a la que hacemos referencia es la que se difunde y distribuye por parte de la Dirección General de la Mujer. Los informes finales que presenta la consultora externa que realiza el estudio de evaluación son dos, correspondiendo a los dos estudios que se realizan (elaboración de un sistema de indicadores y evaluación propiamente dicha) como ya se ha reflejado anteriormente en este mismo capítulo. En el primer informe “Elaboración de un sistema de indicadores para evaluar los resultados obtenidos con el II PIOM” se exponen los criterios utilizados para seleccionar los indicadores, la propuesta metodológica y el diseño del estudio. En el informe referido a la “Evaluación del II Plan de Igualdad de Oportunidades para las Mujeres de la Comunidad de Madrid” se recogen los resultados a través del sistema de indicadores que posteriormente son resumidos en la publicación para su difusión.

2.2. Comunicación de los resultados y destinatarios de los mismos en las evaluaciones mixtas

En la tabla 7.2. b se reflejan estos aspectos en las tres evaluaciones mixtas.

**TABLA 7.2. b: CONTENIDO DE LOS INFORMES DE EVALUACION
EVALUACIONES MIXTAS**

	CONTENIDOS DE LOS INFORMES
Instituto Mujer II (93-95)	<p>“SÍNTESIS DE RESULTADOS DE LA EVALUACIÓN DEL II PLAN DE IGUALDAD DE OPORTUNIDADES DE LAS MUJERES (1993-1995)” (Publicación, Instituto de la Mujer, 1997) (85 págs.)</p> <p>Introducción (3 págs.)</p> <p>1. Estudio de la implementación (15 págs.) (Área jurídica; Educación; Empleo; Imagen Social; Responsabilidades familiares; Participación social y política; Acceso a puestos de decisión; Salud; Marginación; Área internacional)</p> <p>2. Evaluación cuantitativa (13 págs.)</p> <p>2.1. Estereotipos; 2.2. Valores; 2.3. Actitudes hacia el trabajo extradoméstico; 2.4. Calidad de vida percibida; 2.5. Usos del tiempo</p> <p>3. Evaluación cualitativa (16 págs.) (Los puntos de este apartado se desarrollan con los objetivos de las áreas siguientes: Área jurídica; Educación; Empleo; Participación social y política; Acceso a puestos de decisión; Salud; Marginación)</p> <p>ANEXO: Grado de cumplimiento de las actuaciones correspondientes a las 10 Áreas del PIOM (21 págs.) (En unas tablas se consignan las actuaciones por cada área y se señala con una cruz, por cada una de ellas si están iniciadas o no iniciadas)</p>

<p>Cataluña II (94-96)</p>	<p>“INFORME DE EVALUACIÓN DEL II PLAN DE ACTUACIÓN DEL GOBIERNO PARA LA IGUALDAD DE OPORTUNIDADES PARA LAS MUJERES” (242 + 10 págs.) ICD. Documento interno. Versión revisada 25/4/1997</p> <p>METODOLOGÍA SEGUIDA EN LA EVALUACIÓN (10 págs.)</p> <ol style="list-style-type: none"> 1. Finalidad de la evaluación 2. Etapas seguidas en el proceso de evaluación y fuentes de información 3. Identificación de las fuentes a evaluar 4. Aspectos considerados en el análisis de las acciones a evaluar 5. Período de realización y equipo evaluador <p>ÍNDICE</p> <p>EVALUACIÓN GLOBAL DEL PLAN (60 págs.)</p> <ol style="list-style-type: none"> 1. Evaluación del cumplimiento del Plan (A. Estructuración del Plan; B. Grado de implementación de las acciones; C. Grado de consecución de los objetivos; D. Análisis del grado de implementación según tipo de política y distribución territorial) 2. Evaluación de la participación de los diversos actores responsables de implementar las acciones del Plan 3. Evaluación de las acciones dirigidas a los diversos colectivos destinatarios del Plan 4. Recursos destinados al Plan por ámbito, actor y colectivo destinatario <p>EVALUACIÓN DEL ÁMBITO 1 DE SENSIBILIZACIÓN (20 págs.)</p> <p>EVALUACIÓN DEL ÁMBITO 2 LEGISLATIVO(23págs.)</p> <p>EVALUACIÓN DEL ÁMBITO 3 SOCIO-CULTURAL (27 págs.)</p> <p>EVALUACIÓN DEL ÁMBITO 4 DE LA COEDUCACIÓN (29 págs.)</p> <p>EVALUACIÓN DEL ÁMBITO 5 DE LA FORMACIÓN OCUPACIONAL, PROFESIONAL Y EL TRABAJO (35 págs.)</p> <p>EVALUACIÓN DEL ÁMBITO 6 SANITARIO (14 págs.)</p> <p>EVALUACIÓN DEL ÁMBITO 7 DE LOS SERVICIOS SOCIALES Y COMUNITARIOS (33 págs.)</p>
---------------------------------------	--

<p>País Vasco II (95-98)</p>	<p>MEMORIA DE ACTUACIÓN DE LOS PODERES PÚBLICOS EN 1997 EN EJECUCIÓN DEL II PLAN DE ACCIÓN POSITIVA PARA LAS MUJERES (PAPME) EN LA COMUNIDAD AUTÓNOMA DE EUSKADI <i>Emakunde. Documento interno(159 páginas + 42 páginas de Anexos)</i></p> <p>I. INTRODUCCIÓN (10 págs.) Presentación A. Metodología de la evaluación B. Participación de los organismos públicos en la evaluación</p> <p>II. RESULTADOS DE LA EVALUACIÓN A) RESULTADOS GENERALES (101 págs.) 1. Actuación de los tres niveles de la administración pública vasca (organismos que han gestionado -por áreas, subáreas, niveles /categorías y colectivos- y que han colaborado en la ejecución) 2. Colectivos destinatarios (en función de las áreas, subáreas y niveles de intervención -categorías-) 3. Grado de ejecución del II PAPME 3.1. Análisis comparativo entre las acciones propuestas en el II PAPME y aquellas sobre las que los organismos públicos han actuado durante 1997 3.2. Actividades realizadas durante 1997 por los organismos públicos en función de las 9 áreas y de las subáreas 3.3. Distribución de las actividades realizadas durante 1997 en función de las 9 áreas y del nivel de intervención (categorías)</p> <p>B) RESULTADOS ESPECÍFICOS DE CADA CATEGORÍA (28 págs.)</p> <p>III. CONCLUSIONES (21 págs.) A) Consideraciones respecto a la metodología B) Balance de los resultados</p> <p>ANEXOS (42 págs.) (Documentos de solicitud de información a Ayuntamientos, instrumentos de evaluación de las acciones correspondientes a cada una de las 7 categorías, y listado de colectivos destinatarios categorizados)</p>
---	--

La evaluación del **II PIOM estatal** produce, inicialmente, tres informes diferenciados: el informe sobre la “evaluación de la implementación” se entrega en Febrero de 1996, el informe de la “evaluación cualitativa” en Julio de 1996 y el de la “evaluación cuantitativa” en Diciembre de ese mismo año. En el año 1997 se realiza internamente una síntesis de resultados (se entrega en marzo de 1997, el mismo día en que el Consejo de Ministros aprueba el III PIOM). Esta síntesis de resultados se plasma en una publicación a finales del año 1997. Esta publicación, de 83 páginas, resume los principales resultados de cada uno de los estudios, que se reflejan por separado, sin ninguna conclusión de cada parte y sin un punto que integre y dé una visión global del proceso evaluativo en su conjunto. Además de estos tres apartados, la publicación consta de una introducción en la que se explica de manera general el II PIOM y, muy sucintamente, la evaluación del mismo. En esta introducción se dice que “cada una de las vertientes (los tres tipos de evaluación planteados) han sido planteadas, en principio, de manera bastante independiente, de forma que el IM, con ayuda de entidades privadas en algunos casos, ha elaborado sendos informes” (p. 9 de la “Síntesis.”); pero no se fundamenta o justifica el por qué de esa independencia, simplemente se constata. En esta publicación se incluye un anexo titulado “Grado de cumplimiento de las actuaciones correspondientes a las 10 áreas del II PIOM” y que consiste en un cuadro en el que aparecen reproducidas cada una de las 170 actuaciones, organizadas por áreas. En

esos cuadros se consigna, con una cruz, si las actuaciones han sido iniciadas (en una columna), o no iniciadas (en otra columna). De las 170 actuaciones, sólo 11 de ellas constan como “no iniciadas”⁵. Dicha publicación resumen es la que resulta accesible para el público, no los informes directamente. Incluso en la presentación de la publicación se dice que cada uno de sus tres apartados “hace referencia concreta al correspondiente informe, *que obra en los archivos del Instituto*” (p. 9, énfasis nuestro). Efectivamente, los informes originales están en archivos del Instituto⁶, pero no accesibles directamente al público en el Centro de Documentación correspondiente.

Lo que se refleja en la síntesis publicada, y que es un fiel resumen de los tres informes originales, son los resultados arrojados por los tres estudios. En el “Estudio de la implementación”, sólo hay un párrafo de 7 líneas sobre la metodología empleada y otro sobre los objetivos de dicho estudio. Asimismo, en el de la “Evaluación cualitativa”, el párrafo dedicado al objetivo de dicho estudio es de 6 líneas y el de la metodología de 4. En el de la “Evaluación cuantitativa”, sin embargo, sí se especifica, muy profusamente, las características técnicas del muestreo realizado para aplicar el cuestionario utilizado, así como los aspectos medidos en dicho cuestionario y las escalas utilizadas. Es decir se especifica mucho sobre la técnica concreta utilizada, pero ni se fundamenta el estudio, ni se explicitan los criterios de evaluación utilizados ni los procedimientos de evaluación. En definitiva, dicho informe no puede ser calificado como completo, si se tienen en cuenta los requisitos mínimos de un informe de evaluación (*Joint Committee*, 1994: 49).

⁵ Las acciones que constan como no iniciadas son la 1.1.8. “Impulsar, en el marco de la Conferencia de Ministros de Justicia de los países iberoamericanos, el conocimiento y la divulgación de las leyes igualitarias dictadas en España y sus instrumentos de aplicación”; 2.4.3. “Introducir el principio de igualdad de oportunidades en la evaluación del sistema educativo, como uno de los indicadores de la calidad de la enseñanza”; 2.5.3. “Favorecer que en las comisiones que evalúen la investigación y la docencia, o que decidan sobre la distribución de recursos y ayudas del Ministerio de Educación y Ciencia, en sus distintos niveles, se tenga en cuenta la perspectiva de la igualdad de oportunidades entre los sexos”; 2.7.1. Seguir incorporando criterios que favorezcan la participación equilibrada entre los sexos en la Normativa del Consejo Superior de Deportes”; 6.2.2. “Elaborar propuestas de descentralización de las oficinas públicas en sedes o circunscripciones más pequeñas con unificación de la gestión de todas las prestaciones sociales y de una sola sede”; 6.3.5. “Establecer mecanismos de coordinación con las mujeres de la Unión Interparlamentaria, organismo que realiza análisis comparados entre países, sobre la participación política”; 7.2.4. “Elaborar recomendaciones para que sean admitidos/as como candidatos/as a los cursos de formación del INAP, los funcionarios/as que se encuentren disfrutando del año de excedencia especial por cuidado de hijos/as menores”; 7.3.3. “Suscribir un convenio con la Dirección General del Patrimonio del Estado para promover el desarrollo de un Plan de Igualdad de Oportunidades”; 8.6.2. “Elaborar un mapa de riesgos sanitarios de las mujeres, en que se tenga en cuenta variables no sólo relativas al puesto de trabajo sino también variables ambientales, psíquicas y de condiciones de vida de las mujeres”; 8.6.5. “Realizar experiencias piloto en base a diversas medidas preventivas de salud y sociales en el ámbito laboral, formando parte de los proyectos subvencionados por la iniciativa NOW; teniendo en cuenta las recomendaciones de la Comisión Europea en cuanto a la mejora de la calidad del trabajo de las mujeres”; y 8.7.1. Impulsar la creación de un Grupo de Trabajo en el que intervengan representantes del Ministerio del Interior, de Justicia, de Sanidad y Consumo, del de Asuntos Sociales y de Asociaciones de Mujeres, para elaborar un protocolo de malos tratos y agresiones sexuales a las mujeres.

⁶ A efectos de la investigación, nosotros no tuvimos ningún problema para consultar los informes en el despacho del Consejero Técnico de Estudios del IM.

El informe de la evaluación del **II Plan catalán** es bastante completo y parece que refleja bien el proceso de estudio seguido. Es un informe quizá demasiado amplio a efectos de su difusión y divulgación (de más de 250 páginas) y en él se exponen los resultados de evaluación, además de por cada uno de los ámbitos o áreas del plan, de forma global, elaborando de forma general sobre la evaluación del cumplimiento del Plan (estructuración, grado de ejecución, grado de cumplimiento de objetivos y análisis del grado de ejecución de las acciones según el tipo de política, el tipo de actuación y la distribución territorial), la evaluación de la participación de los diversos actores, la evaluación de las acciones destinadas a los diversos colectivos destinatarios y los recursos destinados al Plan por ámbito, actor y colectivo destinatario. Además, se adjunta una breve introducción, aunque está va por separado del grueso del informe, incluso paginada de forma distinta, sobre la “metodología seguida en la evaluación del Plan”. En esta introducción se reflejan cinco puntos, entre los que se encuentra la finalidad de la evaluación, siendo ésta una de los pocos informes analizados que contempla un apartado específico sobre ello. Además, se habla de las etapas seguidas en el proceso de evaluación y las fuentes de información, la identificación de las fuentes a evaluar, los aspectos considerados en el análisis de las acciones a evaluar y el período de realización y el equipo evaluador. También es uno de los pocos informes que expresa explícitamente algo relacionado con los criterios de evaluación (los aspectos considerados en el análisis), ya que se explican cada uno de los aspectos considerados y los criterios empleados por el equipo evaluador para la evaluación de cada una de las acciones⁷, por ejemplo, para el grado de ejecución de las acciones se emplean cuatro categorías: totalmente realizada, parcialmente realizada, en ejecución y no realizada. Aquí también es donde se explican los criterios seguidos para evaluar el tipo de política y las categorías utilizadas de “promoción”, “igualdad” y “paridad”. Lo que sí falta en este informe, al igual que en la mayoría de los informes analizados, es un apartado de conclusiones finales y otro de recomendaciones derivadas del estudio, aunque después de cada apartado (evaluación en cada ámbito) existe un subapartado en el que se refleja una valoración global del cumplimiento del Plan en cada ámbito.

Respecto a la difusión del informe catalán, se realizó una rueda de prensa pero no se hizo una publicación específica para divulgarlo, aunque la responsable del Instituto Catalán de la Mujer comenta en la entrevista realizada, en diciembre de 1997, que “va a estar en el centro de documentación para consulta”. Curiosamente, el informe final de evaluación es del mes de Mayo de ese mismo año, es decir, el informe de evaluación no estaba todavía disponible en el centro de documentación siete meses después de su presentación. Es posible que este hecho esté relacionado con las dificultades que tuvimos para que nos enviaran una copia del informe para la realización de esta investigación⁸ (lo obtuvimos después de 13 meses de petición a través de gestiones

⁷ Como ya se ha explicado anteriormente, el foco de la evaluación del II Plan catalán se encuentra en las acciones del Plan, no tanto en el Plan de forma global. Sin embargo, es una de las experiencias en las que existe un mayor grado de análisis del Plan en su conjunto y no sólo de cada actuación o área, como ocurre en otras experiencias. Es decir, no se evalúa el diseño del Plan en sí mismo o los objetivos del mismo, aunque se recoge información global muy valiosa para ello (quiénes participan, a quiénes va dirigido, tipo de política, recursos dedicados a cada cosa, etc.).

⁸ Uno de los temas que deja pendiente esta investigación, y probablemente cualquier investigación que requiera recopilar y analizar documentalmente informes de evaluación, es la falta de criterios en las administraciones públicas españolas sobre el carácter público de los informes de evaluación. Mientras no tuvimos ningún problema en la obtención de una copia (o en poder consultarlos “in situ”) de los informes

telefónicas y faxes). Respecto a la difusión, también se mencionan unas jornadas realizadas a propósito de los 10 años de igualdad en junio de 1997, en las que se aprovechó para contar los resultados de la evaluación por cada ámbito. Es decir, no fueron unas jornadas específicas para divulgar la evaluación, sino que en cada una de las conferencias en torno a cada ámbito concreto, además de un ponente, había una representante del Instituto que contaba los resultados de evaluación para cada ámbito. A estas jornadas acudieron responsables de los diferentes departamentos de la Generalidad y otros agentes implicados en el desarrollo del Plan⁹. En definitiva, el informe de evaluación y sus resultados no parece haber tenido mucha difusión en sí mismo, sino ligado más bien al proceso de elaboración del III Plan. Se insiste mucho en la participación que han tenido todas las fuerzas sociales (ONG, sindicatos, patronales, etc.) en la elaboración del III Plan, es decir se ha contado con ellas para esta elaboración pero, sin embargo, no parecen haber sido destinatarios primarios del estudio de evaluación del II Plan.

Las “memorias de actuación de los poderes públicos en ejecución del **II Plan de Acción Positiva para las Mujeres de Euskadi**” que realiza *Emakunde* anualmente, para presentar en el Parlamento vasco, son informes muy completos y que reflejan exhaustivamente el proceso seguido de evaluación a lo largo del año. Como ya se ha mencionado, estos informes tienen la virtud de producir información anualmente, por lo que se tiene una información sobre el seguimiento y la evaluación del II PAPME de forma continuada. Esto hace que potencialmente la información que produce el proceso de evaluación sea más útil y utilizada. Además de los resultados, se ofrece mucha información sobre el proceso metodológico de evaluación. Son informes, además, que tienen un claro carácter público, ya que se realizan con el objetivo de ser presentados al poder legislativo.

Como se puede observar en la tabla 7.2. b, además de una introducción en la que se explicita la metodología de evaluación y el grado de participación de los organismos públicos en la misma, los resultados se exponen de forma general (analizando el Plan en su conjunto) en torno a tres ejes: la actuación de los tres niveles de la administración pública vasca (comunidad autónoma, diputaciones forales y ayuntamientos), los colectivos destinatarios de las actividades del PAPME y el grado de ejecución del Plan en el año concreto en el que se realiza la memoria concreta (en este caso el año 1997).

en algunos de los casos estudiados, en otros encontramos reticencias a la obtención de los informes por nuestra parte. Mientras en algún caso el motivo de que finalmente pudiéramos obtener una copia tuvo que ver con nuestra pertenencia a la Universidad y el apoyo institucional que esto nos daba, en otros la obtención se realizó a nivel personal y de manera privada. Aunque esto excede los límites de este trabajo, uno de los temas que queda pendiente de análisis y estudio en nuestro país es el carácter público que los informes de evaluación (de programas y políticas públicas), su difusión y divulgación y el estatus administrativo que éstos deben tener, además de los criterios que deberían utilizarse y el marco legal de los mismos.

⁹ De alguna forma se “aprovecharon” los actos del décimo aniversario (“se divulgaba la evaluación y las líneas estratégicas que nos estábamos marcando cara al tercer plan”). En este sentido, no sólo no hubo una divulgación específica del estudio de evaluación, sino que tampoco parece que en seis conferencias de dos horas de duración en la que había un ponente y además se discutían las líneas del III Plan y se debatían con un audiencia media de 150-160 personas, se pudieran tratar con mucha profundidad los resultados de la evaluación.

Además se ofrecen unas conclusiones, no sólo respecto a los resultados, siguiendo el mismo esquema que en la exposición de los resultados, sino también respecto a la metodología de evaluación. La exposición de la información está muy bien estructurada y es, por lo tanto, muy clara y fácil de seguir. Los tres apartados del informe tienen una clara ligazón entre sí y resulta especialmente claro el proceso lógico de extracción de conclusiones, cuestión no tan evidente en los informes de otras experiencias analizadas.

La experiencia de evaluación vasca es una de las pocas en las que encontramos claramente definidos los destinatarios de la evaluación, quedando claro que el destinatario primordial y primario de los informes es el propio Parlamento Vasco. En segunda instancia, los destinatarios del proceso de evaluación y sus informes son los diferentes niveles e instituciones de las administraciones públicas vascas, incluido claro está, el propio *Emakunde*. En tercera instancia se encuentran los colectivos y asociaciones de mujeres que tienen alguna implicación en el II PAPME. La difusión de estos informes, al ser presentados al Parlamento Vasco, está asegurada. Existe además bastante transparencia al respecto y ésta es una de las comunidades autónomas en las que no han puesto ningún obstáculo para la obtención de información.

2.3. Comunicación de los resultados y destinatarios en las evaluaciones internas

En la tabla 7.2. c se reflejan estos aspectos en las dos evaluaciones internas.

TABLA 7.2. c: CONTENIDO DE LOS INFORMES DE EVALUACION EVALUACIONES INTERNAS

	CONTENIDOS DE LOS INFORMES
Aragón I (94-96)	<p>“INFORME SOBRE EL GRADO DE CUMPLIMIENTO DEL PLAN DE ACCIÓN POSITIVA PARA LAS MUJERES EN ARAGÓN. 1994-1996” (Documento interno, Instituto Aragonés de la Mujer, Enero 1997) (93 págs.)</p> <p>(Documento sin índice) INTRODUCCIÓN (25 págs.) Organismos implicados en el desarrollo del Plan Análisis global de las acciones (acciones realizadas y sin realizar, grado de cumplimiento de las acciones del Plan, acciones del Plan desarrolladas por el IAM) ACCIONES MÁS RELEVANTES EN CADA UNA DE LAS ÁREAS DEL PLAN (44 págs.) (1. Legislación igualitaria; 2. Cultura e imagen; 3. Educación y reparto de responsabilidades; 4. Relaciones laborales, empleo y formación; 5. Salud; 6. Servicios sociales; 7. Asociacionismo; 8. Cooperación internacional) PUBLICACIONES Y OTROS MATERIALES ELABORADOS O EDITADOS POR EL IAM ENTRE 1994 Y 1996 (5 págs.) ANEXO: CUADROS Y GRÁFICOS (23 págs.). Sin índice ni identificación o numeración de tablas.</p>
La Rioja II (96-99)	NO SE HA RECIBIDO NINGÚN INFORME DE EVALUACIÓN

El “Informe sobre el grado de cumplimiento del **Plan de Acción Positiva para las Mujeres en Aragón 1994-96**” consta de una introducción sobre el propio informe y su sentido, de un apartado en el que se especifica la implicación de otros organismos en el desarrollo del Plan, y otro en el que se hace un análisis global de las acciones (grado de cumplimiento y acciones desarrolladas) y se especifican las categorías utilizadas (acciones realizadas -en su totalidad, habituales o parcialmente- y acciones sin realizar). Además, se especifican las acciones más relevantes en cada una de las ocho áreas del Plan. En este sentido, es un informe que refleja fielmente el proceso de análisis realizado; la cuestión es que dicho análisis fue muy limitado, como las propias responsables reconocen. La base del proceso es una ficha que se rellenaba cada vez que

se realizaba una acción. Se rellenaron más de 400 fichas en los tres años (94, 95 y 96) y se hizo el informe con esta base. El informe consiste, sobre todo, en contar lo que se hizo, y aunque existe una cierta base empírica, no se realiza todo el análisis que hubiera sido posible.

Como ya se ha comentado anteriormente, el proceso de evaluación realizado se hace en un momento determinado del desarrollo del plan (en el cuarto mes de su tercer y último año), dado un cambio de responsables políticos, y con el fin de tener una visión global y saber qué se había realizado hasta entonces. Según el propio informe “el IAM realizó una puesta al día del grado de ejecución conseguido hasta esa fecha en cada una de las acciones del Plan, así como una priorización de objetivos a conseguir en el resto del año, incluyendo acciones para cada uno de los Departamentos; todas ellas fueron presentadas a las nuevas representantes del Consejo Rector y aprobadas de común acuerdo” (p. 2). Es decir, en su momento, el informe del proceso de evaluación se utilizó internamente, tanto por el propio Instituto (su Consejo Rector), como por parte de cada uno de los Departamentos del Gobierno de la Diputación General de Aragón. Este mismo proceso de evaluación y la información que de él se extrae, además de la que se sigue recogiendo a lo largo del año 1996, sirve para elaborar un informe de la totalidad del Plan que se presenta en Enero de 1997. Este informe se utiliza igualmente de forma interna y sirve como referencia para el proceso de elaboración del II Plan.

La evaluación del **II Plan riojano** se realiza a lo largo del proceso. Anualmente se produce una memoria y también se ha encargado una evaluación final externa para el final del Plan y que será entregada a lo largo del año 2000. El último informe anual, el correspondiente al año 1999, también se encarga a una empresa externa, debido a la carga interna de trabajo que había representado en los tres años anteriores.

3. EL IMPACTO DE LOS RESULTADOS Y LOS PROCESOS DE EVALUACIÓN EN LA POLÍTICA Y EN LOS ORGANISMOS DE IGUALDAD

En este apartado se analiza el impacto que han tenido las experiencias de evaluación, tanto en los propios planes y políticas de igualdad de género, como en los organismos de igualdad impulsores de dichas políticas. Ya se ha visto que el concepto de impacto de las evaluaciones está inextricablemente unido al de uso; el impacto de las evaluaciones depende del uso que se haga de ellas.

Pero, además de en los usos de los **resultados** de la evaluación, aquí se pretende analizar también los usos de los **procesos de evaluación**, sirviendo éstos para introducir una cultura y una metodología de evaluación y para generar una mayor comprensión de la intervención pública en general además de una mayor conciencia sobre cómo se abordan los problemas públicos.

En este sentido, se analizan indicios de: 1) Los efectos de los mismos procesos de evaluación, incluida una mayor comprensión y/o apreciación de la evaluación, el posible fortalecimiento o institucionalización de la función de evaluación y el aprendizaje sobre dicha función de evaluación. 2) Una mayor comprensión o nuevas perspectivas en una situación, programa u organización (que sería lo que se ha llamado “uso conceptual”). 3) Las acciones que se llevan a cabo como resultado de la evaluación (llamado “uso instrumental”). Así, buscamos algún indicio o “rastros” de utilización en estos tres sentidos para poder dar algo de luz sobre este aspecto, además también de estar atentos a encontrar algún tipo de “uso persuasivo” o de uso en otras instituciones y que vaya más allá del programa evaluado¹⁰.

3.1. El impacto de los resultados y los procesos de evaluación externa

En el caso de la evaluación del **II Plan andaluz**, el estudio se desarrolla de una forma bastante “externa” y no se da un buen ajuste a las expectativas de las responsables del Instituto. Esto, unido a un problema de retrasos en los plazos y a una concepción del estudio de evaluación más de investigación académica que de consultoría profesional por parte del equipo evaluador, hace que la evaluación no haya tenido especial impacto en la normalización y fortalecimiento de la función de evaluación en el Instituto Andaluz de la Mujer.

¹⁰ Como se explica en el capítulo 4, se utiliza también la clasificación de Carol Weiss respecto a los tipos de uso: instrumental, conceptual, persuasivo y el que influye en otras instituciones, más allá del programa evaluado. (Weiss, 1998b: 23; punto 4.3.1., capítulo 4).

Tanto la evaluación del diseño como la de la implementación cumplen un papel importante y arrojan unos resultados que parecen útiles para una mejor comprensión de las acciones emprendidas. En el caso de la evaluación de impacto, no se cumplen las expectativas del IAM, ni de lo reflejado en el propio proyecto inicial de evaluación: el análisis del grado de satisfacción de las usuarias es interesante como información -los resultados corresponden a un grado de satisfacción alto- pero resulta una información algo tangencial respecto a conocer el impacto y los efectos netos del Plan.

La evaluación del **I Plan canario** no parece cumplir un papel importante respecto a los efectos de la propia evaluación. El estudio se encarga externamente porque, en aquel momento, la plantilla del ICM era prácticamente inexistente (se reducía a 4 personas, incluida la directora y su secretaria), por lo que tampoco existe ningún contacto del equipo evaluador con el *staff* que posteriormente va a componer la plantilla técnica del ICM y a encargarse de la función de evaluación en dicho Instituto. El equipo técnico, compuesto por seis personas, que entra en la plantilla del Instituto lo hace el día 4 de noviembre de 1996, sólo 5 días antes de la presentación del estudio al Consejo Rector del ICM. Esto significa que no hubo ningún tipo de intercambio real ni, por lo tanto, aprendizaje del *know how* de la evaluación ni de las dificultades que el equipo evaluador encontró en la realización del estudio.

Respecto al impacto en lo que se refiere a una mayor comprensión de las políticas de igualdad y de una determinada situación o problema, la evaluación canaria simplemente aportó una idea general sobre lo que se había hecho y no se había hecho al amparo del I Plan de Igualdad. Por lo que, en cierto modo, ofreció una visión global de en qué consistía un plan de igualdad ya que, entre otras cosas, este plan era el primero que se aprobaba e implantaba en Canarias. Es decir, el impacto a este respecto puede calificarse de muy bajo.

Por último, las acciones que se realizaron como consecuencia de la evaluación realizada incluyen la realización de unas jornadas de presentación de dicho estudio de evaluación en las que se invitaron a diferentes mujeres y representantes de las instituciones canarias implicadas. Esto suponemos que contribuyó a lograr una mayor visibilidad del ICM y del propio Plan de igualdad y a realizar otra actividad que fomentara la implicación de las diferentes instituciones y Consejerías en las políticas de igualdad, e incluso a cumplir una cierta función de rendimientode cuentas en estos otros departamentos respecto a las dificultades de recopilación de información que había tenido el equipo evaluador. Esta misma información del estudio de evaluación sirvió también de alguna manera en las jornadas de debate abierto que posteriormente se realizaron para la elaboración del II Plan canario y en las que participaron, además de las instituciones, diferentes colectivos y asociaciones de mujeres canarias. En definitiva, aunque no sobre el aprendizaje de la evaluación ni del propio plan, la experiencia de evaluación canaria contribuyó a dar una cierta visibilidad a las políticas de igualdad canaria y dio pie, en cierto modo, a la elaboración del II Plan. Por lo tanto, aunque no en sí misma, la evaluación sirvió de excusa para otras actuaciones y cumplió su papel de rendimientode cuentas, sobre todo en lo referente a las restantes instituciones y departamentos canarios que no eran el propio ICM. En cierto modo, la evaluación del I plan canario tuvo un cierto uso persuasivo (cuando se utiliza para concitar apoyos para

el plan evaluado), al haber sido utilizada para aportar una mayor visibilidad a las políticas de igualdad y al Instituto Canario de la Mujer, y como reclamo a los diferentes colectivos para la elaboración del II plan.

En el caso de la evaluación del **II Plan castellano-manchego** resulta difícil la valoración del impacto de esta experiencia de evaluación, debido a que ésta es la más reciente en el tiempo¹¹. En cualquier caso, al analizar el impacto general de este estudio de evaluación, hay que tener en cuenta el prácticamente solapamiento temporal del estudio de evaluación del II Plan y el proceso de elaboración del III. El hecho de que fuera el mismo equipo técnico externo el que dirigiera ambos procesos (evaluación del II y proceso de elaboración del III), facilita el hecho de que el primero tuviera un cierto impacto sobre el segundo a pesar de su práctica simultaneidad. Esta misma circunstancia, sin embargo, creemos que no facilita el “aprendizaje” de la institución (Dirección General de la Mujer y otras instituciones autonómicas implicadas), ni del proceso de evaluación en sí mismo ni de sus resultados, ya que la institución impulsora de los planes ni siquiera tiene que “utilizar” la evaluación para la mejora del plan siguiente, ya que la propuesta del Plan siguiente se hace también por parte de un equipo externo. En definitiva, creemos que en este caso el impacto potencial de la evaluación es mayor para las políticas de igualdad que para su organismo impulsor. En este sentido, el equipo externo que realiza la evaluación, como es lógico, señala haber utilizado información de la evaluación para dirigir el proceso de elaboración del II Plan y para elaborar la propuesta que se presenta. Este impacto de la evaluación no sólo hay que buscarlo respecto al plan, sino también respecto a la propia evaluación, ya que, y también como es lógico, la propuesta de evaluación que aparece en el III Plan está claramente influida por la experiencia de evaluación del II Plan y esto hace que sea una propuesta bastante más desarrollada que lo que suele aparecer en otros planes, incluyendo un listado de “indicadores mínimos propuestos”. Es decir, la experiencia de evaluación del II Plan claramente sirve para proponer la evaluación siguiente.

El impacto de la evaluación del **I Plan castellano leonés** parece haber sido alto en opinión de su responsable política, aunque este impacto ha de ser matizado. En general, la “evaluación” resulta útil no como evaluación sino como estudio de diagnóstico de la situación de las mujeres en Castilla y León. No se tenía ningún dato al respecto y, por lo tanto, la información resulta muy útil para la formulación del II Plan, ya que el estudio ayuda a que se conozca mejor el problema público que se quiere abordar. Pero realmente no aporta ninguna información relacionada con la intervención pública que supuestamente se está evaluando por lo que resulta prácticamente imposible que tenga impacto sobre el rendimiento de cuentas; sobre la mejora de los planes y la iluminación de acciones futuras en el sentido de, como se ha dicho anteriormente, aportar un mayor conocimiento sobre el problema que no sobre el plan.

¹¹ Los mayoría de los datos que se obtienen de la Dirección General de la Mujer de Castilla-La Mancha (entrevista a su directora general y una jefa de servicio, datos sobre el seguimiento del II Plan, etc.) previamente a la realización de la evaluación. Sin embargo, decidimos incluir esta experiencia de evaluación en el estudio por su interés y características de evaluación externa.

Es decir, el estudio realizado tiene un cierto impacto tanto en lo que se refiere a las políticas de igualdad (ofrece datos sobre el problema a abordar) como también al organismo de igualdad, ya que el hecho de plantearse la evaluación les hace detectar que los planes deben ser más concretos y realistas para que puedan ser evaluados y que deben implantar un sistema continuo de seguimiento y evaluación para el plan siguiente. Así se detectó a raíz de la evaluación, lo ambicioso y genérico del I Plan, y lo difícil de evaluar, por lo que se subsana ese problema en el II que pasa a ser “absolutamente concreto, realista y temporalizado” y al que se considera que “va a ser un Plan muy fácil de evaluar”. Lo curioso es que se detecta ese problema y de hecho se cambian cosas para hacerlo más evaluable, pero en ningún momento se es consciente de lo que realmente significa hacer una evaluación, ya que llaman “evaluación” a un tipo de estudio que claramente no lo es. Incluso cuando se reconoce por parte de la responsable política que es difícil hacer una conexión entre los cambios que demuestra la evaluación que ha habido en las mujeres de Castilla y León en los años de vigencia del Plan y las acciones del Plan, parece que la causa de que no se pueda hacer dicha conexión es lo poco concreto y realista que era el I Plan y la ausencia de indicadores para evaluarlo, no el hecho de que el objeto de estudio sea otro (el problema de la desigualdad y no la intervención sobre el mismo).

Lo que sí es evidente es que la información del estudio realizado, aunque no sea evaluación, sirve para conocer mejor el problema sobre el que hay que intervenir, y que la experiencia de evaluación genera una voluntad política clara (al menos en cuanto intenciones)¹² sobre la función de evaluación; así cuando la responsable política se refiera a como va a ser la evaluación del II Plan dice textualmente que “va a ser una evaluación continua” y que “estamos trabajando en la elaboración de un sistema informático” (en realidad se refiere al seguimiento continuo del Plan, más que a la evaluación).

La evaluación del **II Plan gallego** no parece haber tenido tanto impacto como hubiera sido el esperado y el deseado. Existe un déficit de difusión de los resultados de la evaluación y esto puede deberse, en parte, al cambio de directora de la Agencia de Igualdad durante el proceso de evaluación, y al hecho de que la persona que recibe el informe final es otra distinta a la que había encargado dicho trabajo de evaluación y, en este caso, parece que con una idea distinta sobre el papel que puede y debe cumplir una evaluación¹³.

¹² La entrevista con la responsable política de la Secretaría Regional de la Mujer de Castilla y León se realiza el 9 de julio de 1997 cuando estaba recién aprobado el II Plan de Igualdad y todavía no se había puesto en marcha.

¹³ El cambio de responsables políticos en mitad de los procesos de evaluación es un factor de riesgo especialmente en países con poca cultura evaluativa, ya que normalmente los responsables políticos no tienen una idea preconcebida de los procesos de evaluación y la van construyendo a lo largo de la realización del estudio. La tarea de explicación, persuasión y convencimiento sobre el sentido y significado de la evaluación por parte de los equipos técnicos que llevan a cabo la evaluación es, por lo tanto, muy importante. Así, si el cambio de responsable se da hacia el final del proceso de evaluación - como es en este caso- el equipo técnico no tiene muchas veces la oportunidad de ese trabajo clave para el éxito de una evaluación.

Sin embargo, y debido a cómo se planteó el proceso de evaluación desde un principio, implicando a la “Comisión de Evaluación” que era una comisión en el nivel técnico en el que estaban representadas las diferentes instituciones autonómicas que tenían algo que decir en el Plan, y debido al trabajo y el esfuerzo que en ello se invirtió, sobre todo en lo referido a las fichas de recopilación de información, creemos que la evaluación del II Plan tuvo un cierto impacto en lo que se refiere al seguimiento y a la evaluación del Plan siguiente, generando un sistema de seguimiento y un “aprendizaje” sobre evaluación de los diferentes departamentos e instituciones. Este “aprendizaje” puede considerarse especialmente importante por parte de la propia Agencia de Igualdad, ya que tuvo a dos técnicas muy implicadas desde un principio en el proceso de evaluación. Además, en la intervención que tuvo la entonces secretaria general de la Agencia en una mesa redonda en el Seminario de Evaluación de Políticas de Igualdad en la Universidad Menéndez Pelayo en verano de 1998 (unos meses después de que se entregara el informe final de evaluación), se expusieron una serie de conclusiones finales relacionadas con la propia evaluación (importancia de la evaluación, necesidad de que la gente que participa tenga conocimientos de evaluación, necesidad de estandarizar un modelo periódico de recogida de información, necesidad de estructuras externas que cooperen en el seguimiento y evaluación, y dar tiempo para la valoración antes de comenzar otro plan). De hecho, la propia presentación de la experiencia de evaluación gallega en dicho Seminario ya es una forma de impacto en este sentido.

Por lo demás, desconocemos si los resultados de la evaluación han tenido realmente impacto en la elaboración del III Plan y en una mayor comprensión de las políticas de igualdad por parte de quienes tienen la responsabilidad de impulsarlas. Sin embargo, teniendo en cuenta la poca difusión dada a la propia evaluación, creemos que este impacto es muy bajo.

El impacto del estudio que se realiza para evaluar el **II Plan madrileño** creemos que, como evaluación que pretende la mejora de las políticas y el rendimiento de cuentas, es muy bajo. Sin embargo, aunque no como evaluación, el estudio realizado viene a cubrir una carencia existente de estudios que analizaran la situación social de las mujeres madrileñas. Es decir, incluso cumple en cierto modo un papel de ayuda a la formulación de políticas de igualdad de género (ya que ayuda a conocer mejor el problema -la desigualdad por motivos de género en un determinado territorio- sobre el que hay que intervenir).

Por otro lado, al publicar un resumen de los resultados del estudio, éstos han tenido una difusión alta entre los colectivos de mujeres de la Comunidad de Madrid. Pero estos, por su propia naturaleza no han podido tener mucho impacto en una mayor comprensión de las políticas de igualdad (quizá lo han tenido algo respecto a una mayor comprensión del problema a abordar), ni tampoco sobre la propia función de evaluación. Dicha función no ha podido verse fortalecida ni normalizada ni como consecuencia del estudio se ha aprendido más sobre evaluación. Por el contrario, sorprende la falta de crítica, tanto interna como externa, que ha recibido dicha evaluación; este hecho es otro indicador más de la falta de cultura evaluativa en las administraciones públicas españolas. De hecho, la evaluación del III Plan madrileño se plantea en los mismos términos que la del II, por lo que podemos concluir que se ha dado más bien poco

aprendizaje en evaluación. Otra cosa es que el estudio podría ser una excelente base (y por lo tanto podría tener un gran impacto en ello) para desarrollar un observatorio regional de igualdad que estudiara de forma periódica la situación de la mujer madrileña, al estilo del trabajo que el IM realiza con “Mujeres en cifras”. Sin embargo, esto no debería obviar la necesidad de evaluación de los planes y actuaciones.

Aún considerando la absoluta necesidad e importancia de obtener información actualizada sobre el colectivo destinatario sobre el que se pretende intervenir y el problema que se quiere abordar, parece que, en este caso, los resultados del estudio de evaluación no causan un especial impacto sobre las políticas de igualdad madrileñas, o al menos sobre la última responsable de impulsarlas, ya que al entrevistar a la Directora General y preguntarle sobre qué datos de la evaluación le habían sorprendido y cuáles se habían utilizado en concreto para la propuesta del II Plan, no recordaba ninguno¹⁴.

3.2. El impacto de los resultados y los procesos de evaluación mixta

La experiencia de evaluación del **II PIOM estatal** tiene como efecto una cierta normalización de la función de evaluación en el IM. Al menos, no se cuestiona la evaluación del III PIOM (de hecho se refleja ya un cierto compromiso de evaluación en el propio Acuerdo de Consejo de Ministros que lo aprueba), y existe un equipo técnico que se ocupa de ello. Esto tiene alguna relación con la también normalización de la tarea de producir “Mujeres en cifras”, publicación que produce la Subdirección de Estudios y Documentación y que refleja periódicamente (aproximadamente cada dos años) la base de datos sobre mujer que se renueva de forma permanente. “Mujeres en cifras” lleva produciéndose desde el año 1987, por lo que la “normalización” de su producción es muy anterior a la evaluación. Sin embargo, tanto el departamento como los responsables técnicos de ambas tareas son los mismos los equipos que realizan una y otra también. La “normalización” de la tarea de evaluación se produce cuando se incorpora en el III Plan la evaluación en los mismos términos que se había realizado en el II. Este mismo hecho, sin embargo, es el que nos lleva a interpretar que se ha dado poco aprendizaje sobre la evaluación en torno a los PIOM estatales y el IM. Creemos que el Instituto ha sido muy poco autocrítico con su propia evaluación y, por lo tanto, no ha estado abierto al aprendizaje que la experiencia de evaluación le hubiera podido aportar. Si esto hubiera sido así, se hubieran incorporado mayores cambios y mejoras más sustanciales a un diseño de evaluación que era manifiestamente mejorable.

Sí es cierto que el hecho de haber evaluado el II PIOM hace que el IM haya adquirido cierto “liderazgo” en evaluación de políticas públicas de igualdad de género,

¹⁴ Entrevista a la Directora General de la Comunidad de Madrid el 21/8/97, ocho meses después de la entrega del Informe final y cuatro meses después de la publicación del resumen.

aunque éste sea ciertamente débil¹⁵. Sin embargo, creemos que este liderazgo se produce más por su posición estratégica del Instituto respecto a las comunidades autónomas que por un verdadero “liderazgo metodológico”. Es más, y como se desarrollará más detenidamente en los capítulos siguientes de este trabajo, creemos que el IM hubiera podido cumplir un mayor “liderazgo metodológico”, y que hubiera podido consistir en generar metodología de evaluación, coordinar y ayudar a intercambiar las experiencias de evaluación en las diferentes comunidades autónomas, crear espacios de encuentro, intercambio y reflexión, etc. En nuestra opinión, sin embargo, el IM ha dejado pasar esta oportunidad y, en este momento, el “liderazgo metodológico” en evaluación es mucho mayor en algunas comunidades autónomas que en el propio IM.

El diseño de la evaluación del II PIOM estatal no permite tener una visión integral del diseño, proceso y resultados del Plan, por lo que resulta difícil que se haya generado una mayor comprensión de las políticas de igualdad de género en ese sentido. Con la “evaluación de la implementación” simplemente se tiene una idea de qué se ha hecho bajo el paraguas del Plan. Con la “evaluación cualitativa” nos hacemos una cierta idea del estado de opinión general de ciertos colectivos respecto al mismo y con la “evaluación cuantitativa” nos hacemos una idea de cómo evolucionan los usos del tiempo, las actitudes y los valores de los españoles y españolas respecto al género; es más creemos que la “evaluación cuantitativa” es una excelente y rigurosa investigación sobre cómo medir la evolución de la sociedad española respecto a cuestiones tan “intangibles” pero tan cruciales como las estereotipias de género, los valores y las actitudes. Y en ese sentido creemos que aporta una mayor comprensión sobre la medición de estos extremos, cuestión sumamente importante en el análisis del problema de la desigualdad de género y su evolución. En este caso, aunque aporta una mayor comprensión sobre la situación de desigualdad, no aporta una mayor comprensión sobre las políticas de igualdad porque no son su objeto de estudio, ni las evalúa en ningún caso. La responsable política del IM, en la entrevista que realizamos sobre la evaluación y ante la pregunta sobre qué habían aprendido con la evaluación y para qué les había servido en este sentido, no parecía tener una visión global de la aportación de la evaluación del II PIOM a la mejor comprensión del Plan y de la política de igualdad, e insistía más en la idea abstracta de la necesidad de evaluar y del necesario rendimiento de cuentas.

Respecto a las acciones llevadas a cabo en los planes como consecuencia de la evaluación, creemos que éstas no han ocurrido, al menos como consecuencia directa e inmediata de la evaluación. Entre otras cosas, porque no ha existido una traducción práctica de la evaluación hacia sugerencias o recomendaciones a llevar a cabo para la mejora de las diferentes actuaciones o del plan en su conjunto.

¹⁵ El único indicador tangible de este liderazgo en evaluación es la celebración de un encuentro de dos días y medio sobre “Evaluación de las políticas de igualdad entre mujeres y hombres” en la Universidad Internacional Menéndez Pelayo, entre el 20 de Junio y el 1 de Julio de 1998, patrocinado por el IM y cuyos directores fueron Rocío Fdez. Ballesteros, catedrática de la Universidad Autónoma de Madrid y Mariano Álvaro, Consejero Técnico de Estudios del IM y protagonista técnico de la evaluación del II PIOM estatal. Este seminario fue el resultado de las peticiones sobre seminarios, encuentros o reuniones sobre evaluación que algunas comunidades autónomas formulaban en las Conferencias Sectoriales sobre Asuntos Sociales.

La evaluación del **II Plan catalán** parece tener un impacto importante tanto en el Instituto Catalán de la Mujer como en las políticas de igualdad catalanas, en opinión de su responsable política. Así como nosotros destacaríamos el impacto real que el proceso de evaluación ha producido en el ICD, matizaríamos en cierto modo el producido en las políticas de igualdad, aunque por supuesto creemos que realmente sí ha habido impacto y que el proceso de evaluación ha servido para que las responsables de impulsar las políticas de igualdad de género en el ICD se hayan cuestionado y reflexionado sobre los planes (por ejemplo, respecto a los destinatarios de los mismos) y que se hayan tomado decisiones respecto al III Plan debidas a la evaluación. Nuestra matización tiene que ver con la constatación, por un lado, de que el III Plan ha sido elaborado en un proceso en que la evaluación del II, como es lógico, ha sido un factor entre los muchos que se han tenido en cuenta. En las propias palabras de la responsable del ICD, “hubo una primera negociación con Departamentos, la evaluación, el IV Programa de Acción Comunitaria, la Plataforma de Pekín y las conferencias que se hicieron (en el marco de la universidad de verano y con motivo de la celebración de los 10 años de las políticas de igualdad), las sugerencias que ya salieron ahí, se elaboró el primer borrador, se mandó a las ONG, han mandado enmiendas, se han incorporado y se ha vuelto a mandar el documento a los Departamentos que han vuelto a hacer enmiendas”. Por otro lado, el hecho de que la evaluación no tuviera una gran divulgación nos hace interpretar que entre los destinatarios primarios de la evaluación no estaban las ONG y colectivos de mujeres; creemos que ni siquiera entre los restantes Departamentos del Gobierno de la Generalidad ha habido una divulgación específica del estudio de evaluación, aunque sí probablemente estos Departamentos estaban contemplados al menos como destinatarios secundarios y probablemente recibieran información sobre la misma. De hecho, la evaluación ha cumplido una función importante de petición de cuentas a los Departamentos implicados. Con esto queremos decir que parece que las ONG y otros colectivos, e incluso los propios Departamentos de la Generalidad no han tenido las mismas oportunidades que el ICD de “aprender” de la evaluación antes de hacer sus aportaciones al III Plan; en definitiva, el impacto de la evaluación en las políticas hubiera podido ser probablemente mayor si se hubieran comunicado mejor los resultados de la misma.

En cualquier caso, la evaluación del II Plan catalán parece haber tenido efectos en el fortalecimiento y apreciación de la función de evaluación. Esta función de evaluación se considera clave para planes posteriores, incluido que se piense en términos de indicadores de evaluación y asignación presupuestaria concreta por parte de cada Departamento. También hay una referencia explícita a una “evaluación de impacto” en el III Plan, cuestión que no se había podido plantear en la evaluación del II, y que se considera que será posible al haber planificado previamente con unos indicadores de actividad y de evaluación.

Respecto al impacto en una mayor comprensión de las políticas de igualdad, la evaluación del II Plan también cumple en cierto modo su papel, ya que ayuda a tener una visión global del II Plan de actuación respecto a una serie de dimensiones (peso por área, tipo de actividad, agentes, destinatarios, presupuesto destinado) que ayuda a revisar y reflexionar sobre el Plan. Es decir, al evaluarse de alguna forma el diseño del Plan en su conjunto, es más fácil que haya un cierto aprendizaje y una mayor comprensión sobre la forma de intervenir para combatir la desigualdad de género. De hecho, las responsables del ICD reclaman haber hecho una reflexión sobre los

colectivos destinatarios a los se dirigen, llegando a la conclusión de que les resulta más difícil llegar al colectivo juvenil.

Como acciones concretas debidas a la evaluación, estarían las referidas a la estructura y planificación del plan siguiente, como son la introducción de un mayor compromiso explícito en evaluación y la inclusión de indicadores de actividad y de asignaciones presupuestarias por parte de los diferentes Departamentos. Las otras cuestiones que se reclaman por parte de las responsables del ICD como útiles, como la reflexión sobre los colectivos destinatarios y el poder pedir cuentas a los diferentes Departamentos, no tenemos constancia de que se hayan concretado en acciones concretas.

En lo que se refiere a la evaluación del **II PAPME vasco**, hay que constatar aquí la actitud de aprendizaje constante que tienen en *Emakunde* respecto a la evaluación. Esta es una de las comunidades autónomas que ha demostrado un interés más sincero por nuestra investigación, siendo paradójicamente una de las que más podía “ofrecer” en el tema de evaluación. La evaluación ha sido definida como un tema prioritario en la institución, y esto ha tenido un impacto fuerte que trasciende el propio *Emakunde*, probablemente hacia otras instituciones vascas y también hacia otros organismos e igualdad de otras comunidades autónomas. Por lo tanto, estamos ante la experiencia analizada con un mayor impacto en lo que se refiere a los efectos sobre la propia evaluación, tanto en una mayor apreciación de la misma y en un aprendizaje claro sobre ella, como en cuanto a la normalización y el fortalecimiento de la función de evaluación. Esta capacidad de aprendizaje de evaluación en la misma institución, se puede constatar no sólo en los esfuerzos invertidos por *Emakunde* en los procesos de evaluación, sino también en los procesos de reajuste y perfeccionamiento del propio sistema de evaluación que se aprecia de la evaluación del I PAPME al II.

Además, las memorias anuales presentadas al Parlamento Vasco sirven para rendir cuentas y han tenido un cierto uso persuasivo, además de que han aportado una gran visibilidad a *Emakunde* como institución y a las políticas de acción positiva y de igualdad de género. También el uso conceptual ha sido relativamente alto, ya que las evaluaciones y especialmente, las tipologías de actuaciones desarrolladas en las mismas han permitido a las responsables de las políticas reflexionar sobre las mismas y saber dónde se encuentran sus puntos débiles.

Finalmente también se ha dado un cierto uso instrumental, en el sentido de que se han llevado a cabo algunas acciones como consecuencia de la evaluación, especialmente en lo referido a la coordinación con otros departamentos y las otras administraciones públicas vascas (Diputaciones y Ayuntamientos), para la puesta en práctica y gestión del plan y la propia recopilación de información.

3.3. El impacto de los resultados y los procesos de evaluación interna

En el caso de la evaluación interna del **I Plan aragonés**, la experiencia de evaluación, además de ayudar a situarse a las responsables políticas que la emprenden, dadas las especiales circunstancias ya comentadas (cambio de responsables y vacío de poder durante más de cinco meses), sirve también para iniciar, de forma más sistemática, y desde el Instituto Aragonés de la Mujer, la recopilación de información sobre los planes. Esta experiencia en evaluación va a servir a las responsables técnicas del IAM a enfrentar posteriormente la tarea de hacer el seguimiento y evaluar el II Plan aragonés.

También va a ayudar a tener una idea general y global sobre lo que se ha hecho -y no se ha hecho- al amparo del I Plan aragonés de igualdad. Aunque este impacto sobre una mayor comprensión de las políticas de igualdad no sea especialmente alto, sí se da en cierto modo y merece, por lo tanto, una cierta consideración.

Respecto a las acciones o decisiones que se llevaron a cabo como consecuencia de la evaluación habría que mencionar la reorientación del propio I Plan durante el año 1996 (la evaluación se hizo con un cierto carácter formativo), la “reafirmación” en los pesos relativos otorgados a determinadas acciones (las referidas a empleo), y la tendencia a definir actuaciones más concretas en el II Plan e intentar preestablecer unos indicadores de evaluación según cada una de las actuaciones. Es decir, para lo limitado de la experiencia de evaluación analizada, su impacto en este sentido, aunque no sea muy alto, es considerable.

CAPÍTULO 8

EL DISCURSO SOBRE LA FUNCIÓN DE EVALUACIÓN EN TORNO A LAS POLÍTICAS DE IGUALDAD DE GÉNERO

INTRODUCCIÓN

En la parte anterior (parte III) se han descrito y analizado las once experiencias de evaluación objeto de estudio, según una serie de criterios establecidos en torno a tres ejes: el planteamiento de las evaluaciones (contexto, propósitos, tipología, estructuras y recursos), los elementos claves de dichos procesos (agentes, momentos, criterios y procedimientos), y su posible utilización. En esta parte IV se realiza una valoración global de dichas experiencias en su conjunto.

En esta parte IV se incluye también, en este capítulo ocho, la valoración que realizan las responsables políticas de los organismos de igualdad sobre las experiencias de evaluación en concreto, y en general, de la función de evaluación de las políticas de igualdad de género en el seno de dichos organismos. Además de cómo se hayan realizado de hecho las evaluaciones, resulta imprescindible para una valoración completa, captar la perspectiva institucional frente a dichas experiencias de evaluación, comprender de una forma más global el contexto en el que se generan, así como saber cómo se entiende la función de evaluación y el papel que cumple en los organismos de igualdad. Al comprobar que eran prácticamente inexistentes los documentos institucionales o declaraciones explícitas que, en los organismos de igualdad, trataran el tema de la evaluación, la mejor forma de obtener dicha perspectiva fue a través de la captación del discurso que sobre la evaluación tengan sus máximas responsables.

Por lo tanto, el discurso de las personas al frente de los organismos de igualdad representa la perspectiva institucional que en dichos organismos se tiene ante la evaluación. Pero además, y por otro lado, también resulta crucial la posición y la actitud de dichas responsables ante la evaluación, ya que estos organismos son los principales protagonistas e impulsores de los procesos de evaluación de los planes de igualdad. De esta manera, el liderazgo que pueda ejercer la máxima responsable de los organismos encargados de la evaluación es importante para entender el desarrollo de la función de evaluación de las políticas de igualdad de género en un determinado ámbito. La importancia de este liderazgo se acrecienta ante la ausencia de una cultura desarrollada de evaluación: la posición que tome la responsable política del organismo de igualdad ante la evaluación, la comprensión de su significado y sus posibilidades y la importancia relativa que se le dé respecto a otras funciones en la institución, resultará clave para analizar y comprender los procesos de evaluación que se han dado. De hecho, de dicha posición institucional ante la evaluación depende, en gran parte, que se comience a instaurar una cultura de evaluación en torno a las políticas y en los organismos. Como se ha comentado ya anteriormente, esta posición

institucional ante la evaluación se puede captar, sobre todo, a través del discurso de sus responsables políticas y, en parte también, del de las responsables técnicas de los procesos de evaluación, en caso de que éstos se hayan dado.

Se realizaron entrevistas abiertas semiestructuradas a 16 directoras de organismos de igualdad, de las cuales once lo eran de los organismos que llevan a cabo los procesos de evaluación analizados. En realidad, se entrevista a todas las responsables políticas de los organismos de igualdad autonómicos, con la excepción de dos de ellos (el navarro y el cántabro), además de a la directora del IM central. Todas las entrevistas se realizan en los años 1997 y 1998¹. El objetivo de las entrevistas era la captación del discurso político en torno a los procesos de evaluación: qué papel cumple la evaluación de las políticas de igualdad en los organismos que las impulsan y para las propias políticas, qué esperan de y cómo utilizan la evaluación, cuáles son y han sido sus principales dificultades, etc. Además de los concretos procesos de evaluación metaevaluados, también se captó el discurso sobre la evaluación de las políticas de igualdad de género en general y sobre la función de evaluación en sus respectivos organismos. Es por esta razón por la que se incluyeron también los discursos de las responsables políticas de los organismos autonómicos en los que no existían experiencias de evaluación de segunda generación.² En definitiva, no se trataba de recopilar información sobre los procesos de evaluación en sí (para esto se recopilaba información de los equipos técnicos y de los documentos existentes), sino de saber las percepciones que sobre la evaluación tenían -y desde un punto de vista político y no técnico- las máximas responsables de los procesos de evaluación. Lo que se refleja en este capítulo es, esencialmente, el análisis del discurso extraído en dichas entrevistas, contrastándolo, en algunos casos, con las entrevistas estructuradas (en algunos casos presenciales y, en otros, telefónicas) mantenidas con el personal técnico encargado de las evaluaciones en los organismos de igualdad.

¹ Véase en el anexo correspondiente la fecha de las entrevistas y los nombres de las entrevistadas.

² Los guiones de las entrevistas aparecen también en el anexo.

1. EL DISCURSO SOBRE LA EVALUACIÓN Y SOBRE LAS POLÍTICAS DE IGUALDAD DE GÉNERO DE LAS RESPONSABLES POLÍTICAS

En general, se aprecia que el discurso sobre la evaluación por parte de las responsables políticas de los organismos de igualdad no está muy elaborado, es decir, no es un discurso previamente construido ni que haya sido explicitado formalmente. Esto parece normal si se tiene en cuenta que la práctica evaluativa es muy novedosa y que, además, no existe cultura de evaluación ni referencias teóricas o prácticas en torno a las políticas de igualdad y, por lo general, ni siquiera tampoco en el entorno institucional de las administraciones públicas españolas.

Lógicamente, el discurso sobre su evaluación está inextricablemente ligado al de las políticas de igualdad de género y éste, a través de los objetivos y estrategias de dichas políticas, a los conceptos de acciones positivas, perspectiva de género, transversalidad de las políticas, etc. De hecho, en muchos casos, la falta de discurso elaborado sobre evaluación se suple con uno sobre las políticas que se pretenden evaluar. Es decir, en muchas ocasiones, aunque el foco de las entrevistas era la evaluación de las políticas, el discurso derivaba hacia las propias políticas.

Otra conclusión general del análisis realizado es que no se puede hablar de un discurso común o monolítico ante la evaluación. Esto es una consecuencia clara del hecho de que el discurso sobre la evaluación no está muy elaborado y ha sido construido en parte sobre la marcha. Sin embargo sí se pueden encontrar algunas ideas comunes y tendencias en los posicionamientos, que tratamos más adelante en este capítulo. Por otro lado, resulta interesante analizar las diferencias discursivas. Efectivamente, una de las primeras cuestiones a destacar en el análisis de las entrevistas es el hecho de que hemos encontrado ciertas diferencias entre los discursos captados de las 16 responsables entrevistadas, y que dichas diferencias pueden ser categorizadas, de forma general, en cinco polos o factores:

1) Las responsables en cuyo organismo se ha hecho evaluación y ellas han tenido implicación en ello vs. las responsables en cuyo organismo no se ha hecho evaluación o ellas no han tenido implicación en ello

Las responsables entrevistadas de los organismos de igualdad andaluz, aragonés, canario, castellano-leonés, catalán, gallego³, riojano y vasco son las que encargaron (o establecieron el mandato de) las evaluaciones metaevaluadas e hicieron un seguimiento de las mismas.

³ En el caso gallego, la responsable entrevistada fue la persona que encargó la evaluación del II Plan, aunque no la misma que recibió el informe en marzo de 1998, fecha de su finalización.

En cinco casos (Asturias, Baleares, Extremadura, Murcia y Comunidad Valenciana) no se realizó evaluación de la llamada “segunda generación” de planes o la que se hizo no puede ser calificada como tal⁴. En el caso de Castilla-La Mancha se hizo una evaluación, pero posteriormente a la entrevista realizada a su responsable, es decir, en el momento de la entrevista, la responsable todavía no tenía experiencia en evaluación en su organismo. Por último, tanto en el caso del IM central como de la Dirección General de la Mujer la Comunidad de Madrid, sus responsables llegan al cargo cuando el proceso de evaluación está ya iniciado⁵. Por lo tanto, en ambos casos, el liderazgo inicial del proceso de evaluación se encuentra más en el personal técnico de los organismos que son quienes realmente impulsan los respectivos estudios de evaluación.

Como es lógico, el discurso sobre evaluación de las responsables implicadas e “iniciadoras” de las experiencias analizadas, es más elaborado que el resto, es decir, se deja entrever que dichas responsables, en el momento de las entrevistas, habían pensado más y de forma más profunda en el papel que cumple o puede cumplir la evaluación de las políticas de igualdad.

2) Las responsables con más tiempo en el cargo vs. las que llevan menos tiempo en él

También se aprecian diferencias en las responsables que llevan más tiempo en el cargo. De las 16 personas entrevistadas, tres -las directoras de los Institutos Andaluz, Catalán y Vasco- llevaban casi diez años en el cargo en el momento que fueron entrevistadas. Ninguna de las restantes asume su cargo antes de 1995 y en algunos casos llevaban poco tiempo en el cargo en el momento de la realización de la entrevista. Esta diferencia de antigüedad marca un discurso distinto, más elaborado no sólo respecto a la evaluación sino también respecto a las políticas de igualdad, la perspectiva de género, etc.

Por lo tanto, con alguna excepción, también tiende a ser más elaborado el discurso de las responsables que llevan más tiempo en el cargo.

3) Las responsables que tienen un pasado político anterior, especialmente relacionado con organizaciones feministas vs. las que no lo tienen

Parece que la experiencia política y/o feminista previa al cargo también influye en el discurso de las responsables. Además, así como las primeras directoras de los organismos de igualdad se destacaban por tener experiencia previa en organizaciones feministas (por lo tanto las que llevan más tiempo en el cargo suelen tener esta característica), no es así con las directoras de los organismos de igualdad a partir de

⁴Véanse los criterios de selección de experiencias a metaevaluar expuestos en el capítulo cuatro.

⁵ En el caso de la Comunidad de Madrid, aunque el concurso público para la adjudicación de la primera parte del estudio de evaluación se falla en Noviembre de 1996 (cuando ya está incorporada Asunción Miura, Directora General bajo cuyo mandato se realiza la evaluación), el proceso ya había sido técnicamente iniciado anteriormente -a principios del año 1996-.

1995, especialmente si los partidos que asumen el gobierno son de corte conservador, ya que entre las filas o simpatizantes del Partido Popular es más difícil encontrar mujeres que tengan una militancia feminista.

En este caso, pues, las directoras que han tenido un pasado de militancia feminista tienen un discurso sobre las políticas de igualdad de género -y esto también se traduce en el discurso sobre la evaluación de las mismas- que el resto.

4) Las responsables en gobiernos de otros partidos vs. las responsables en gobiernos del partido conservador

En el período en el que se realizaron las entrevistas (entre mayo de 1997 y septiembre de 1998), tanto el gobierno central como el de nueve de las 15 comunidades autónomas a cuyas responsables de los organismos de igualdad se entrevistan, están a cargo del Partido Popular (Aragón, Asturias, Baleares, Castilla y León, Galicia, Madrid, Murcia, La Rioja y Comunidad Valenciana)⁶.

Aunque las diferencias son menores de lo que cabría esperar, ya que el discurso pro igualdad de género se ha ido generalizando en todas las tendencias políticas incluidas las conservadoras, se encuentran discursos diferenciados en el sentido de que las responsables de organismos cuyos gobiernos son conservadores tienen un discurso menos elaborado, más “reciente” y algo más “ingenuo” respecto a las políticas de igualdad de género. En cualquier caso, tienen un discurso poco conservador respecto al género y esto, y en algunos casos se reconoce, les trae más de algún problema en el seno del partido al que representan. En definitiva, son las encargadas de impulsar políticas que son progresistas en el nombre de un partido conservador y esta circunstancia, de por sí, puede ser contradictoria.

“Porque además, y esto te lo digo... grábalo o no lo grabes, como quieras, pero el estar en determinado grupo político, también influye mucho. (...) No hay que engañarse ¿eh?, hay un cierto sector que es muchísimo más reacio. Entonces para ese sector estás haciendo demasiado, y para el otro grupo, te estás quedando corto. ¿Cuándo lo haces bien?”

En el discurso de las responsables conservadoras, la edad aparece como un factor clave a la hora de encontrar apoyo o comprensión de las políticas que están llevando a cabo. Para ellas, en su entorno político hay diferencias según la edad: la gente joven de sus partidos -casi todas ellas lo son- tiene más asumido el discurso de la igualdad que la gente mayor. Esto lleva a que en sus discursos aparezca a menudo la afirmación de que ellas no han sido discriminadas porque son jóvenes y es distinto, aunque se dejen entrever algunas contradicciones al respecto.

⁶ El gobierno canario es una coalición entre el PP y Coalición Canaria. En el caso del Instituto Canario de la Mujer -adscrito a la Consejería de Empleo y Asuntos Sociales- está dirigido por una mujer perteneciente a Coalición Canaria. Los gobiernos de las comunidades autónomas de Aragón, Asturias y Baleares cambian de signo político en las elecciones de 1999, pasando a ser coaliciones de partidos autonómicos con el PSOE.

“Pues decirte que hay una parte de, de... no sé como decírtelo... experiencia personal, ¿no?, porque habiendo muchas diferencias, porque las hay, por ejemplo mi caso personal, yo no creo que a mí me hayan discriminado nunca, pero es que yo posiblemente o tú en este caso, a lo mejor somos gente privilegiada”

Y más tarde, la misma responsable ante la pregunta de si se ha sentido alguna vez discriminada:

“Muchas, muchísimas. Y salir de reuniones con impotencia y dando un puñetazo encima de la mesa muchísimas. Muchas, la mayoría. Cuesta mucho, es que cuesta mucho. Cuesta mucho”

En cualquier caso, parece que las responsables conservadoras entrevistadas se han enfrentado, por primera vez, a una determinada forma de ver y definir un problema público en el momento en el que asumen el cargo y, a lo largo de su experiencia en él, van asumiendo un discurso que realmente no se diferencia tanto del de otras responsables representantes de otros partidos políticos. De hecho, la mayoría de las responsables (de diferentes tendencias ideológicas) piensa que las actitudes ante la desigualdad de género dependen más de factores personales que ideológicos, incluida la idea de que no necesariamente por ser mujer alguien debe tener una postura de género asumida⁷. Creemos que éste puede ser un posible indicador de cómo la posición de “femócrata” puede hacer asumir posiciones feministas que antes no estaban tan presentes o elaboradas. En definitiva, parece que a estas responsables les falta algo más de “rodaje” en una arena de alguna forma novedosa para el ámbito político en el que se mueven.

“Y entonces todo esto me sirve para de alguna forma integrar esfuerzo y si lo tenía yo claro cuando vine, ahora lo tengo todavía mucho más claro”

5) El bagaje profesional de las responsables políticas

Al ser la evaluación de programas y políticas públicas algo tan nuevo y reciente en las administraciones públicas españolas, y al no haber una formación específica sobre la misma, el bagaje profesional y de estudios de las responsables condiciona de alguna manera la forma de entender la evaluación.

En este caso, hemos encontrado alguna diferencia entre las responsables que vienen de un bagaje profesional relacionado con las ciencias sociales (más que las que vienen de las ciencias jurídicas, de la salud o las relacionadas con el mundo empresarial), que parecen entender algo mejor lo que significa evaluar programas públicos. Especialmente, si vienen del ámbito del trabajo social, por la tradición que

⁷ En más de un caso aparece en el discurso la idea de que hay “mujeres-hombres”, especialmente si han llegado a posiciones de poder que, por haber llegado a dichas posiciones, pueden llegar a ser más “machistas” que algunos hombres que, por factores personales, pueden ser más sensibles al discurso pro igualdad.

en España han tenido estos estudios de tratar el tema de la evaluación y de dar importancia a la documentación y gestión de la información de los programas, parecen tener un discurso algo más elaborado sobre la evaluación. Este es el caso de las responsables entrevistadas de Canarias y La Rioja, en ambos casos trabajadoras sociales.

Estas cuatro dimensiones o polos interactúan entre sí. Es decir, en algún caso, una responsable tiene un discurso muy elaborado sobre las políticas de igualdad de género porque, aunque no tenga mucha antigüedad en el cargo, tiene un pasado de militancia feminista y tiene muy incorporado el discurso de género, y además tienen un bagaje profesional que le permite tener una idea más aquilatada de lo que es la evaluación de programas. En general, respecto al discurso de la evaluación en concreto el primer polo o dimensión mencionada (haber estado directamente implicada en la evaluación) parece ser el más importante y para el discurso de las políticas de igualdad en general, el tercero mencionado, ya que haber tenido o estar cercana a una militancia feminista ofrece muchas más probabilidades de tener un conocimiento profundo sobre cuestiones de género.

Una vez expuestas las diferencias en los discursos y las causas que hemos identificado de las mismas pasamos a exponer las características comunes o tendencias que hemos encontrado. Para exponer este discurso de las responsables políticas sobre la evaluación de las políticas de igualdad de género, diferenciamos cuatro puntos. En primer lugar, las ideas asociadas a la evaluación además de sus actitudes hacia la misma, es decir, qué piensan sobre y cómo se sitúan ante la evaluación. Estas las agrupamos en siete ejes: expectativas mágicas, necesidad de la evaluación, propiedad para los organismos, confusión con análisis de la situación de la mujer, visión técnica, lo cuantitativo y el seguimiento. En segundo lugar, exponemos los porqués de la evaluación, es decir, la utilidad simbólica que las responsables políticas le dan a la evaluación (para qué creen ellas que puede servir, para qué les ha servido a ellas), así como sobre qué y cómo deben evaluarse las políticas de igualdad. En tercer lugar, se identifican y analizan qué cosas proponen que se deben evaluar y cómo. Y por último, también se exponen las dificultades percibidas por estas responsables a la hora de poner en práctica un proceso de evaluación de un plan de igualdad de género, ya que el constructo de “dificultades” es uno de los principales hilos conductores de su discurso.

2. LA EVALUACIÓN: IDEAS ASOCIADAS Y ACTITUDES ANTE LA MISMA

En este punto se analizan las ideas que las responsables entrevistadas tienen asociadas a la evaluación, así como las actitudes ante la misma. En su discurso hemos encontrado ciertas expectativas mágicas sobre la evaluación, contribuyendo así a un discurso simbólico sobre la necesidad de la evaluación, normalmente poco elaborado: la evaluación debe hacerse más por una necesidad incuestionable que por una conveniencia real. En general, la evaluación se concibe como algo sobre todo técnico y que está asociada de forma especial a la medición de los resultados de los planes evaluados, y además de forma numérica y cuantitativa. Finalmente, la evaluación de los planes es algo propio de los organismos de igualdad, aunque a veces se conciba sólo o bien como evaluación externa o bien se confunda con lo que sería el seguimiento de la gestión de los planes. A continuación, se desarrollan estas ideas más pormenorizadamente.

2.1. Entre las expectativas mágicas sobre la evaluación y una visión más realista

El discurso que sobre evaluación tienen las responsables políticas de los organismos de igualdad en el nivel central y en el autonómico contiene ciertas expectativas mágicas sobre evaluación, especialmente si en los organismos de igualdad no se tiene experiencia evaluadora. Es más, a mayor experiencia y a mayor implicación de la responsable política en la evaluación, visión más realista y ajustada de los procesos evaluativos.

De esta manera, para las responsables con menos experiencia en evaluación, cuando se trata de hablar sobre la evaluación, su potencia y posibilidades son múltiples: se habla de medición de impactos, de cómo la evaluación permite saber en qué se ha cambiado y, sobre todo, en la necesidad de la evaluación para la planificación de los planes subsiguientes. Sin embargo, para las responsables con mayor implicación en los procesos de evaluación, es decir, las que saben más sobre evaluación, tienen una visión mucho más ajustada y realista sobre la misma. Hablan mucho más abiertamente de las dificultades que entraña la evaluación y mencionan la práctica imposibilidad de la medición de impactos, o más sabiamente aún, no de la imposibilidad, sino de la poca pertinencia de invertir la cantidad de esfuerzos y recursos que ello conlleva cuando, a lo mejor, la información que arrojaría no sería tan importante o crucial; es decir, porque no merecería la pena. Por ejemplo, para la responsable catalana, los estudios de impacto son carísimos y para la responsable vasca, la evaluación que se realizó para medir el impacto de una campaña de sensibilización sobre los malos tratos, no mereció la pena.

Respecto a la utilización de la evaluación, también las responsables menos experimentadas tienen expectativas algo mágicas. Incluso, en algunos casos, llegan a afirmar contundentemente el uso de sus evaluaciones en términos en los que resulta

imposible, dados los estudios realizados. Por ejemplo, el discurso de una responsable en cuyo organismo se ha realizado un estudio sobre la situación de la mujer durante los años de vigencia del plan pero en el que no se ha estudiado ningún aspecto relacionado directamente con el plan, es el siguiente:

“Por eso la evaluación a mí me parecía fundamental, por ejemplo, para insistir en determinadas cuestiones un poco más en este tercer plan, que es lo que más nos está costando, cómo recoger cuestiones de otros planes sin repetir... cómo introducir en este tercer plan alguna cuestión que nos parezca fundamental de los planes anteriores, o por lo menos del anterior que es del que hemos hecho la evaluación”.

Sin embargo, en este ejemplo concreto, dadas las características del estudio en cuestión, sólo se ha podido obtener una orientación general sobre el problema a abordar (algunos indicadores de desigualdad de género), pero en ningún caso información sobre el plan evaluado en concreto.

Aquí también hay que mencionar el componente simbólico y retórico que ha ido adquiriendo el discurso sobre la evaluación en los últimos años. La evaluación es buena y necesaria, incluso dicho de forma contundente, aunque no se haya hecho nunca ni se hayan invertido recursos o esfuerzos en ello. Aunque se trate del discurso de responsables en cuyos organismos no se hubiera realizado ningún tipo de evaluación en el momento de la entrevista o se hubiera evaluado de forma claramente deficitaria, con muy poca utilidad final, la necesidad, conveniencia y potencia de la evaluación, al menos en el nivel simbólico, aparece de forma clara.

“... cualquier programa político tiene precisamente como una fase muy importante la evaluación del programa. O sea, que nosotros ... , porque es obvio que los programas se tienen que evaluar”

“Hombre, a mí me parece que trabajar sin evaluar, y no entiendo cómo se puede trabajar sin evaluar, es decir, el saber en cada momento para qué está sirviendo lo que estamos haciendo, a mí me parece que es fundamental, pero en todos los aspectos de cualquier trabajo, porque si no es dar palos de ciego y además es no rentabilizar el trabajo que se está haciendo”

“Yo creo que es absurdo llevar una acción a cabo si no se evalúa” (por parte de una responsable de un organismo en el que no se había hecho nada de evaluación)

Sin embargo, las responsables experimentadas son más realistas también en lo que se refiere a su utilización. De acuerdo con la teoría más reciente⁸, para la elaboración de planes subsiguientes, la evaluación es sólo una de las fuentes de las que se bebe, entre otras muchas. Por ejemplo, para una de estas responsables experimentadas, la evaluación del II

⁸ Véase capítulo uno, punto 1.3.4.

Plan se utilizó para la elaboración del III, pero también se utilizaron, en igual o mayor medida, un estudio de la situación social de la mujer en esa región, la Plataforma de Acción de la IV Conferencia Mundial de la Mujer en Pekín y el IV Programa de Acción Comunitario de la UE.

2.2. La inevitable necesidad de la evaluación vs. la conveniencia de la misma

Relacionada con la idea anterior, mientras las responsables menos experimentadas en evaluación fundamentan la necesidad de la evaluación porque sí y como parte inherente al propio proceso de intervención (sin explicar las razones del por qué o para qué va a ser necesitada), las responsables más experimentadas, sin embargo, afinan más e insisten no tanto en la ineludible necesidad de la evaluación, sino en su conveniencia para muchos otros fines: como propio instrumento de intervención para ir sensibilizando sobre la perspectiva de género, para “pinchar” y pedir cuentas al resto de las unidades gubernamentales, para difundir y hacer más visibles las políticas de igualdad de género e incluso para que las administraciones públicas “se vayan acostumbrando a evaluar”.

2.3. La evaluación como algo propio de los organismos de igualdad

En todo momento, las responsables políticas asumen que la responsabilidad de la evaluación de los planes reside en los organismos de igualdad que los impulsan y ellas dirigen. Es decir, pueden reivindicar la necesidad de colaboración del resto de las unidades gubernamentales en la recopilación de información y, como se ha comentado antes, la evaluación puede servir para conseguir precisamente una mayor implicación de dichas unidades en las políticas y planes de igualdad. Pero la concepción, el diseño y el análisis valorativo corresponde a los organismos de igualdad. En definitiva, aunque ellas se sienten más impulsoras que ejecutoras de las políticas de igualdad⁹, y todo su discurso gira en torno a la necesaria implicación de otros en dichas políticas, la evaluación es netamente suya. Además de impulsoras, “diseñadoras”, coordinadoras y seguidoras (“vigilantes”), son claramente las “evaluadoras” de dichas políticas.

Sin embargo, este discurso no aparece con la suficiente fuerza. Así como tienen claro que su función principal es el impulso de las políticas de igualdad, no ocurre lo mismo con la función de evaluación que aparece, más bien, como algo residual o tangencial. Es decir, si se habla de evaluación, ellas asumen la responsabilidad, pero esta función no se ve como algo específico y crucial de los organismos de igualdad, salvo en el caso del Instituto Vasco

⁹ Aquí existe una excepción y es la Dirección General de la Mujer de Extremadura que en realidad funciona más como un organismo ejecutor que impulsor de las políticas de igualdad. De hecho, en Extremadura hubo un primer plan aprobado en el año 1991 y nunca se ha llegado a formular un segundo, entre otras cosas porque “no tienen tiempo” de hacerlo, ya que están ocupadas en la ejecución directa de muchas actividades.

de la Mujer. Para ellas la evaluación es una prioridad y, por lo tanto, es lógico que se note en el discurso de su responsable todos los esfuerzos y energías invertidas en ello.

2.4. Una evaluación más integral vs. una evaluación asociada exclusivamente a los resultados

Muchas de las evaluaciones que se han realizado se centran, sobre todo, en el “grado de cumplimiento” del plan que, como se ha visto anteriormente, consisten en una evaluación de los “productos” o *outputs* del plan -con diferente grado de sofisticación según la experiencia de evaluación de la que se trate-. En algunos casos esta evaluación consiste exclusivamente en una descripción o enumeración de lo que se ha hecho en cada una de las áreas del plan evaluado. Sin embargo, en el discurso de las responsables, se deja entrever una asociación de la idea de evaluación a una evaluación de los efectos o impactos del plan (*outcomes*), es decir, lo que se consigue modificar por el plan. En realidad, esta idea tiene que ver, de alguna forma, con las expectativas mágicas sobre la evaluación que se han encontrado en el discurso de algunas de las responsables. Sin embargo, no aparece prácticamente (exceptuando las responsables con más experiencia en evaluación) la idea de una evaluación de diseño y de proceso y que contemple la evaluación de las estrategias de transversalidad, coordinación, etc.

Esta visión de la evaluación como medición de efectos tiene que ver también con la confusión, que se encuentra en muchos discursos, entre la evaluación de los planes y las políticas con el diagnóstico de la situación social de la mujer. En cualquier caso, si se hablara de evaluación de efectos, se trataría de la medición de los efectos netos (lo que ha ocurrido por el plan y no debido a otros factores), y sin embargo de lo que se habla, en muchos casos, es de la necesidad de tener datos sobre la realidad sobre la que hay que intervenir. Es decir, se confunde la evaluación de la acción de los poderes públicos con el estudio o investigación del problema que se quiere abordar. En el siguiente *verbatim* se puede apreciar que de lo que se habla es de un análisis del problema de la igualdad, sobre datos que demuestren que la desigualdad de género sigue existiendo y hay que actuar sobre ella, pero no sobre si es efectivo o no el plan correspondiente.

“Yo creo que esta es la principal fuerza de la evaluación, del plan, el que te dé argumentos para continuar las políticas públicas para la igualdad, porque hay una tendencia a decir que ya esta todo solucionado: “las mujeres trabajan, se han incorporado al mundo laboral, no se que más queréis”. Este discurso esta ahí, y con la evaluación te das cuenta de que sí, de que queremos más porque la igualdad de oportunidades no existe en muchos casos, se ha avanzado mucho pero hay muchos puntos flojos aún, con lo cual te da esta fuerza la evaluación sobre todo frente a los departamentos para que destinen presupuesto y que tengan presente las acciones de igualdad en toda su planificación propia.”

Aquí se deja entrever la necesidad de investigación y datos que existe en torno a las políticas de igualdad. Hacen falta estudios e investigaciones sectoriales, observatorios de

igualdad y perspectiva de género en todas las investigaciones relacionadas con el problema (educación, sanidad, cultura, medioambiente, trabajo, etc...). Y esta necesidad es la que hace que se “aprovechen” situaciones de evaluación “*para matar dos pájaros de un tiro*” como dice una de las responsables de uno de los organismos que hace un estudio de la mujer en lugar de una evaluación de su plan. Esto hace que se den situaciones confusas en las que no se hace bien ni una cosa (evaluación) ni la otra (investigación), y que hace que las responsables políticas tengan una idea equivocada de lo que es y significa la función de evaluación de las políticas públicas.

En algunos casos sí se distingue bien la evaluación de la investigación en el discurso de las responsables, como es el caso de Andalucía, aunque el hecho de que se aproveche el estudio de evaluación para introducir en él un estudio sobre la situación social de la mujer paralelo, hace más fácil que haya una cierta confusión. Así, cuando se le pregunta a su responsable sobre qué información de la evaluación le interesa más, responde, en primer lugar, que la evolución de las mujeres.

2.5. La evaluación como algo técnico

Aunque es evidente que, además del componente técnico que indudablemente tiene, la evaluación tiene un importantísimo componente político, y aunque esto está ampliamente reconocido por la comunidad científica y desarrollado en la teoría y la práctica de evaluación¹⁰, no parece que esté plenamente asumido por la mayoría de las responsables políticas. Más bien la idea que se extrae de forma general en el discurso es que la evaluación es algo exclusivamente técnico. En muchos casos, ellas mismas se declaran inexpertas y apelan a los profesionales o a los técnicos que la han realizado.

“Lo que pasa es que de la evaluación más de lo que dice el libro (publicación sobre la evaluación), yo no te puedo decir”

De hecho, y esto nos parece un indicador claro de esta idea, en uno de los casos se nos denegó inicialmente la entrevista con la directora del organismo de igualdad, alegando que para hablar de evaluación debíamos entrevistarnos con el responsable técnico de la misma. Sólo después de insistir en que lo que interesaba era su posición como responsable política de la evaluación, la directora accedió a ser entrevistada.

2.6. La evaluación como medición cuantitativa y relacionada con indicadores

Claramente y de forma dominante, la evaluación se entiende como medición cuantitativa. La idea de evaluación está asociada a la idea de “indicadores” y, en este concepto, se incluye no sólo la medición cuantitativa sino también la idea de tener

¹⁰ Véase capítulo uno, punto 1.3.2.

clarificado qué se quiere evaluar y los baremos para hacerlo (saber qué es lo que se puede ser valorado como satisfactorio). Pero siempre con números, índices o porcentajes. Es lo que tiene credibilidad y lo que ofrecería seguridad a las responsables tanto interna como externamente, es decir para justificar las políticas de igualdad de género y la acción pública de la que son responsables. Para eso ellas necesitan:

“poner índices numéricos que avalen el cumplimiento” o

“que me cambien las estadísticas”

Aunque en algunos casos se mencionan los estudios de evaluación que han utilizado técnicas cualitativas de recopilación y análisis (entrevistas y grupos de discusión), esto aparece de forma muy tangencial y, de hecho, en el discurso de las responsables, se equipara prácticamente la evaluación a la medición numérica.

2.7. La evaluación o bien como seguimiento o bien como evaluación externa

En el discurso de las responsables políticas muchas veces se utiliza el concepto de evaluación como sinónimo del seguimiento del plan, es decir con el control de su gestión y la simple comprobación de que las cosas se están haciendo tal y como se tenía previsto, del calendario y del gasto. Efectivamente, el seguimiento también ha sido y es una carencia de los organismos de igualdad que, además, tiene que ver con los sistemas de información y con la recopilación de información, condiciones esenciales para una evaluación. De hecho, de la existencia de un buen seguimiento y unos adecuados sistemas de información depende la posibilidad y el éxito de la evaluación. Pero la evaluación va mucho más allá del seguimiento y el control de la gestión y, esto, no parece estar del todo asumido por las responsables políticas de los organismos de igualdad.

Esto también tiene que ver con la equiparación que también se hace, en algunos casos, de evaluación a estudios de evaluación externa. La evaluación no sólo es algo técnico, sino que debe ser algo externo, ya que esta externalidad otorga credibilidad al estudio realizado. Si se analizan estas dos circunstancias, se comprueba, una vez más, cómo la función de evaluación no llega a estar asumida del todo como algo propio, como una función interna y propia de los organismos de igualdad. Es decir, o bien se limita al seguimiento y control de la gestión si se concibe de forma interna y, si no, es algo que debe hacerse externamente, por parte de expertos que son los que conciben, diseñan y valoran los estudios realizados, es decir, en los que el organismo no tiene mucho que decir. De hecho, son muy pocas las experiencias analizadas en las que se asume la evaluación como algo propio y la función de evaluación está realmente asumida y desarrollada de manera interna -aunque se requiera de asesoramiento externo en un momento dado-.

3. ¿PARA QUÉ SIRVE EVALUAR? LA UTILIDAD DE LA EVALUACIÓN

También resulta interesante analizar el discurso de las responsables políticas en lo que se refiere a las funciones que para ellas cumple la evaluación y para qué les sirve y les resulta útil. Para exponer el análisis, se tienen en cuenta las tres funciones básicas de la evaluación: mejora, rendimiento de cuentas e iluminación para acciones futuras.

Respecto a la función de **mejora** de la evaluación, el discurso de las responsables tiene un alto carácter simbólico: se menciona esta función de forma un tanto vacía y abstracta y realmente no es el porqué al que más se apela. Las responsables cuyos organismos han trabajado más la evaluación hablan de “corrección de tendencias” y de autoaprendizaje. Sí se menciona, en algunos casos, la idea de que la evaluación les ha servido para hacer planes más realistas, menos generales y abstractos, es decir, la evaluación sirve para mejorar la estructura de los planes y hacer una mejor planificación de los mismos. Incluso esta tendencia puede llegar a ser peligrosa, ya que podría llegar a implicar que, para que fueran evaluables, se tendería a hacer planes “demasiado” sencillos, como se podría dejar entrever en el siguiente extracto de discurso:

“...pusimos unas actuaciones y unos objetivos sencillos y, bueno, que a la hora de evaluar nos diera un alto grado de cumplimiento”

La **responsabilidad** o **rendimiento de cuentas** es una función que está algo más presente en el discurso de la evaluación, aunque en algunas ocasiones de forma un tanto sesgada y que no es sino el reflejo de la propia cultura política de los españoles y españolas. En algunos casos no se menciona el rendimiento (de ellas como responsables de las políticas públicas de igualdad de género), sino la petición de cuentas, es decir se señala cómo la evaluación debe servir para pedir cuentas a las asociaciones o colectivos que obtienen una subvención, pero en ningún momento esto se relaciona con un rendimiento de cuentas hacia la ciudadanía. En otras ocasiones, cuando se menciona el rendimiento de cuentas, se hace hacia las instituciones -por ejemplo, el Parlamento- y esto no incluye tampoco la ciudadanía.

“Pero no es la intención, justificarse no es la intención. Tu te puedes justificar ante el parlamento si tu quieres, pero no ante la ciudadanía, ¿me entiendes?”

Incluso, en otras ocasiones se habla del rendimiento de cuentas al partido político al que se pertenece, sin mencionar en ningún momento otro tipo de obligación. Resulta curioso cómo la fuerte disciplina de partido existente en nuestro sistema político llega hasta tal punto que hace que una responsable política sienta ese rendimiento de cuentas más importante que el que debe hacer al propio poder legislativo o a la misma ciudadanía a la que se supone que está sirviendo.

En otros casos sí se habla de dar cuenta de los resultados de la política aunque, insistimos, la ciudadanía no aparece en ningún momento como objeto de dicho rendimiento de cuentas ni siquiera como constructo relacionado con la evaluación de las políticas. Sí se menciona el rendimiento de cuentas a los Consejos de los organismos de igualdad y al Parlamento (en los casos en los que es así -Andalucía, Cataluña y País Vasco-). Y en la otra dirección -la petición de cuentas- se menciona a las otras unidades o departamentos gubernamentales que se supone deben tener una responsabilidad en las políticas de igualdad horizontales. Esto sí aparece en el discurso de las responsables, especialmente en las más experimentadas. La evaluación sirve, además, para implicar horizontalmente y para pedir cuentas a otros departamentos.

“te sirve para empezar a meter dedos en ojos de otros departamentos”

Esta petición de cuentas a los otros departamentos, además, es una forma de intervenir, convirtiéndose la evaluación en un propio instrumento para dicha intervención. La evaluación sirve para sensibilizar...

“.. prepararles una ficha para además darles ideas”

Una reflexión que realizamos a raíz de analizar la idea que las responsables políticas de los organismos de igualdad tienen sobre el rendimiento de cuentas, es el fuerte componente gubernamental de las políticas públicas de igualdad en España y la poca centralidad y participación real que el movimiento social y de mujeres tiene en las mismas. Para las responsables políticas el rendimiento de cuentas y responsabilidades se hace a las instancias “superiores”, bien sea al poder legislativo, al gobierno o, incluso, al partido político. Sin embargo, hacia los iguales -otros departamentos- o al movimiento social y de mujeres se les pide, y no se les rinde, cuentas.

La tercera función de la evaluación -**la iluminación para acciones futuras**- es la que menos aparece en el discurso de las responsables. Las únicas que parecen estar preocupadas por reflexionar acerca de lo que significan los planes, por categorizar y saber distinguir lo que se entiende o no por acciones positivas, son las vascas y, en cierto modo, las catalanas. Pero, en general, no parece que la evaluación pretenda utilizarse para hacer una reflexión sobre cuáles deberían ser las políticas de igualdad de género en un futuro, o cuál podría ser su posible evolución.

Además de las tres funciones básicas de la evaluación, en el discurso de las responsables se deja entrever una preocupación y un papel claro de la evaluación en la proporción de **datos e información**. Ya se ha comentado la falta que parece existir, en torno a las políticas de igualdad de género, de datos e información específica que permita una lectura profunda del problema público que se pretende abordar -la desigualdad de género- y que justifique la existencia y necesidad de las acciones públicas en pro de la igualdad. Y esto se deja entrever en el discurso de las responsables. En general, están ávidas de datos, de investigación, y de un mejor conocimiento del problema y de la realidad sobre la que deben actuar. Es posible que ellas estén convencidas de la necesidad de actuar en determinados ámbitos, pero necesitan mucha información fiable y creíble para convencer

a los demás ya que el éxito de su política depende de esta capacidad de convencimiento. Esta necesidad de información y datos sobre el problema a abordar parece tan acuciante que hace que se le “pida” a la evaluación funciones que no son estrictamente suyas: el hecho de que se evalúen los planes y las políticas no obvia la necesidad de investigación sobre el problema de la desigualdad de género.

Unido a esto, también se menciona la función de **documentar los programas** y planes que, de por sí, son novedosos; y la idea de **difundir publicando**, aunque esto se encuentra únicamente en uno de los casos, ya que la utilización de la evaluación parece que se concibe más de forma interna. Por último, también encontramos la idea de que la evaluación sirve o ha servido, en algún caso, para **hacer más visible** el organismo de igualdad y el propio plan evaluado, e incluso para **“reafirmarse”**.

Sin embargo, ocurren dos fenómenos en el discurso de las responsables que tienen que ver con la utilidad de la evaluación y que consisten, por un lado, en lo que podríamos llamar **“ilusión de utilidad”** y, por otro, en la **poca capacidad de sorpresa** que las responsables otorgan a la evaluación.

Respecto a la **“ilusión de utilidad”**, las responsables dicen que utilizan la evaluación más de lo que en realidad parecen hacerlo. En algunos casos, se habla de lo útil que ha resultado la evaluación, sabiendo que el informe de evaluación estuvo listo casi a la vez que se aprobara un plan, por lo que es difícil pensar en una utilidad real. En cualquier caso, le dan más utilidad y potencia a las evaluaciones de la que, objetivamente, pueden tener. En el fondo, parece más una cuestión de fe que algo realmente basado en la realidad. Como muestra, se expone aquí un extracto del discurso de una de las responsables, ante la pregunta de la entrevistadora sobre cómo se relacionaban unos datos de la evaluación (el incremento de mujeres universitarias) con las medidas de su plan en concreto (se trata de una experiencia en la que se hace un estudio de la situación de la mujer y no una evaluación del plan).

“En algunos casos sí y te soy sincera, en algunos casos se ve clarísimo que hay una relación, en otros no, pero digo yo que algo habrá contribuido el plan,(...) algo tendrá que ver, tampoco, te soy sincera, no quiero decir con esto que sí puede inferir de una forma causa y efecto inmediato, pero sí desde luego que algo habrá tenido que... yo quiero creer que algo ha tenido que ver todas esas acciones. Pero también te soy sincera, porque como el plan no estaba temporalizado y no estaba concreto, los indicadores son un poco también genéricos, y por lo tanto no te puedo decir a ciencia cierta que esto es por esto, pero algo habrá contribuido”. (subrayado mío)

Es decir, a la evaluación se le asignan bondades que realmente no ha tenido, como por ejemplo, en este caso, constatar o demostrar no ya una relación de causa-efecto, sino una cierta relación entre variables (entre la existencia del plan y la evolución de la situación de las mujeres).

Por último, sorprende la **poca “capacidad de sorpresa”** que han tenido las evaluaciones para las responsables políticas. Se supone que una evaluación, para resultar

realmente útil, debe aportar algo nuevo, alguna información que antes no se tuviera y que “enseñe” algo sobre el plan o política que se esté evaluando. Pues esa capacidad de sorpresa no se deja entrever en el discurso y, solo ante una pregunta directa sobre si había algo que les hubiera sorprendido de los resultados de la evaluación, las responsables contestan, en muchos casos, algunos datos tangenciales. Incluso en algún caso dicen no acordarse o no tenerlo en la cabeza presente en ese momento.

4. ¿QUÉ SE DEBE EVALUAR Y CÓMO?

Como se verá más adelante, en torno a las experiencias de evaluación analizadas, existe poca claridad respecto a QUÉ se quiere evaluar. Esto tiene también su claro reflejo en el discurso de las responsables políticas. En general, las responsables apelan a la falta y a la necesidad de “**indicadores**” preestablecidos, pero no señalan cuáles. En realidad lo que están reclamando es que se les diga qué hay que evaluar de las políticas de igualdad. Incluso hablan de indicadores comunes para todos los planes, dejando entrever un concepto de evaluación única, técnica, que sirva para todos los espacios, momentos, políticas y territorios. En realidad y de nuevo, algo muy ajeno y externo a ellas y a sus responsabilidades, algo que mágicamente “cae” del cielo y que no tiene demasiado que ver con una reflexión interna. En definitiva, reclaman que alguien, con autoridad técnica, les diga qué y cómo hay que evaluar. Probablemente también la apelación a datos comunes tenga más que ver con la necesidad de datos e investigación que antes se ha comentado, que con la propia necesidad de evaluación.

“...evaluar tiene problemas, y de hecho en el Instituto, incluso en más de una reunión, hemos estado hablando que a ver si hacemos unos indicadores que sean iguales para todas las Comunidades Autónomas porque si no hay unos desfases y unos desbarajustes que son mortales ¿no?”

Aquí, en realidad, más que a la evaluación de los planes se refieren a realizar sistemas de información u observatorios de igualdad que se coordinen entre sí y que recojan datos a través de los mismos indicadores.

Este discurso contrasta con el de las responsables más experimentadas que sí parecen haber hecho una reflexión más seria sobre qué quieren evaluar (Instituto Catalán e Instituto Vasco de la Mujer), estableciendo categorías y clasificaciones sobre el tipo de acciones que se realizan en los planes y preguntándose sobre qué es y qué no es una acción positiva.

Al no existir, en la mayoría de los casos, una reflexión previa sobre qué se debe evaluar, se suple esta carencia apelando a criterios “metodológicos” o técnicos, es decir al cómo.

“Introducimos no sólo lo que era la evaluación de las medidas, cómo se habían hecho, cómo no se habían hecho, como se habían implicado las distintas consejerías, sino también se incorporan estudios un poco de encuestas y estudios de campo ¿no?”
(Subrayado mío).

Sin embargo, y aunque sea de forma poco clara, sí se deja entrever qué es lo que preocupa a las responsables como objeto de la evaluación, es decir, sobre qué se debe evaluar. Aparece el concepto de “grado de cumplimiento”, que en realidad tiene que ver sobre todo, con la evaluación de *outputs* o productos y, en cierto modo, con el propio

control de la gestión, especialmente en lo que se refiere al control de si las demás unidades gubernamentales cumplen con su parte o no. Y de ahí, prácticamente se pega un salto a cuestiones de impacto, confundiendo, como ya se ha señalado anteriormente, la medición de los “efectos netos” del plan (lo que se ha modificado por efecto del plan y no de otros factores) con la “monitorización” del problema sobre el que hay que actuar (desigualdad o discriminación basada en el género) haciendo un seguimiento de su evolución.

Cuestiones como el diseño de los planes, si éstos son o no el instrumento adecuado y si se consigue la perspectiva de género en el resto de las políticas públicas, no aparecen realmente como objeto de la evaluación en el discurso de las responsables. Ni siquiera la implicación del resto de las unidades gubernamentales, y su propia capacidad de persuasión y convencimiento, que es algo que les preocupa enormemente, se concibe como algo a evaluar, algo que pueda ser convertido en criterio de evaluación. Esta carencia la expone magistralmente la responsable vasca cuando dice que eso es algo que hay que contemplar en la evaluación aunque no se refleje en las acciones concretas de los planes:

“...para evaluar nuestra propia parte de impulso, no hay acciones para ello, eso debería estar reflejado en la evaluación”

Efectivamente, si se concibe la evaluación de los planes exclusivamente como la evaluación de las acciones que en ellos se propone, se pierde una perspectiva más global de las políticas de igualdad y que también debería estar contemplada en su evaluación. Sin embargo, y salvo en este ejemplo que hemos puesto, esto no aparece de ninguna manera en el discurso del resto.

5. DIFICULTADES PERCIBIDAS PARA LA EVALUACIÓN DE LOS PLANES DE IGUALDAD

Las responsables señalan toda una serie de dificultades a la hora de evaluar que aquí se exponen organizadas en torno a tres factores. Dificultades relacionadas 1) con la falta de cultura de la evaluación, 2) con la propia complejidad de los planes de igualdad y, 3) con la transversalidad o concepción de horizontalidad de las políticas de igualdad de género. Curiosamente, casi todas **las dificultades mencionadas no tienen que ver con los procesos de evaluación en sí, sino con factores previos o en torno a la misma** (o bien falta de cultura o costumbre de hacer evaluación, o bien dificultades que provienen por los propios planes y políticas y no por su evaluación). Sólo en el caso del País Vasco, probablemente debido a su larga experiencia en evaluación y a la antigüedad de la responsable de *Emakunde* en el cargo, el discurso es más elaborado respecto a la evaluación y se mencionan dificultades inherentes al propio proceso, como son la dificultad de encontrar la “unidad de medida” (o de evaluación), la dificultad de establecer tipologías válidas de acciones o, incluso, el peligro de que las administraciones públicas se engañen a sí mismas, no teniendo suficiente autocrítica. En el resto, las dificultades mencionadas son más genéricas y, como se ha mencionado anteriormente, no relacionadas directamente con los propios procesos de evaluación, sino con el contexto general de los planes a evaluar. Se describen a continuación, y de forma pormenorizada, las dificultades señaladas por las responsables políticas.

5.1. Dificultades relacionadas con la falta de cultura de evaluación

Las dificultades mencionadas por las responsables relacionadas con la inexistencia de una cultura de evaluación son las siguientes:

- Falta de referencias y de experiencias previas de evaluación. La evaluación de los planes es algo nuevo y que tienen que “inventarse” porque prácticamente no existían referencias previas conocidas.

- Falta de equipos expertos. En España existen todavía, hoy por hoy, pocos profesionales especializados en la evaluación de programas. Como decía una de las responsables, si además de evaluación deben saber sobre políticas de igualdad de género, es decir, comprender y ser sensibles a la perspectiva de género y la estrategia de transversalidad o *mainstreaming*, es mucho más difícil contar con personas expertas que puedan suplir la falta de referencias mencionada anteriormente. En palabras de otra de las responsables (refiriéndose al equipo externo que realiza la evaluación), “*aprenden con nosotras*”.

- Falta de una cultura de recopilación y sistematización de datos y falta de unos buenos sistemas de información en las administraciones públicas. Esto se hace aún más difícil cuando se depende también de otros para recopilar la información, como se verá más adelante.

- Falta de considerar la evaluación desde un principio. Dado que la evaluación no es una función normalizada en el quehacer cotidiano de las organizaciones públicas, es difícil considerarla desde un principio y, el no hacerlo así, conlleva dificultades de diseño, recopilación de información, etc. Curiosamente, esta dificultad relacionada con el tiempo no se encuentra en el discurso de todas las responsables, sino de las que ya han tenido la experiencia de la evaluación. Precisamente, encontramos una visión distinta sobre el tiempo entre las que no han tenido tanta experiencia y que es, precisamente, un indicador de este factor de dificultad.

“...pero no planteamos todavía nada de evaluación, porque además nos queda hasta el año 2000” (en una entrevista realizada a finales del año 1997)

“...se ha hablado pero no hay nada atado” (entrevista realizada a la mitad del período de vigencia del plan del que se trataba de evaluar).

5.2. Dificultades relacionadas con la complejidad de los planes

Se señalan una serie de dificultades que tienen que ver con la propia complejidad de los planes de igualdad de género. Esta complejidad viene dada por el propio contenido de los mismos, por su integralidad y las diferentes áreas o aspectos que contienen. Una de las consecuencias de esta complejidad e integralidad es la generalidad o abstracción con la que son formulados: se tocan muchos aspectos de forma general. Las dificultades mencionadas por las responsables son las siguientes.

- La generalidad y abstracción de los planes. Esta característica los hace muy poco evaluables, ya que al contar con objetivos abstractos, su “traducción” en algo operativo que pueda ser medible o aprehensible es muy difícil. Esta es la dificultad que más se menciona por parte de las responsables. De hecho, reconocen que los intentos de evaluación son los que han hecho que los planes subsiguientes sean más “realistas”.

- La dificultad de medir los impactos. Los efectos que se pretenden conseguir con los planes de igualdad de género son profundos y a largo plazo; de hecho, se pretende una modificación profunda de las estructuras y las relaciones sociales y esto, de por sí, requiere de mucho tiempo. Por lo tanto son efectos muy difícil de medir y no existe una teoría clara sobre cuáles serían los pasos intermedios o los indicios que pudieran indicar que esos cambios se van produciendo. Las responsables más experimentadas en evaluación son conscientes de la dificultad de poder comprobar de forma fiable dichos efectos.

- La dificultad de medir los costes económicos de las acciones planteadas en los planes de igualdad. En los planes hay acciones que son más fácil de ponderar

económicamente y otras que menos. Esta dificultad sólo se menciona por una de las responsables más experimentadas. Pero de hecho, en esta generación de segundos planes, sólo se calcula el presupuesto empleado en el plan en muy pocas de las experiencias analizadas. En la mayoría, ni siquiera se tiene una idea real de cuánto dinero se ha empleado para la ejecución del plan evaluado.

5.3. Dificultades relacionadas con la estrategia de la transversalidad de las políticas de igualdad de género

El hecho de que las políticas de igualdad de género se conciban como políticas transversales -algo que debe implicar también al resto de las políticas públicas- conlleva una serie de dificultades que se traducen también, y como es lógico, a la fase de su evaluación. De esta manera, muchas de las dificultades de la evaluación se encuentran en la propia naturaleza de las políticas y no en la evaluación en sí misma. Esta realidad se reconoce por parte de las responsables políticas, de tal forma que el discurso sobre las dificultades de la evaluación se traslada muy fácilmente hacia las políticas que deben ser evaluadas. Además de las dificultades específicas de la evaluación de las políticas de igualdad de género, esto también quiere decir que el discurso político sobre la evaluación no está muy elaborado por lo que, en general, las responsables políticas “huyen” hacia terrenos en los que se encuentran más cómodas.

Hay dos temas que aparecen recurrentemente en el discurso de las responsables y que tienen que ver con la estrategia de la transversalidad. Uno de ellos tiene que ver más específicamente con la evaluación y es el de las **dificultades en la recopilación de información**: para recopilar la información necesaria para el seguimiento y la evaluación de los planes, por mucho que se gestione dicha información en los propios organismos de igualdad, se necesita recurrir al resto de las unidades gubernamentales para obtenerla. Esta es una de las mayores dificultades con la que se han encontrado y también tiene que ver con el segundo tema recurrente que es la **falta de implicación** del resto de las unidades gubernamentales en las políticas transversales de igualdad de género. Esta falta de implicación se traduce en desinterés, incompreensión y tergiversación de lo que son o deben ser las políticas de igualdad de género. Las dificultades en la recopilación de información tienen su origen en dicha falta de implicación.

Para entender cómo las responsables políticas explican el origen de estas dificultades y para organizar su discurso, vamos a exponer las dificultades señaladas en torno a dos polos: 1) La propia dificultad del concepto de “perspectiva de género” (siendo dicha perspectiva de género la meta u objetivo y la estrategia para conseguirlo la transversalidad o *mainstreaming*) y 2) Las estrategias para la trasmisión de dicha perspectiva de género.

1) La **perspectiva de género** es un concepto de difícil aprehensión, no sólo por su complejidad, sino también porque implica una revisión profunda de las estructuras sociales,

construidas sobre unas determinadas relaciones entre los géneros. Su asunción implica modificar dichas estructuras sociales y realizar un nuevo contrato social entre los géneros. Aplicado a las políticas públicas, la perspectiva de género significa analizar la realidad, formular y diseñar las políticas, ejecutarlas y evaluarlas teniendo en cuenta a las mujeres, analizando diferencialmente los posibles impactos e implicaciones para uno y otro género. En palabras de una de las responsables, significa ejercitar “otra mirada” a la realidad, visibilizando a las mujeres y trabajando por una distribución real de derechos y deberes más equitativa entre los géneros.

La transmisión de esta idea es difícil y, esencialmente, es lo que tienen que hacer los organismos de igualdad y sus respectivas responsables, respecto a sus compañeros de gobierno. Ya que no se trata de hacer políticas verticales dirigidas a la mujer, sino que las políticas públicas en general se hagan con perspectiva de género. Las responsables se quejan de que este concepto no está realmente asumido:

“no hay un convencimiento profundo de lo que son las políticas de igualdad”

Sin embargo, lo que se encuentra en el análisis del discurso de las responsables es que, en algunos casos (que coinciden claramente con las responsables que menos tiempo llevan en su cargo y no vienen de la acción política feminista), el concepto de “perspectiva de género” no está ni siquiera realmente aprehendido por las propias responsables; por lo tanto, aplican una visión ingenua y simplista del mismo. Así, la “perspectiva de género” consiste, para algunas de las responsables, en que las diferentes unidades administrativas hagan cosas para mujeres, de forma que dichas actividades puedan consignarse como parte del plan de igualdad respectivo. Estos son algunos extractos de los discursos que corroboran esta visión:

“Ni las mismas personas, muchas veces, de otras consejerías saben, se dan cuenta que algunas de las acciones que están haciendo, están precisamente incidiendo en aspectos de la mujer, y a veces me lo dicen, “no, no aquí se hace igual para los hombres que para las mujeres”, bueno se hace igual para los hombres que para las mujeres, pero ¿se está haciendo esto? sí, pues si se está haciendo esto, ¿y qué participación hay?, pues igual de los hombres que de las mujeres, o más de las mujeres, bueno pues ya tenemos que tener en cuenta que entonces es algo que esta favoreciendo, por lo menos ya se está trabajando en ese ámbito en el campo de la mujer; entonces a lo mejor lo que es importante que ya no se tenga en cuenta desde la labor del organismo, puesto que ya hay otro que lo está haciendo”.

“No es que no lo hagan, lo hacen, pero a lo mejor lo hacen inconscientemente, no se dan cuenta de lo que están haciendo”.

“Y dicen: no, es que yo no hago nada para la mujer. Y dices: ¿cómo que no? ¿no acabas de hacer un curso de no se qué?, ¡ah bueno!, ¡claro!, sí, en ese aspecto sí; digo: pues ya está. O sea, cuesta mucho meterlos en... el cerebro y en la cabeza políticas de igualdad ¿no?”.

Esta visión poco elaborada de la perspectiva de género va, en algún caso, paralela incluso a la no asunción real de la base de partida de cualquier política de igualdad y que es la existencia de discriminación respecto a las mujeres, al menos no en todos los niveles o sectores. Entre las responsables de organismos de igualdad pertenecientes a partidos conservadores, que llevan poco tiempo en el cargo y que no proceden de la acción política feminista, existe una tendencia a la externalización de la discriminación. Es decir, ellas son unas privilegiadas que no han sido discriminadas nunca, aunque reconocen que hay otros colectivos de mujeres que sí lo están.

“...los funcionarios somos los únicos trabajadores que no sufrimos discriminación respecto a la mujer”

Las responsables consideran que hay áreas de gobierno más “sensibles” que otras a las cuestiones de género. Las áreas más sensibles, suelen ser bienestar social, educación, trabajo, sanidad, etc. y las que menos economía, industria, agricultura, etc. Además por el propio contenido del trabajo, por lo que las áreas más sensibles estarían más “acostumbradas” a trabajar en relación con las mujeres, para las responsables en realidad muchas veces esto depende más de la sensibilidad personal de las personas al cargo. Preguntadas sobre en qué se basa o que es lo que da esa mayor sensibilidad (casi todas coinciden que, aunque hay mayor probabilidad de encontrar a políticas mujeres sensibilizadas, esto no siempre coincide habiendo políticos hombres sensibilizados y mujeres políticas menos sensibles), casi todas señalan factores de carácter personal, haciendo referencia a la formación y a las conductas interiorizadas de cada uno/a.

2) Si el concepto de “perspectiva de género” es difícil, más aún lo es la **estrategia** para llegar a la misma, sobre todo lo relacionado con la **transversalidad** o **mainstreaming**. En algún caso, no sólo se apela a la dificultad intrínseca sino que también se es consciente de las propias limitaciones, justificándolas a través de la globalización: nadie sabe cómo hacerlo.

“Es cómo conseguir llevar a cabo esa estrategia, el mainstreaming, para mí eso sería muy importante y realmente eso es una cosa que nadie ha dado todavía en el clavo. Porque yo he ido a diferentes seminarios, incluso en la UE en el que se ha debatido esto, y muy bien, el discurso es muy bonito teórico, y cuando llegas a la pregunta de: ¿Cómo hacer esto?, no hay una respuesta”.

La estrategia de la transversalidad o *mainstreaming* tiene que ver, sobre todo con una tarea de **convencimiento** o **persuasión** al resto de los compañeros de gobierno para que hagan suyos los objetivos de las políticas de igualdad, ya que no se trata de que las acciones las lleven ellas sino las unidades administrativas correspondientes. Así,

“...el diseño de las acciones no es responsabilidad tuya”

Para que el resto de las unidades administrativas asuman los objetivos y acciones que les corresponden, primero hay que convencer a los respectivos responsables políticos de la necesidad y conveniencia de dichas políticas. Las responsables asumen que esa es una de las tareas más importantes que tienen encomendadas, aunque se quejan enormemente de lo difícil y arduo que a veces resulta. E incluso a veces, dejando entrever que esa no debe ser una de sus principales funciones.

“Además ¿sabes lo que ocurre?, que la persona, generalmente, que está al frente de un organismo de igualdad, llega a un momento a que solo se le asocia con mujer, es decir, fulanita de tal, solo sabe hablar de mujeres. Oiga no, fulanita de tal, puede hablar de mujeres, pero además puede hablar de derecho que para eso tiene la formación detrás, y ya te encasillan o esta es una histérica que solo sabe hablar de lo mismo. Hay que ser bastante más hábil para que no te metan los goles por ahí. Y suele ser un tema que quema. Quema mucho.”

Como se puede observar, estas dificultades mencionadas tienen más que ver con las políticas de igualdad de género en general y no tanto con la función de la evaluación en particular. Curiosamente, este convencimiento y persuasión que deben realizar y que parece crucial y para el que encuentran muchas dificultades, no es objeto de evaluación.

Sin embargo, resulta muy interesante el análisis de las estrategias que utilizan para realizar esta tarea de convencimiento y persuasión, ya que esto podría dar lugar a reflexionar sobre qué se debería evaluar de estas estrategias. A la hora de preguntarlas por las estrategias que utilizan para convencer y persuadir a los compañeros de gobierno para que se impliquen en la puesta en práctica de los planes, vemos que son verdaderas maestras, especialmente las que llevan más tiempo en el cargo. Analizamos a continuación las estrategias mencionadas por las responsables entrevistadas.

En primer lugar, cabe mencionar lo que llamaríamos las estrategias **formales** para conseguir esa otra estrategia general de la transversalidad. Entre ellas se mencionan, por parte de las responsables:

- Tener una buena posición en la estructura gubernamental, esencialmente, estar en el área de Presidencia, ya que esto ofrece un mejor punto de partida para dicho convencimiento.
- Establecer compromisos políticos y estructuras de seguimiento de los planes que reflejen dicho compromiso político. Esto se expresa desde *“haciendo que pongan dinero”* como una forma de compromiso, hasta en la necesidad de que las estructuras en torno a los planes se formen con gente con responsabilidad política y no sólo técnica. En general, todas coinciden en recalcar la importancia de que el nivel político tenga presencia en las estructuras de seguimiento de los planes. Pero algunas también expresan la importancia de los técnicos, y de buscar apoyos en los diferentes

departamentos a través de técnicas (aquí sí hablan en femenino) sensibilizadas. Además de ser el personal imprescindible para el desarrollo de los programas, suele ser el personal con mayor estabilidad.

“Hay que preparar a la gente técnica que es la que va quedando”

- En esta línea, algunas comunidades autónomas han hecho especial hincapié en realizar campañas de sensibilización y formación específica para el personal interno de las administraciones públicas.

- Curiosamente, sólo una de las responsables habla de la importancia de tener el movimiento social detrás. Y decimos curiosamente porque, en realidad, contar con el movimiento social y de mujeres como aliado, podría ser una de las mejores estrategias para ser escuchadas, o al menos podría ser un elemento de presión para tener en cuenta. Este podría ser otro indicador del carácter fuerte y casi exclusivamente gubernamental de las políticas públicas de igualdad de género en España.

En segundo lugar, se puede hacer también un resumen de las estrategias más **informales** que las responsables de los organismos de igualdad aducen para conseguir dicho convencimiento y persuasión, especialmente dirigido al resto de sus compañeros de gobierno. Estas son las siguientes:

1. Apelar a estudios, datos, estadísticas.

Esta es una de las estrategias que más se mencionan. Y tienen una necesidad acuciante de ellos. Es la idea de poder demostrar de forma contundente que existe la discriminación de género y hay que actuar sobre ella. En esto, las responsables dejan entrever que, en la mayoría de los casos, su autoridad sobre un tema se pone en entredicho, sino es con el aval de unos datos “objetivos” que demuestren lo que se dice. Se hace caso no a lo que ellas dicen, sino a lo que dicen los datos. El análisis de esta estrategia nos lleva a reflexionar sobre cómo existen determinadas políticas públicas, como la que nos ocupa, que están de forma continua y reiterada en entredicho, a diferencia de otras que no están tan cuestionadas. La siguiente estrategia también está relacionada con esta idea.

2. Repetición y reiteración.

Es la estrategia de no cejar en el intento, no desesperarse y repetir y reiterar las cosas, con paciencia. Se trata de llevar a cabo un trabajo diario y cotidiano con las personas a las que se ha de convencer. En sus propias palabras:

“Yo creo que la constancia y pertinaz presencia, no decaer nunca, (...) cuando voy a ver a un consejero o a un viceconsejero, que me lo pongo como obligación, porque es mi obligación”

“Presencia constante del discurso de convencimiento”

“Paciencia revolucionaria (...), hacer de misioneras”

“... tienes que volver a empezar el discurso desde un principio, siempre estás empezando,(...) y tienes que empezar otra vez a convencerlos de la necesidad de estas políticas, o sea que es lo eternamente puesto en cuestión”.

“¿Cómo se lucha?. Pues, siendo muy cabezota. Siendo muy tenaz, y repitiendo y machacando las cosas siete veces y si no me lo haces pues a la próxima ocasión volvértelo a plantear. Siendo muy cabezota. No dándolo por perdido, porque si te das por vencida ya... Intentando volver al mismo tema, dándole la vuelta o por otro lado y llorando a veces, sin que te vean ¿eh?. Llegando y dando puñetazos y diciendo: "este tío es que no se ha enterado de nada"”

3. Apelar a la justicia y a la razón

Para las responsables entrevistadas, se trata de explicar y fundamentar las cosas. En palabras de una de las responsables, *“trabajar con la razón”*. Es decir, se trata de apelar a la justicia, pero utilizando los razonamientos y la fundamentación, no la imposición directa.

4. “Mano izquierda”, no enfrentarse frontalmente

Esta es también una de las estrategias más mencionadas. Se trataría de tener posturas no enfrentadas con los hombres. En definitiva una estrategia de *low profile*, evitando la confrontación directa, curiosamente acorde con una forma más “femenina” de ser y estar en el mundo.

“Y como digo yo, hay que ir trabajando, no hacer mucho ruido y cuando se dan cuenta ya han caído, no, no, no es que hayan caído, caído literal, pero hay que ser bastante sutil en este tema ¿eh?”

“Y ahí está la clave de todo, el implicar a la gente. Lo que pasa que eso es lo que te decía yo, hay que ir con mucha mano izquierda. Hay que ir haciéndolo sin que a veces se note mucho lo que estás haciendo”

“Darle la vuelta a las cosas, tarea de convencimiento, nunca frontalmente. De modo indirecto”

“Tu objetivo es convencer. No te puedes poner a la gente en contra. La palabra feminista sigue teniendo mala prensa. Hay que hacer equilibrios”.

Estas cuatro primeras estrategias son las más mencionadas y, con diferentes matices se señalan por prácticamente todas las responsables. Las siguientes están menos generalizadas, pero también se menciona por parte de alguna responsable como estrategias para el convencimiento.

5. Dar protagonismo a los otros compañeros de gobierno

Curiosamente, esta estrategia de la transversalidad puede implicar estar dispuestas a perder cierto protagonismo político. Se trata de que hagan cosas los demás y, por lo tanto, también “se lleven los laureles” ellos. Tendría también que ver con la idea de que para implicar realmente, primero hay que “dar cancha”.

“...es importante políticamente ser generoso, sabes, en el sentido de no querer acaparar protagonismo de ningún tipo y saber que las políticas de género no son cosa de uno, sino cosa de varios”.

“...invitar a los responsables a comprometerse”.

6. Recurrir o apelar a la experiencia personal de los responsables a los que hay que convencer

Significa estar atentas al factor personal de dichas personas, apelando a su experiencia personal, por ejemplo, si tienen hijos e hijas.

7. Tener las ideas claras

La responsable que menciona esta estrategia -y que es una de las más “antiguas”- es consciente de que es un tema complejo, con claras imbricaciones ideológicas y con consecuencias profundas sobre la propia concepción de la estructura y las relaciones sociales. Por lo tanto, tener claro qué es lo que se quiere y qué significa exactamente la perspectiva de género es esencial para no perder la orientación hacia las metas a conseguir.

“Paciencia, constancia e ideas claras”

8. Capacidad de respuesta: estar muy atentas y aprovechar los resquicios

Esto significa capacidad de respuesta, es decir, estar siempre atentas a lo que ocurre en la sociedad y a lo que está dispuesta la administración, además de estar siempre con toda la disponibilidad para actuar en los momentos más adecuados.

9. Utilizar el humor

Esta estrategia también se menciona por parte de una de las responsables más “experimentadas”. Incluso esta estrategia ha sido formalizada realizando campañas de sensibilización en clave de humor. Pero es una estrategia que no debe utilizarse en todos los contextos.

“El humor es una tercera vía para introducir mensajes, pero no sirve para el Parlamento”

CAPÍTULO 9

LOS FACTORES QUE FACILITAN Y DIFICULTAN LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS DE IGUALDAD DE GÉNERO

HACIA UNA EVALUACIÓN INNOVADORA DE DICHAS POLÍTICAS

INTRODUCCIÓN

En este capítulo se valoran las experiencias de evaluación analizadas, tratando de concluir cuáles son los factores que han facilitado y dificultado la puesta en práctica y la eficacia de dichos procesos de evaluación. El objetivo de este capítulo es, pues, extraer tendencias de estas primeras experiencias de evaluación de los planes de igualdad para apuntar cuáles son las líneas hacia las que debería tender la evaluación de las políticas públicas de igualdad de género.

Una vez realizado el análisis sobre las once experiencias de evaluación, se pueden distinguir diferentes tipos de factores influyentes en las mismas. Por un lado, se distinguen factores que facilitan o dificultan la evaluación, pero que son previos a la misma, es decir, no son inherentes a los propios procesos de evaluación, sino que son condiciones que intervienen e influyen en dichos procesos. Por otro lado, estarían los factores que forman parte de la evaluación y que son elementos de la misma, y que influyen de una forma determinante en su desarrollo. Finalmente, nos encontramos con aquellos factores que determinan el impacto de dichos procesos de evaluación.

Por lo tanto, la valoración se hace respecto a los tres tipos de factores mencionados. En primer lugar, se valora la influencia, tanto negativa como positiva, que tienen algunos **factores previos** a los propios procesos de evaluación y que tienen que ver, en algunos casos, con características específicas de las políticas de igualdad de género (por ejemplo, la existencia de los planes como instrumento de política o la idea de políticas transversales y horizontales) y, en otros, con características institucionales (por ejemplo, falta de cultura y experiencia previa en evaluación de programas o la fortaleza institucional relativa de los organismos de igualdad).

En segundo lugar, se señalan los principales problemas identificados en los procesos de evaluación analizados y que tienen que ver con la propia concepción, diseño y ejecución de las respectivas evaluaciones **-factores inherentes** a la misma-. Aquí, la base lógica para la valoración de las experiencias se compone de seis grandes criterios:

1. La centralidad de los procesos de evaluación en el organismo de igualdad correspondiente: esta centralidad es necesaria para introducir, en primera instancia, una cultura y una práctica de evaluación que, posteriormente, pueda conducir a la normalización de su función dentro de la institución pública.

2. La adecuación al contexto y la claridad de los propósitos de las evaluaciones: la evaluación se da en un contexto determinado, con un fuerte carácter político, y al cual debe responder. La evaluación debe ser siempre contextual, partiendo y respondiendo al contexto en el que se ha producido. Los propósitos de la evaluación deben ser claros, estar previamente explicitados y responder a dicho contexto.

3. Claridad y centralidad del QUÉ se evalúa (los criterios de evaluación): es necesario que los criterios de evaluación estén claramente definidos, no sólo porque determinan las variables a estudiar, sino también porque a través de los mismos se construye la base lógica con la que se van a emitir los juicios posteriormente. Además deben ser el principal elemento organizador de los estudios de evaluación: de éstos ha de surgir la selección específica de técnicas e instrumentos para recopilar, analizar y valorar la información extraída.

4. Adecuada gestión de los recursos de evaluación, incluyendo lo siguiente:

- el buen aprovechamiento de las posibilidades que ofrecen los diferentes tipos de evaluación (incluidos la existencia de agentes externos), utilizando las sinergias de su utilización conjunta,

- el establecimiento de unas buenas estructuras de coordinación que permitan recoger información fiable y participada,

- la buena e imprescindible gestión de los tiempos y el calendario de la evaluación y,

- la inversión de suficientes recursos en la misma.

5. Elaboración suficiente de la información que se recopila y analiza en los procesos de evaluación. En evaluación, además de recopilar y analizar información sistemáticamente, se deben emitir juicios sobre la misma, a la luz de los criterios de evaluación previamente elaborados, y este proceso, debe concluir en una serie de orientaciones hacia la acción, para la mejora, la responsabilidad y el mejor conocimiento de las políticas evaluadas.

6. Buena comunicación tanto de los procesos como de los resultados de la evaluación: ésta es una condición necesaria para la posterior utilidad y utilización de la evaluación.

Por último, se identifican dos **factores relacionados con el impacto** de las evaluaciones. Primero, se valora la pertinencia y utilidad final de los procesos de evaluación metaevaluados, incluyendo cómo los procesos de evaluación se engarzan -o no- en las estructuras, funcionamiento y quehacer de las instituciones que los promueven. Segundo, se valoran las pocas posibilidades existentes de realizar una síntesis de resultados

de las evaluaciones de planes de igualdad de género, así como la conveniencia de realizar metaevaluaciones de los procesos de evaluación.

1. FACTORES PREVIOS QUE FACILITAN O DIFICULTAN LA EVALUACIÓN: CUESTIONES ESPECÍFICAS DE LAS POLÍTICAS DE IGUALDAD

En este punto se desarrollan el primer tipo de factores, que son denominados “factores previos” y que se organizan en dos polos. Por un lado, los factores que facilitan la evaluación: la concepción de planes, el aval internacional, la fortaleza institucional, el reconocimiento simbólico de la evaluación y la institucionalización de sus compromisos. Por otro, los factores que la dificultan: generalidad de los planes, carácter horizontal y transversal, su fuerte carácter persuasivo y no sancionador, la falta de cultura y experiencias previas, la falta de equipos preparados y la falta de normalización en el quehacer cotidiano de las administraciones públicas. A continuación se desarrollan más pormenorizadamente.

1.1. Los factores que facilitan la evaluación

La concepción y estructuración en planes de igualdad

Como ya se ha visto en el tercer capítulo, las políticas de igualdad de género en España se articulan, de forma generalizada en el nivel central y autonómico, en **planes de igualdad**. Estos son el principal instrumento por el que se configuran las políticas, y en ellos se integran toda una serie de medidas para promocionar la igualdad desde los diferentes sectores gubernamentales. Además, cuentan con una dimensión temporal, teniendo establecido un determinado periodo de vigencia.

La propia concepción de planes, como una realidad estructurada, es decir como un conjunto de objetivos y medidas ordenadas por sectores, y a cuyas acciones les son asignados responsables y recursos, facilita la evaluación. Independientemente del hecho de que los planes estén mejor o peor formulados, lo están de una forma estructurada, ordenada y global. En un documento, que además es aprobado por Consejo de Gobierno, se expresan las intenciones y las acciones que se pretenden llevar a cabo para hacer realidad dichas intenciones. La propia existencia de planes lleva a pensar en su seguimiento y evaluación, más que si se tratara de programas, medidas o acciones “sueltas”, con diferentes calendarios y distintas fundamentaciones.

La existencia de planes, en determinados periodos de vigencia que, además, debe ser continuada por planes sucesivos, ha hecho que las *femócratas* hayan comenzado a pensar en la evaluación de los planes de igualdad como algo necesario y deseable. No sólo se necesita tener información para la elaboración del plan siguiente, sino que se debe rendir cuentas también al Parlamento Vasco.

Ya en el capítulo tercero se mencionaban algunas características promovidas o facilitadas por el hecho de que existan planes de igualdad: una mayor conciencia de los poderes públicos, un mandato formal de informar sobre el desarrollo y los resultados del plan al Consejo de Gobierno que lo ha aprobado (especialmente si además hay un compromiso con el poder legislativo, como en el caso de *Emakunde*), un mayor compromiso de los mismos respecto a la igualdad de género, una visión global y estratégica, una forma de hacer operativa la idea de la transversalidad u horizontalidad, una mayor coordinación interdepartamental y una mayor potencialidad para la participación, así como para el propio seguimiento y la evaluación. Además del propio seguimiento y evaluación, queremos destacar aquí otras dos de estas características. En primer lugar, los planes hacen más conscientes a los poderes públicos de lo que están haciendo en materia de igualdad y, por lo tanto, fomentan también, de alguna manera, una evaluación de aquello que se realiza. Por otro lado, la coordinación interdepartamental hace no sólo que haya que actuar coordinadamente sino también que cada unidad gubernamental implicada en el plan necesita saber y conocer lo que están haciendo el resto de unidades. La evaluación, lógicamente, puede ser una vía para ello.

La institucionalización de compromisos respecto a los planes y a su evaluación

Además de “ordenar” la acción de los poderes públicos respecto a la igualdad de género y de aumentar la conciencia de los mismos sobre lo que hacen, los planes se aprueban en Consejo de Ministros o de Gobierno. Como se ha visto, en algunas comunidades autónomas también se debe informar al parlamento autonómico correspondiente. Estas circunstancias no sólo implica un compromiso formal importante respecto a la igualdad de género, sino también suele incluir un compromiso respecto a su evaluación, especialmente si existe la obligación de informar al legislativo (esto sólo ocurre en el País Vasco -está recogido en la propia ley de creación de *Emakunde*-, y de alguna forma, en Cataluña y Andalucía, en las que existen comisiones parlamentarias especialmente activas respecto a la igualdad de género).

El hecho de que se institucionalicen los planes tiene un cierto “efecto rebote” en la institucionalización de su evaluación. Aunque esto no está generalizado en la segunda generación de evaluaciones, que son el objeto de esta investigación, sí se generaliza para la tercera¹.

También el hecho de la horizontalidad o transversalidad de los planes hace que se hayan ido estableciendo mayores compromisos respecto a las responsabilidades que cada unidad gubernamental tiene en el plan. No sólo a cada una de estas unidades le va a interesar de algún modo la evaluación de dichas acciones de las que son responsables, sino que el propio organismo de igualdad hará lo posible por conocer si dichas responsabilidades han sido o no asumidas, y si lo han sido de modo efectivo.

¹ Véase capítulo 3.

En cualquier caso, el hecho de que existan compromisos respecto al seguimiento y a la evaluación de los planes, es un factor facilitador, aunque no sea condición suficiente para la misma.

Aval internacional y grado de acuerdo tanto en la necesidad de las políticas de igualdad de género como en la estrategia a utilizar

Ya se ha visto como la influencia internacional - sobre todo a través de la labor de los organismos internacionales y supranacionales- fue un factor importante en la entrada en la agenda política del problema de la desigualdad de género. Esta influencia se ha mantenido e incluso aumentado, no sólo a través de las instituciones creadas para ello (DAW, INSTRAW, UNIFEM, CEDAW, ETC.), sino también de las Conferencias Mundiales de la Mujer, organizadas por las Naciones Unidas y que desde 1975 se han ido realizando periódicamente (México, Copenhague, Nairobi y Pekín), y que han producido una serie de documentos que pretendían el compromiso de las naciones miembros.

El impacto que ha tenido la IV Conferencia Mundial de la Mujer en Pekín en 1995 ha sido especialmente importante y es conveniente destacarlo. Allí se adoptaron por consenso, por parte de las 185 delegaciones gubernamentales participantes -y, por lo tanto, el impacto es significativo por la cantidad de gobiernos representados- dos documentos: la Declaración de Pekín y la Plataforma para la Acción. La Plataforma de Acción es un documento en el que se describen los problemas más relevantes y los obstáculos que impiden el avance de las mujeres, el origen y las causas, los objetivos que hay que alcanzar para mejorar la situación de las mujeres y las medidas para conseguirlo. Para esto se describen doce áreas de especial interés, entre la que se encuentra la de “Mecanismos de igualdad” cuyo principal objetivo es el de crear o fortalecer mecanismos nacionales y otros órganos gubernamentales e integrar la perspectiva de género en las legislaciones, políticas, programas y proyectos estatales. Entre las medidas institucionales y financieras que se acuerdan está la de que los gobiernos, que son los principales responsables -aunque deben alentar a las ONG-, deben formular estrategias para el cumplimiento de la Plataforma, asignando los recursos suficientes. La Declaración de Pekín es el documento que contiene el mensaje político de los principales compromisos de los Estados miembros, y de ella destacamos algunos extractos:

“Es esencial diseñar, aplicar y hacer el seguimiento con la plena participación de las mujeres, de políticas y programas eficaces, eficientes y que se refuercen entre sí desde una perspectiva de género.” (art. 19)

“El cumplimiento de la Plataforma de Acción requiere del compromiso de los gobiernos y de la comunidad internacional...” (art. 21)

“El éxito de la Plataforma para la Acción también requerirá una movilización adecuada de recursos para el avance de las mujeres a nivel nacional e internacional.” (art. 36)

“Por la presente, adoptamos y nos comprometemos en calidad de Gobiernos a aplicar la siguiente Plataforma para la Acción, garantizando que en todas nuestras políticas y programas quede reflejada la perspectiva de género” (art. 38)

Como se puede observar, esta influencia internacional (en la que también incluimos para España la influencia de la Unión Europea, a través de sus Programas de Acción), avala y legitima las políticas gubernamentales de igualdad de género. Además del aval general de las políticas de igualdad, existe también un acuerdo en las estrategias generales a llevar a cabo, que serían, por un lado, la transversalidad de las acciones para conseguir una perspectiva de género en todas las políticas públicas (*mainstreaming*) y, por otro, el “empoderamiento” de las mujeres². Por último, el compromiso de los gobiernos por estas políticas y estrategias fomenta, en cierto modo, su evaluación, ya que, en teoría, los gobiernos deberían rendir cuentas sobre los avances logrados al respecto.

Todo este movimiento internacional, además de ser un acicate para la elaboración, ejecución y evaluación de políticas de igualdad de género, representa un “cuerpo teórico” en el que se sustentan dichas políticas y, por lo tanto, es un referente para su evaluación. Aunque sea de forma muy general, ya que son documentos y declaraciones cuya naturaleza requiere de formulaciones generales que sean capaces de concitar el consenso entre tantos gobiernos, la Plataforma y Declaración de Pekín y otros documentos internacionales contribuyen a la “teoría de programa” de las políticas de igualdad de género. Y esto siempre representa una excelente base para su evaluación.

La fortaleza y estabilidad institucional de los organismos de igualdad

A pesar de que la estrategia general del *mainstreaming* o de transversalidad de la perspectiva de género (es decir, del grado de acuerdo que existe en la idea de que el fin último es el de conseguir una perspectiva de género “normalizada” en todas las políticas públicas, y no hacer políticas verticales de promoción de la mujer), se sigue abogando, de forma unánime, tanto por la existencia de políticas específicas de igualdad de género, como por la permanencia de los organismos o mecanismos gubernamentales o institucionales específicos de igualdad.

Para el Grupo de Especialistas del Consejo de Europa³, el *mainstreaming* de género no puede ser reemplazado ni considerarse redundante con las políticas específicas de igualdad ni con los mecanismos u organismos específicos⁴. En realidad, cada vez que se menciona la estrategia del *mainstreaming* se hace hincapié no sólo en la complementariedad sino en la necesidad de políticas y organismos específicos. En realidad

² La identificación de estas dos estrategias, que van a marcar a partir de entonces las actuaciones de promoción de la igualdad de género, se considera como el gran fruto del proceso de trabajo que se consumó en la IV Conferencia Mundial en Pekín en 1995. (Cirujano, 2000: 114; López, 2000: 59-84).

³ *Council of Europe, Group of Specialist in Gender Mainstreaming*, 1998: 20-21.

⁴ Lo que en la literatura anglosajona se conoce como “*women’s equality machineries*”.

son consideradas como dos estrategias para un mismo fin: la igualdad de género. Para este grupo sí existen diferencias entre estas dos estrategias. Mientras el punto de partida de las políticas y los organismos específicos es un problema concreto resultante de la desigualdad de género, para el *mainstreaming* es una política que ya existe, de forma que se reorganice su proceso y se tenga el género en cuenta por parte de los actores que normalmente están implicados en dicha política. Para el Grupo de Especialistas, el *mainstreaming* es una estrategia fundamental y tiene un gran potencial para lograr un cambio sostenible, pero es más lenta. Las políticas y acciones específicas actúan de forma mucho más rápida, pero normalmente se limitan a áreas de política muy específicas.

Las razones por las que los organismos de igualdad siguen siendo imprescindibles, son cuatro:

a. En torno a los organismos de igualdad, que han sido capaces de desarrollar políticas y, a la vez, poner de manifiesto cuestiones importantes de igualdad, se ha ido acumulando una experiencia importante en cuestiones de género. De hecho, esta experiencia y conocimiento de las cuestiones de género es una condición necesaria para comenzar una estrategia de *mainstreaming*.

b. Los organismos de igualdad cumplen el importante papel de ofrecer el necesario apoyo a la estrategia del *mainstreaming*. El Grupo dice que esto es especialmente importante en los países en los que no están muy desarrollados los estudios sobre mujer o género. Además, reconocen que el *mainstreaming* es una nueva estrategia que requiere un mayor desarrollo y reflexión y que, incluso en los países en los que los organismos de igualdad no son el principal actor en dicha estrategia, éstos siguen reflexionando y diseminando conocimiento e información sobre temas de género a otros actores y a la sociedad en general.

c. Los organismos de igualdad son directos y específicos, por lo que parten de los desequilibrios existentes en las relaciones de género y se focalizan en cómo la desigualdad puede ser vencida. La estrategia del *mainstreaming* significa introducir la perspectiva de género en una política dada de forma que los efectos de dichas políticas sean más neutrales respecto al género, pero no tiene en cuenta, como punto de partida de las políticas, los desequilibrios reales de género. Por eso siempre se necesitarán las políticas y los organismos específicos que directamente traten con dichos desequilibrios.

d. La estrategia del *mainstreaming* se construye sobre las lecciones aprendidas de las anteriores experiencias en políticas de igualdad. De alguna forma se ha ido reconociendo últimamente que las políticas específicas no son suficientes para construir una sociedad que verdaderamente respete la igualdad de género. La estrategia del *mainstreaming* es el siguiente paso lógico a dar.

Para nosotros, además, existe otra vinculación definitiva entre ambas estrategias y es la idea de que, en realidad, el *mainstreaming* de género no es una estrategia que se haga partiendo de cero o en abstracto sino que, hoy por hoy, y todavía durante muchos años, es una estrategia que se hace y se hará desde las políticas y los organismos específicos de

igualdad. Al menos, este es el caso de España en donde, como ya se ha visto, las políticas de igualdad están articuladas en planes transversales que impulsan y coordinan los organismos específicos de igualdad. En definitiva, el *mainstreaming* se propone *desde* las políticas específicas de igualdad que son, a su vez, impulsadas y sustentadas *desde* los organismos de igualdad. Precisamente, el gran objetivo, tanto de las políticas específicas de igualdad, como de los organismos específicos, es conseguir la perspectiva de género en el resto de las políticas públicas. Una buena parte de sus energías y recursos irán, además de a políticas específicas, destinados a poner en marcha esta estrategia del *mainstreaming*.

Con esto, queremos hacer hincapié en que, no sólo los organismos de igualdad no están en peligro de “extinción” sino que están avalados tanto por las propias políticas de igualdad como por los organismos internacionales que las legitiman y las reclaman. A pesar de que parecen haberse quedado algo obsoletos en su denominación (siempre se trata de institutos o direcciones de la *mujer* cuando el foco de las políticas ha girado hacia las *relaciones de género*⁵), su consolidación e institucionalización en los diferentes niveles gubernamentales es un hecho. En el caso español, y teniendo en cuenta los niveles central y autonómico⁶, la consolidación e institucionalización de estos organismos ha sido evidente en la última década. De hecho, y salvo en contadas excepciones, el marco institucional se ha ido fortaleciendo según los dos factores analizados (tipo de organismo y posición en el organigrama gubernamental), es decir, las modificaciones que se han dado a lo largo de sus años de existencia, han sido siempre para otorgar a dichos organismos una mayor fortaleza y consolidación institucional.

En este momento, pues, se puede hablar en España de una “red” de organismos específicos de igualdad en el nivel autonómico y central fuerte y establemente consolidada, con una asignación presupuestaria y de personal, por lo general, razonable y con una buena legitimación institucional en sus respectivos gobiernos.

Esta consolidación ha ido acompañada, además, con un reconocimiento de dichos organismos como los principales actores de las políticas de igualdad de género. Incluso aunque estos organismos se consideren más como impulsores que como ejecutores directos de dichas políticas, su reconocimiento como “motores” y como “directores de orquesta” de todas las acciones en pro de la igualdad de género, resulta evidente. Y es precisamente este rol de principales “impulsores” de las políticas de igualdad el que provoca un papel clave como “evaluadores” de las mismas.

El reconocimiento simbólico de la evaluación en los últimos años

⁵ Véase el punto sobre “la evolución de la definición del problema” en el capítulo 3 de esta tesis doctoral.

⁶ Insistimos en que esta tendencia de consolidación e institucionalización de unidades de mujer o de género parece también ser la tónica en el nivel local, pero sobre esto hace falta hacer todavía más estudios (Valiente, 1998 b).

En los últimos años se ha ido reconociendo la necesidad de la evaluación, al menos simbólicamente, en torno a las organizaciones y las políticas públicas. Aunque este reconocimiento se quede sólo en el nivel simbólico y discursivo –se habla de la importancia y la necesidad de la evaluación, pero no se hace evaluación–, permite introducir el concepto de la evaluación de forma que se abren posibilidades para que ésta se dé.

De hecho, la evaluación se ha puesto de alguna manera de moda en los últimos años. Incluso en los discursos políticos se habla de evaluar las políticas públicas como una forma de rendir cuentas y como una parte inherente a la propia acción pública. Aunque ese discurso no se lleve a la práctica, sí representa un primer paso para que se hable, se permita y se intente llevar a cabo la evaluación.

Como ya se ha visto anteriormente, la propia concepción de planes de igualdad transversales agudiza, además, este discurso simbólico. Al haber tantas responsabilidades implicadas en dichos planes, resulta fácil recurrir a la necesidad de evaluar aquello que se ha hecho o no se ha hecho en cada una de las áreas o sectores.

Este discurso simbólico de la necesidad de evaluar se traduce, en muchos casos, en la intención de evaluar, y en dar algún tipo de respuesta que cubra dicha necesidad teórica. Aunque a veces las respuestas no sean las más adecuadas, sí se consigue el efecto de dar una respuesta general de hacer una evaluación. Y esto, además, tiene un efecto mimético, al igual que los propios planes. Si un plan de igualdad ha sido evaluado, se da un fenómeno en cascada de respuestas de evaluación en el resto. Esto es, sin duda, un primer paso. Al menos, la evaluación en sí no es cuestionada y goza de una legitimidad en torno a las políticas de igualdad de género.

1.2. Los factores que dificultan la evaluación

La generalidad de los planes

Ya se ha visto cómo los planes de igualdad de género consisten en un conjunto de objetivos y medidas, aprobadas por Consejo de Ministros o de Gobierno, y que deben llevarse a cabo por parte de las diferentes unidades administrativas implicadas. Los primeros planes eran más bien un conjunto de intenciones, en muchos casos poco concretadas o especificadas. Por tal razón, su interpretación podía ser múltiple o simplemente nula (quedándose en una simple declaración de intenciones que no se concretaba en ninguna acción concreta) y, sobre todo, eran muy poco evaluables ya que existían pocas referencias sobre lo que significaba su cumplimiento. Este grado de generalidad se ha ido corrigiendo a lo largo del tiempo, entre otras cosas y como ya se ha señalado anteriormente, debido precisamente a los intentos de evaluación. Sin embargo,

todavía existe en los planes un cierto grado de generalidad, especialmente si se tienen en cuenta los planes de segunda generación que son los que son metaevaluados en este trabajo.

Este grado de generalidad de los planes dificulta su evaluación, ya que lo que denota en realidad es una falta de claridad en los objetivos, no saber exactamente lo que se quiere hacer o no saber cómo hacerlo. Como se ha señalado antes, lo que ocurre es que las interpretaciones de lo que realmente se pretende hacer pueden ser múltiples y, por lo tanto, la evaluación, sobre todo el establecimiento de los criterios de evaluación (QUÉ se va a evaluar), resulta una tarea ardua. En estos casos, es normalmente el propio proceso de evaluación el que acaba definiendo (en realidad, reconstruyendo) los términos de referencia con los que se formuló el plan.

En algunos casos, un cierto grado de generalidad ha sido, inevitable, ya que los planes tuvieron que ser aprobados por toda una serie de agentes críticos quienes, en su mayoría, no tenían muy elaborado ni asumido un discurso de género. Así, la ambigüedad de los planes es lo que a veces preserva de encontrar problemas en su aprobación. Un conjunto de intenciones generales es aceptable por parte de todo el mundo, máxime si el discurso de género a adquirido un alto valor simbólico, como es el caso. Un conjunto de medidas concretas puede encontrar más problemas, aunque eso asegure de una manera más firme su cumplimiento.

Pero insistimos en que la generalidad de los planes, en nuestro parecer, denota sobre todo, ausencia de lo que podríamos denominar “teoría de programa”. Recordamos aquí⁷ que por teoría de programas se entiende el conjunto de premisas o creencias que subyace a la acción, es decir el conjunto de hipótesis sobre las que la gente construye y planifica los programas o los planes (Weiss, 1998a). Es una explicación de las relaciones causales que unen las entradas o *inputs* de un programa con sus salidas, *outputs* o resultados esperados, un modelo sobre cómo un programa se supone que debe funcionar. Por ejemplo, ¿por qué la elaboración de un programa de educación sexual y planificación familiar para jóvenes contribuye a favorecer cambios permanentes en las actitudes y comportamientos de y hacia las mujeres jóvenes⁸?, o mejor aún, ¿cómo debe ser ese programa de educación sexual para que realmente favorezca cambios permanentes en las actitudes o comportamientos?, e incluso, ¿por qué se escogen, precisamente, acciones de promoción del deporte, de información en Centros de Enseñanzas Medias y de educación sexual para favorecer el cambio permanente de actitudes y comportamientos?, ¿por qué esas y no otras? Todo esto es lo que creemos está relacionado con la “teoría de programas” y es lo que normalmente falla cuando un plan queda formulado de forma general.

Lógicamente, a mejor “teoría de programas”, es decir a mayor fundamentación de las acciones que se proponen, más fácil será la evaluación, ya que se contará con puntos de referencia claros para el establecimiento de criterios, la recopilación y el análisis de la

⁷ Véase Capítulo 2, enfoques o modelos de evaluación.

⁸ Ejemplo escogido del II Plan para la Igualdad de Oportunidades de las Mujeres de la Comunidad de Madrid 1993-1995 (Objetivo 6, acción 6.5).

información y su posterior valoración. De hecho, los autores que propugnan un enfoque de evaluación basado en la teoría (Chen 1990; Chen y Rossi, 1983; Bickman, 1990 y la propia Weiss, 1998) dicen que resulta prácticamente imposible realizar la evaluación de un programa si previamente no se ha establecido su “teoría de programa”.

Así, en la segunda generación de planes de igualdad de género en España no hemos encontrado una “teoría de programas” especialmente fundamentada o explicitada. Entre otras cosas, hace falta mucho más desarrollo de la teoría de las relaciones de género y, sobre todo, de cómo traducir en la práctica acciones realmente tendentes a una mayor equidad. Esto hace que no haya habido una clara teoría sobre la que fundamentar los estudios y los procesos de evaluación. Como ya se ha visto antes, las interpretaciones pueden ser múltiples y, en muchos casos, la dificultad a la hora de evaluar significaba en realidad no saber qué había que evaluar, ya que no se tenía una base lógica en la cual fundamentar los criterios de evaluación. De esta manera, se ha tendido a analizar simplemente el “grado de cumplimiento” de las acciones sin cuestionarse esas mismas acciones y su sentido o si realmente contribuían, y en qué medida, al cumplimiento de los objetivos.

Su carácter de horizontalidad e integralidad

Los planes de igualdad tienen un carácter integral, es decir, cubren toda una serie de áreas o aspectos muy diversos entre sí pero que, a su vez, están relacionados. El “*leit-motiv*” y punto común es la promoción de la igualdad de género -o la eliminación de la discriminación basada en las relaciones de género- en todos los aspectos tratados: educación, empleo, salud, participación social y política, medioambiente, mundo rural, etc. Este carácter integral hace que los planes sean, en sí mismos, complejos, simplemente por la cantidad de aspectos que cubren y cómo se relacionan unos con otros. Esta complejidad hace más difícil la evaluación, entre otras cosas, por la propia teoría subyacente a la acción -la “teoría de programas”-, la cual debe ser amplia para cubrir todas las dimensiones de las que son objeto los planes.

Por otro lado, la discriminación que está en la base de las relaciones de género es algo en sí mismo muy complejo y profundamente enraizado en las propias estructuras sociales. Combatir dicha discriminación es también, pues, complejo, ya que implica una mirada distinta a la realidad, una forma diferente de analizarla y abordarla. Por lo que la acción que los poderes públicos se propongan para abordar dicho problema es también compleja y, por lo tanto, difícilmente evaluable. Por último, el carácter innovador de estas políticas también contribuye a moverse en un terreno desconocido y en el que hay que elaborar “*ex-novo*” los parámetros a través de los cuales se va a juzgar dicha acción pública.

El carácter horizontal viene determinado por la idea de canalizar las acciones a favor de la promoción de la equidad de género, a través de las políticas públicas “regulares”, huyendo de realizar una única política “vertical” para las mujeres. Este carácter horizontal, como ya se ha visto, se refuerza con la estrategia del *mainstreaming*, avalada por la

Plataforma de Acción de la IV Conferencia Mundial de las Mujeres de Pekín (1995), cuestionando la neutralidad de género de las políticas públicas y proponiendo acciones para conseguir la perspectiva de género en dichas políticas públicas.

Es indudable que esta “horizontalidad” o transversalidad de las políticas de igualdad de género hace la acción más compleja. En primer lugar, necesita de la implicación de muchos actores y agentes críticos -en torno a cada una de las políticas en las que se pretende incidir (educativa, sanitaria, medio ambiente, de empleo, etc...)- que deben estar convencidos y asumir los objetivos que se proponen. Los planes, además de incluir diferentes sectores o dimensiones -cuyas competencias están distribuidas en múltiples unidades administrativas- suelen incluir también la participación de ONG y del movimiento de mujeres, además de diferentes niveles de gobierno, estableciéndose una compleja red de relaciones intergubernamentales. Pero, además de asumir los objetivos, estos múltiples actores deben ponerse de acuerdo y trabajar coordinadamente para llevar a cabo las acciones precisas. En términos de Pressman y Wildavsky se trataría de la “complejidad de la acción conjunta”. Por lo tanto, no sólo es compleja la elaboración y formulación de estas políticas, sino que, sobre todo, es compleja su ejecución o implementación y, por lo tanto también, su evaluación. Esto hace que existan no sólo una serie compleja de dimensiones a evaluar, sino también un contexto de múltiples relaciones que debe ser tenido en cuenta necesariamente a la hora de evaluar estas políticas.

Su carácter de sensibilización y persuasión

Debido a su carácter transversal y horizontal, las *femócratas*, desde los organismos específicos de igualdad, deben persuadir a sus compañeros y compañeras de gobierno para que asuman los objetivos que desde las políticas de igualdad de género se proponen. Los planes de igualdad, además, no tienen ningún poder sancionador, por lo que ese poder de persuasión acaba siendo una de las estrategias más importantes con la que deben trabajar las *femócratas*. Ya que no tienen ninguna autoridad sancionadora, deben convencer a los demás de que hagan suyos los objetivos que ellas proponen. Esencialmente, es importante que se hagan aliados y aliadas en los diferentes departamentos e instituciones. Porque se trata de conseguir, finalmente, que las acciones las hagan otros. Esto complica aún más la acción y hace que exista una acción política específica -la de la persuasión y el convencimiento- que se convierte en crucial en las políticas de igualdad de género. Dicha estrategia, su diseño, ejecución y resultados debe ser también evaluada, entre otras cosas porque esto puede ser de extrema utilidad para la determinación de los éxitos y los fracasos y para la mejora de las políticas de igualdad de género. ¿Por qué unas unidades administrativas son más sensibles que otras a las cuestiones de género? ¿Cuánto cuenta el factor personal de los responsables en el nivel político y en el técnico? ¿En qué consiste dicho factor personal? ¿Cuáles son los factores que hacen que finalmente un conjunto de personas o un departamento se sensibilice y actúe en consecuencia? ¿Cuáles son las estrategias más adecuadas para dicha sensibilización y persuasión? ¿Son siempre las mismas en diferentes circunstancias? De esta manera, si no se tiene en cuenta y se evalúa

esta dimensión aquí explicada, no se va a comprender nunca una buena parte de la acción transversal en pro de la igualdad de género.

Sin embargo, dicha dimensión o componente de sensibilización y persuasión es harto compleja de evaluar. En primer lugar, hay que definirla y tipificarla: en qué consiste y cómo se lleva a cabo. Y después evaluarla, en su proceso de funcionamiento y sus resultados y, a su vez, unirla y relacionarla con la acción propia de los planes de igualdad. Esta dimensión, además, no ha sido contemplada en ninguna de las experiencias de evaluación metaevaluadas en este trabajo, aunque es una de las preocupaciones expresadas más recurrentes y constantes de las responsables políticas y técnicas de los organismos de igualdad. Aquí reiteramos la necesidad de incluir esta dimensión en la evaluación de las políticas de igualdad de género, aunque esto la haga más compleja.

Falta de cultura y experiencia previa, institucionalización y recursos para la evaluación

Otra dificultad añadida a la evaluación de los planes de igualdad ha sido la falta de experiencia previa en procesos de evaluación y, en general, la falta de una cultura de la evaluación.

La falta de experiencias previas en evaluación de planes de igualdad se traduce en una carencia de referencias o modelos para la evaluación de las políticas de igualdad. Entre otras cosas, los planes de igualdad son realidades muy recientes en el momento en el que se aborda su evaluación y dicho carácter novedoso agudiza aún más dicha falta de referencias. En definitiva se trata de “inventarse” cómo evaluar dichos planes ya que nadie o casi nadie lo había hecho antes.

Además, nos encontramos con una escasísima, si no nula, tradición evaluadora en las administraciones públicas españolas. Esta falta de cultura de evaluación revierte, entre otras cosas, en la ausencia de modelos y de bagaje teórico, así como de un “saber hacer” en evaluación de programas. Esta es la razón por la que, en general, se ha recurrido a modelos utilizados en metodología de investigación en ciencias sociales en general, no sabiendo adaptar o no aplicando modelos y “saber hacer” específico sobre evaluación de programas que sí se había ido desarrollando en otros países y en el nivel internacional.

Esta ausencia de cultura evaluativa también se refleja en la escasa institucionalización de la función de evaluación en España en general. Entre las posibles causas del retraso y de la falta de desarrollo de la función de evaluación se pueden mencionar: que la evaluación llega a España en tiempos de crisis económica y escasez de recursos, que el sistema parlamentario español, tanto en el nivel central como en el autonómico, se basa en una fuerte disciplina de partido que tiende a limitar la función de control del Parlamento, que los Tribunales de Cuentas limitan su papel a una exclusiva supervisión de las cuentas públicas, que el número de profesionales especializados en evaluación es escaso y, que los sistemas de información internos en las administraciones públicas, a pesar de los avances, no son potentes ni del todo adecuados para la evaluación (Ballart, 1993: 221-2).

Ya se ha comentado que en la tercera generación de planes la función de evaluación está algo más institucionalizada ya que se menciona explícitamente y suelen darse compromisos respecto a la evaluación de los planes en el propio documento de aprobación. Sin embargo, éste no era el caso generalizado en la segunda generación que es la que es objeto de este trabajo. Es decir, las evaluaciones realizadas se han hecho sin que hubiese un compromiso o un mandato institucional para ello. De hecho, muchos planes de igualdad no han sido siquiera evaluados. Esta falta de institucionalización dificulta y a veces hace poco comprensible la función de evaluación. No obstante, si se compara con otros sectores de políticas, la posición relativa de las políticas públicas de igualdad de género a nivel central y autonómico, no es tan mala respecto a la función de evaluación: de hecho, aún con grandes deficiencias, se han dado una serie de experiencias y, como ya se ha visto, en la tercera generación de planes está contemplada la evaluación de los mismos de una forma u otra.

Sin embargo, queda aún mucho camino por recorrer para que la función de evaluación se desarrolle con todas sus potencialidades para las políticas de igualdad. Y para acelerar este proceso, los poderes públicos, y en especial las *femócratas* al frente de los organismos de igualdad, deben reconocer la función de evaluación como una de las prioridades de dichos organismos. Esto implicaría dedicar esfuerzos y recursos, formación del personal y búsqueda de ayuda externa especializada; pero también supondría disposición para el intercambio, la reflexión y el aprendizaje. Esta falta de experiencia y de cultura se concreta especialmente en una escasez de recursos para la evaluación, ya que al no ser una función reconocida y normalizada se considera como un “gasto extra”. Ya se ha visto en el capítulo anterior la parquedad de recursos empleados para la evaluación de los planes de igualdad, sobre todo si se analizan las cantidades en términos relativos, comparándolos con los recursos empleados para otras funciones. Como se verá en el punto siguiente, la falta de recursos y, sobre todo, la falta de disposición para emplearlos es uno de los factores que más dificultan las evaluaciones y condicionan su calidad.

Por lo tanto, existe una necesidad de aumentar la cultura y competencia evaluativa, como factor previo que facilite los procesos de evaluación. En este sentido, conviene señalar el potencial papel que pueden cumplir no sólo los organismos de igualdad, sino también otros actores del *policy network* o redes de políticas, sobre todo, las personas o colectivos expertos en temas de género, políticas de igualdad y evaluación (como son seminarios universitarios, institutos de investigación y áreas especializadas).

Falta de equipos preparados

Al no existir cultura ni práctica de la evaluación en España, tampoco existen profesionales preparados y entrenados para realizarla. Esta es una de las dificultades que mencionan algunas de las responsables políticas entrevistadas. Una de ellas mencionaba especialmente la dificultad de encontrar especialistas en evaluación que, además, tuvieran conocimiento -y sensibilidad- sobre temas de género. Como se ha visto antes la forma que en algunos casos se ha tomado para suplir esta carencia ha sido atraer a especialistas en

metodología de investigación social y no tanto en evaluación de programas. Y esta circunstancia ha condicionado enormemente la forma en la que se han realizado algunas de las evaluaciones analizadas.

A la falta de especialistas externos hay que añadir la falta de formación y preparación específica en evaluación de los equipos técnicos internos en los organismos de igualdad. Esto se ha ido supliendo de alguna manera por parte de los organismos de igualdad, facilitando la formación en temas relacionados con la evaluación de las técnicas responsables. Pero el tipo de formación específica recibida ha sido, por lo general, escasa y de corta duración, en parte sobre todo debido a la propia carencia de oferta de este tipo de formación en España⁹: hoy por hoy, todavía resulta difícil formarse como evaluador en España y todavía no existe suficiente demanda de evaluadores en el mercado. En los próximos años es muy probable y esperable que la evaluación vaya consolidándose como profesión y como actividad científica y académica diferenciada, apareciendo asociaciones profesionales, publicaciones especializadas y formación específica. Pero esto todavía no ha ocurrido, ni ocurría durante los años en los que se diseñaron y realizaron las evaluaciones analizadas.

Y ya se ha comentado anteriormente (capítulo 1, punto 1.3.1) cómo, en nuestra opinión la formación y habilidades de un evaluador deben ser de alguna manera diferentes a las de un investigador: debe saber manejarse con habilidad en contextos complejos en los que debe intervenir, ser un buen conocedor de los procesos políticos subyacentes a los programas o políticas, conocimiento y capacidad de trato con grupos y personas, habilidades de comunicación y negociación, ... También ha de ser un buen metodólogo, especialmente preparado para las cuestiones específicas de la evaluación: establecimiento y definición de criterios de evaluación, establecimiento de consenso en torno a los mismos, emisión sistemática de juicios y buena comunicación durante el proceso y de los resultados de la evaluación. Además, claro está, de conocer las técnicas de recopilación y análisis de la información para aplicar la más apropiada en cada momento. En definitiva, se requieren equipos específicamente preparados en evaluación de programas. No poder contar con ellos también añade dificultades a la tarea de evaluar.

No pensar la evaluación a tiempo, debido a la falta de normalización de su función

Ya se ha visto como, en general, se hace poca evaluación de programas, planes y políticas en las administraciones públicas españolas. Esto hace que esté muy poco normalizada la función de evaluación y que se considere más como una actividad “extra”.

⁹ Prácticamente no existe formación específica de post-grado en evaluación en España. Durante dos cursos académicos (1998-1999 y 1999-2000) se realizó un Experto en Evaluación de Políticas, Programas y Proyectos de Cooperación al Desarrollo (IUDC-Universidad Complutense de Madrid) que ha acabado por reconvertirse en un curso general sobre la Gestión del Ciclo de Proyecto ante la demanda laboral en ese campo.

Además, en algunos de los casos analizados, era la primera vez que se hacía evaluación en el seno de los organismos de igualdad.

Por otro lado, y debido a esta falta de normalización de la función de evaluación o, si se quiere la falta de costumbre de evaluar, existe una tendencia a pensar en la evaluación exclusivamente de forma finalista, es decir, se piensa en la evaluación como la última fase a tener en cuenta en un plan o programa. Esta circunstancia hace que, en muchos casos, se piense en la evaluación sólo al final del período de vigencia del plan que se pretende evaluar. El hecho de que exista una última fase final de evaluación no obvia el que haya que pensar en la evaluación desde un principio, en realidad, desde el mismo diseño del plan. No hacerlo significa que la evaluación pierda una buena parte de su potencia y potencial para ser útil. Si no se definen los criterios de evaluación al principio -aunque luego se tenga la suficiente flexibilidad como para ajustarlos sobre la marcha- y no se establecen los mecanismos para la recopilación de información, será muy difícil recopilar una información que, al final, se hace irrecuperable. Esto se agudiza si no existen unos buenos sistemas de información y el plan no está bien documentado, como en la mayoría de estos casos ocurre (ya que mucha de la información y su registro depende de otras unidades administrativas distintas a los organismos de igualdad).

Además, el hecho de pensar la evaluación desde un principio ayuda a reforzar los propios objetivos del plan, ya que si se piensa sobre qué se quiere evaluar se está haciendo un repaso de la teoría subyacente al plan, sobre lo que es o no importante, a lo que se debe dar prioridad y, sobre todo, este ejercicio de pensar en los criterios de evaluación ayuda a hacer los objetivos menos abstractos y ambiguos. Sin embargo, si esto se hace al final, además de la irrecuperabilidad de una información muy valiosa, el “repaso” antes mencionado se hace “a posteriori”, teniendo mucha menos fuerza. Es más, pensar en la evaluación hacia el final lo que hace es contribuir a no cuestionar ni los objetivos ni el plan en su conjunto, sino más bien limitarse sólo al “grado de cumplimiento” de dichos objetivos, es decir, saber si se han hecho las actividades o acciones previstas., limitando enormemente el potencial de la evaluación.

Por lo tanto, el hecho de anticiparse a la evaluación, debido principalmente a su falta de normalización como función propia de las administraciones públicas, es uno de los factores que más dificulta la misma.

2. FACTORES INHERENTES AL PROPIO PROCESO DE EVALUACIÓN

En este punto se exponen los factores que tienen que ver con el propio proceso, es decir, que son inherentes al mismo y que determinan claramente el producto evaluativo. En este caso no se trata de cuestiones relacionadas con la cultura organizativa de las organizaciones responsables de la evaluación, con el diseño o forma de ejecución de los planes a evaluar o con la falta de normalización de su función y la ausencia de profesionales para llevarla a cabo, sino de factores que tienen que ver directamente con la forma en la que se producen los procesos de evaluación. Es decir, no nos referimos aquí a factores en torno a las evaluaciones sino a aquellos directamente relacionados con su diseño y ejecución.

En este apartado, pues, señalamos los factores que, en nuestro parecer, han influido más, y de forma más directa, en los procesos y los productos de evaluación analizados. De esta manera aquí se señalan la falta de claridad en los criterios de evaluación, los problemas relacionados con los flujos de información y con los calendarios de los procesos, la importancia de los agentes responsables de dichos procesos, el déficit de elaboración de conclusiones, valoraciones y recomendaciones, además de la comunicación deficitaria.

2.1. Falta de claridad en el propósito de la evaluación: Evaluación vs. investigación de políticas de igualdad o la necesidad diferenciada de ambas actividades

El primer factor inherente a los procesos de evaluación que influye negativamente en los mismos es la falta de claridad a la hora de plantearse la evaluación: ¿PARA QUÉ y POR QUÉ se hace la misma? ¿Cuál es su función? ¿Se hace realmente para mejorar los planes y políticas evaluadas, ser más conscientes de ellas y controlar responsabilidades? ¿Se hace verdaderamente con la intención de facilitar la toma de decisiones sobre los planes y para facilitar el aprendizaje organizativo? Más allá: ¿alguien se ha planteado conscientemente estas cuestiones? Con esto queremos decir que, en general, no es que las responsables de las políticas de igualdad de género se hayan planteado dichas preguntas y, deliberadamente, hayan contestado negativamente a las mismas, sino que muy probablemente ni siquiera se las hayan formulado. Como se desarrolla en el punto 9.1.2. de este capítulo, una de las razones fundamentales de esta falta de conciencia es una función de evaluación muy poco desarrollada en las administraciones públicas españolas, que se traduce en una casi total ausencia de cultura y práctica evaluativa.

Un indicador evidente de esta falta de claridad es el propio planteamiento general de algunos de los estudios de evaluación. Una de las primeras conclusiones que se obtienen del análisis de las experiencias de evaluación de los planes de igualdad, es la necesidad de

distinguir entre hacer investigación sobre la situación de la mujer y hacer evaluación de los planes y las políticas de igualdad de género. En general, estos dos extremos tienden a confundirse y en algunos casos no llegan a diferenciarse bien, e incluso, en algunos casos, se han diseñado las evaluaciones de los planes en términos de investigación de la situación de las mujeres. Este último ha sido el caso, sobre todo, de la Comunidad de Madrid, la Junta de Castilla y León y, en parte, del propio IM. En estas evaluaciones se pretendía evaluar el impacto de los planes: para ello se establecieron una serie de medidas o indicadores y se midieron los mismos al inicio y al final del plan, analizando las diferencias o evolución durante el período de vigencia del mismo. Evidentemente, este tipo de indagación proporciona información sobre la evolución de la situación de la mujer en una serie de variables, pero no permite ningún tipo de conexión entre dichos datos y el plan. Es decir, dicha evolución puede haber sido producida por muchas otras variables, además del plan (la propia evolución de la sociedad). En otras palabras, no se puede llegar a la conclusión de que, en la ausencia del plan, no hubiera pasado lo mismo, ni se ofrece ninguna información rigurosa sobre los “efectos netos” del plan (los efectos debidos al plan y exclusivamente al plan y que sin él, no se hubieran producido). En el caso de las evaluaciones de Madrid y Castilla y León (en esta última se utilizan los mismos indicadores que en Madrid, ya que se realiza por parte del mismo equipo externo), los indicadores se establecen tomando como base el plan; es decir, y en este caso, las variables que definen la situación de la mujer se definen por los aspectos que los poderes públicos han considerado susceptibles de mejora o cambio. Aún así, el hecho de que se recoja información sobre, por ejemplo, el mayor número de mujeres universitarias al final que al principio del período de vigencia del plan, o el igual número de mujeres empleadas en la Comunidad de Madrid, no significa ni mucho menos que haya sido el plan el causante o provocador de esta situación, no por su potencial capacidad para hacerlo, sino porque no se ha obtenido ninguna información que permita establecer una mínima conexión entre los datos y las acciones de los poderes públicos. En el caso de la “evaluación cuantitativa” del II Plan central del IM, las variables medidas son más específicas y tienen que ver con una investigación concreta respecto a los estereotipos sexuales, valores, actitudes, calidad de vida y usos del tiempo; aquí también se recoge información, a través de dos encuestas en el inicio del plan y el final y se analiza la evolución de una serie de índices relacionados con las variables estudiadas.

En ambos casos (Madrid y Castilla y León, por un lado y el IM, por otro), se trata de excelentes investigaciones sobre la situación social de las mujeres. Incluso en el caso del IM, es una investigación de tipo innovador, en la que se construyen una serie de índices muy interesantes para medir la (des)igualdad entre los géneros, basados en aspectos normalmente no recogidos como son valores, estereotipos y actitudes. Además, tomando como base los usos del tiempo se construyen unos indicadores o índices de igualdad social, calidad de vida, autonomía personal, trabajo doméstico y productividad. Pero es evidente que no se trata de evaluaciones de los planes que se supone están evaluando. La función de investigación sobre la situación de la mujer es una de las funciones específicas y más importantes de los organismos de igualdad. En nuestra opinión es necesario llevar a cabo investigaciones, pero no hay que “disfrazarlas” de evaluación, no es bueno ni para la función de investigación ni para la de evaluación, ya que en realidad, ni se están dedicando los suficientes recursos a la investigación, ni se están haciendo realmente evaluaciones.

Como se ha visto en el primer capítulo de esta tesis doctoral, existen muchas razones para diferenciar las funciones de investigación y evaluación, pero aquí se resaltan especialmente dos. En primer lugar se diferencian en el objeto, siendo el objeto de la investigación la situación de la mujer y su evolución a lo largo del tiempo (es decir, el problema público de la desigualdad de género) y el objeto de la evaluación, los planes, es decir cómo abordan los poderes públicos el problema de la desigualdad y cómo influyen dichas actuaciones en el problema público. La evaluación debe centrarse en la respuesta que los poderes públicos dan a un determinado problema, y no en ese problema en abstracto. En segundo lugar, también resaltamos las diferencias entre las funciones que ambos tipos de indagación pueden cumplir: la evaluación debe cumplir especialmente dos funciones que son la de mejora de la actuación de los poderes públicos y la del control de responsabilidades. La investigación pretende, sobre todo, construir conocimiento y contribuir genéricamente al mejor conocimiento y definición de los problemas que los poderes públicos deben abordar. En el caso de las políticas de igualdad de género, es evidente que la investigación sobre la situación de la mujer contribuye, definitivamente, a una mejor formulación de las políticas de igualdad. Pero eso no obvia la necesidad de evaluación de dichas políticas. Una cosa es monitorizar el problema que hay que abordar y otra, que es precisamente lo que le confiere un carácter específico a la evaluación, valorar la actuación de los poderes públicos para abordar dichos problemas. Esta valoración debe hacerse por una cuestión de responsabilidad pública -los poderes públicos deben rendir cuentas y responsabilidades sobre su gestión- y también, porque la evaluación es una fuente de aprendizaje para ser más conscientes y conocer mejor las actuaciones que se realizan, pero además también para mejorarlas a través de información sobre los procesos, los éxitos y fracasos, etc. La continua monitorización del problema a abordar -que lo puede cumplir la investigación- es necesaria siempre como una fuente importante para el reajuste de las políticas; pero dicha actividad no cumple las funciones que puede y debe cumplir la evaluación de dichas políticas y planes, ya que es una fuente de información diferente.

Resulta evidente que tanto la función de investigación como la de evaluación son clave para los organismos de igualdad. Y que ambas actividades se benefician y potencian mutuamente. El hecho de que existan “observatorios de igualdad”, cuyo objetivo es recoger datos periódicamente sobre la situación de la mujer en cada uno de los territorios, resulta sumamente útil e interesante para la evaluación de las políticas de igualdad, pero el hecho de su existencia no implica que se esté haciendo evaluación. Por otro lado, la evaluación de las actuaciones, planes y políticas ofrece una información que puede contribuir sustancialmente a la identificación de variables a ser estudiadas en posteriores investigaciones. Mientras no se diferencien estas dos funciones de investigación y evaluación y a cada una se le otorgue el papel que les corresponde, no se aprovechará el enorme potencial que ambas tienen en la elaboración de las políticas públicas de igualdad de género.

2.2. Falta de una visión global de la actuación de los poderes públicos ante el problema de la desigualdad de género: la evaluación de las políticas vs. la evaluación de los planes

Todas las experiencias de evaluación consignadas en torno a las políticas de igualdad de género central y autonómicas han sido, o han pretendido ser, evaluaciones de los planes de igualdad aprobados por los respectivos gobiernos¹⁰. Es decir, ninguna de las experiencias de evaluación se ha planteado como una evaluación más global de la política de igualdad, sino como una evaluación concreta de un determinado plan. Incluso, considerando esa evaluación concreta de planes, uno de los problemas detectados en las mismas es que se tiende a hacer evaluación de las medidas o actuaciones contempladas en los planes, y a veces se carece de una visión más global que tenga en cuenta el plan en su conjunto.

Es cierto que los planes de igualdad formulados pretenden cubrir, impulsar y coordinar todas las actuaciones gubernamentales en materia de igualdad de género, por lo que se puede concluir que la principal forma de evaluar las políticas de igualdad es evaluar dichos planes. Esto es evidente, pero creemos que es importante no perder de vista la idea de evaluar las políticas, y no sólo los planes de igualdad, aún cuando éstos sean el principal instrumento a través del cual se articulan dichas políticas.

La razón de insistir en esto es que la evaluación de políticas representa una visión más global que incluye, además, la evaluación de los propios planes y que facilita una mirada que normalmente no suele desarrollarse en las evaluaciones de los planes, que resulta una visión más estrecha o “encajonada”. La evaluación de los planes facilita el cuestionamiento del desarrollo y de los resultados de dichos planes, pero no suele cuestionar el papel de los planes ni los planes en sí mismos. Cuestiones relacionadas con la adecuación o no de los planes de igualdad como instrumento propio de las políticas, con la efectividad e idoneidad de la estrategia transversal y con el papel y el desarrollo de las funciones propias de los organismos de igualdad, incluyendo el desarrollo adecuado de la función de evaluación (y de la de impulso, y de la coordinación, etc.), son cuestiones que no se han planteado en ninguno de los procesos de evaluación estudiados.

La distinción entre programas o planes y políticas públicas no suele encontrarse de una manera clara en la literatura al uso sobre evaluación. Normalmente se habla indistintamente de evaluación de programas y evaluación de políticas. Sin embargo, se pueden encontrar algunas claves para su distinción en la discusión que realiza Carol Weiss sobre la “teoría de programa” (Weiss, 1998a: 46-71). El diseño de políticas tiene que ver con responder a problemas públicos sobre la base de alguna “teoría de política” que define las condiciones iniciales (cuál es el problema), los efectos finales (cuál es el objetivo), y los medios para llegar de unas a otros (acciones de programa o política). Estas acciones son las que se traducen en programas o planes y consisten en teorías de programa que se deben

¹⁰ En algunas comunidades autónomas se han realizado evaluaciones parciales de programas concretos, como es el caso de Asturias, del programa de talleres itinerantes socioculturales (1996), y los programas NOW (que en el caso de esta Comunidad Autónoma se realizan por parte del mismo que los gestiona). En general, todos los programas que cuentan con fondos europeos suelen requerir la existencia de una evaluación, que en muchos casos se da como una mera formalidad para “cubrir el expediente”, pero que en otros sirve para proporcionar una información útil e ir generando una cierta cultura evaluativa en torno a los organismos de igualdad. Pero estas experiencias de evaluación siempre se refieren a programas o actuaciones muy concretas, y ni siquiera se plantean como una evaluación parcial de los planes.

articular de alguna forma con el marco de política más amplio que es el que le da su razón de ser o fundamentación. Este “enganche” o reconstrucción de la “teoría de programas” surge cuando se parte de la visión más amplia que supone la perspectiva de la evaluación de políticas. Y de esto todavía están muy lejos las experiencias de evaluación analizadas, en las que como se ha comentado antes, predomina la visión de evaluar exclusivamente las medidas de un plan, y no tanto el plan como una unidad global.

Estas cuestiones relacionadas con la teoría de los programas y de la implementación de los mismos¹¹ pueden, desde luego, plantearse en evaluaciones de planes más completas, que contemplen, sobre todo, el propio diseño de los planes y las variables relacionadas con el desempeño de los organismos de igualdad en el proceso de puesta en práctica de dichos planes. Pero insistimos en que la visión más amplia y global de la evaluación de las políticas de igualdad, resulta especialmente interesante para políticas públicas innovadoras (y éste es el caso de las políticas de igualdad), y siempre va a cubrir más aspectos que la evaluación de los planes. Es evidente que este tipo de evaluación es una evaluación que interesa, sobre todo, a los responsables políticos y gubernamentales pero que, en general, puede hacer más consciente a la “comunidad de política”¹² de lo que tiene entre manos.

2.3. Falta de reconocimiento del carácter político de la evaluación

Esta conclusión tiene también que ver, al igual que la primera, con la falta de claridad en el fin y las funciones de la evaluación. El discurso sobre las funciones de las evaluaciones analizadas contiene, por lo general, razones suficientemente bien esgrimidas sobre la necesidad del rendimiento de cuentas, la mejora de los planes y la iluminación para acciones futuras, aunque en muchas ocasiones estas razones se quedan en el nivel discursivo y no llegan realmente a desarrollarse y cumplirse en la práctica. Quizá la función que mejor se cumpla, por encima de las otras, ha sido la del control de responsabilidades (es decir, la evaluación sirve, sobre todo, para presentar datos y rendir cuentas de la gestión tanto al ejecutivo sobre todo, como al legislativo -es el caso especialmente del País Vasco y, en menor medida, de Cataluña y Madrid), aunque esto tampoco puede generalizarse.

¹¹ Se recuerda aquí la interesante propuesta de Weiss de lo que ella llama las “teorías del cambio”, y que incluyen la “teoría de programa”, distinguiéndola de la “teoría de la implementación del programa”. Mientras la “teoría de programa” se refiere a los mecanismos que median entre la entrega (y la recepción) del programa y el surgimiento de los efectos de interés, la “teoría de la implementación” es la teoría subyacente a lo que se requiere para traducir los objetivos en la entrega del servicio en curso o en las operaciones del programa. El mecanismo operativo de cambio de la “teoría de programas” no son las actividades del programa en sí (lo que se hace), sino la respuesta que dichas actividades generan. El de la “teoría de la implementación” es el de la entrega y reparto de los servicios y las actividades del programa. Es decir, mientras la primera teoría pone el énfasis en las respuestas de la gente a las actividades, la otra en cómo se llevan a cabo dichas actividades.

¹² La “comunidad de política” (*policy community*) estaría compuesta por todas las personas o grupos que afectan o se ven afectados por la política, o que tienen algún interés en la misma, incluidas las *femócratas*, otros actores gubernamentales, el movimiento de mujeres, las personas expertas en temas de igualdad, seminarios o institutos universitarios interesados en los estudios de mujer, etc.

Una de las razones que aquí se identifican como clave de la falta de claridad en el fin y las funciones de las evaluaciones estudiadas, es la ausencia de reconocimiento del carácter político que todo proceso de evaluación tiene. Por lo general, la evaluación de los planes de igualdad se ha considerado como un asunto eminentemente técnico, y esto es algo que inevitablemente hace que se desvirtúen los procesos de evaluación, haciéndolos en muchos casos residuales e inútiles para los fines que supuestamente se habrían propuesto en un principio.

En el primer capítulo de este trabajo, ya se exponía el evidente carácter político de la evaluación y la centralidad de dicho carácter para la práctica y la teoría de la evaluación. No sólo la evaluación parte de un contexto claramente político (en el que hay grupos y personas interactuando y con diferentes intereses y posiciones en el programa que se va a evaluar) y sólo tiene sentido una respuesta de la evaluación a dicho contexto, sino que tanto lo que se evalúa (programas o políticas) como la decisión de evaluar son claros productos de decisiones políticas. Además, la evaluación se hace para facilitar el proceso de toma de decisiones y el aprendizaje organizativo (ambos procesos claramente políticos), y toma una determinada postura respecto a los programas o políticas que está evaluando.

Este “reconocimiento político” de la evaluación tiene una serie de implicaciones, como son la necesaria credibilidad de los equipos evaluadores, la oportunidad en el tiempo de la evaluación y la necesidad de flexibilidad durante su proceso. En general, significa también reconocer que los intereses de los evaluadores y los de los gestores públicos y responsables políticos respecto a la evaluación son normalmente distintos y muchas veces entran en conflicto, que hay que entender el valor simbólico y de cambio de la información que ofrecen las evaluaciones, y que es necesario extender las evaluaciones más allá de los objetivos de unos planes o programas y comprender los procesos organizativos que subyacen a toda situación evaluativa. También implica que hay que tener en cuenta a todos los agentes críticos (*stakeholders*) de los programas -especialmente a los más vulnerables- y reconocer que existen diferentes necesidades de información en diferentes puntos del ciclo de una política. Por lo general, las experiencias analizadas no han tenido especial preocupación por estas cuestiones, sino que se han centrado más bien en los elementos técnicos de la misma. Y esta circunstancia se ha traducido en un mal enfoque de la mayoría de los procesos de evaluación, obviando una serie de circunstancias y elementos cruciales para conseguir evaluaciones realmente útiles.

2.4. Percepción de la evaluación como una función secundaria: el papel crucial de los organismos de igualdad en la evaluación de las políticas de igualdad de género

Como se ha señalado en el tercer capítulo de este trabajo, los organismos de igualdad, tanto a nivel central en primer lugar como autonómicos más adelante, han sido los

principales protagonistas e impulsores de las políticas de igualdad de género en España. Incluso cuando se adoptan los planes de igualdad como principal instrumento de política, aunque estos planes involucran a muchos otros actores gubernamentales, los organismos de igualdad siguen siendo los absolutos protagonistas y principales actores del conjunto de actuaciones que proponen, siendo los formuladores, impulsores y seguidores de las intervenciones en torno a la igualdad de género y siendo su capacidad de persuasión y convencimiento un elemento clave en el éxito de dichas políticas. Por lo tanto, los planes de igualdad son impulsados por los organismos de igualdad, que en el nivel central y en el autonómico (salvo excepciones)¹³ son más impulsores que ejecutores de políticas. Teniendo en cuenta esta realidad, y si una de sus principales funciones, además de la información y sensibilización y la de investigación en temas de igualdad, es, sobre todo, el impulso, coordinación, seguimiento y evaluación de los correspondientes planes, la evaluación debería ser una de las funciones más importantes en dichos organismos de igualdad. Sin embargo, la función de evaluación (con la excepción del Instituto Vasco de la Mujer *Emakunde* que han hecho una apuesta explícita por la evaluación y declaran que ésta es una de sus prioridades), es prácticamente residual en la mayoría de los organismos de igualdad que realizan las evaluaciones analizadas.

Aquí se insiste no sólo en el protagonismo crucial que deben tener los organismos de igualdad en los procesos de evaluación de las políticas y los planes (esto ha sido así hasta ahora ya que en todos los casos estudiados, el impulsor y en algunos casos también ejecutor de cada evaluación estudiada ha sido precisamente el organismo de igualdad correspondiente), sino en que *la función de evaluación debe ser claramente una de sus funciones principales*. De hecho, puede ser precisamente esta función de evaluación la que ayude a definir diferencialmente el papel que deben cumplir los organismos de igualdad en las políticas transversales de igualdad de género. Es decir, aunque los organismos de igualdad vayan perdiendo protagonismo en la propuesta, formulación y ejecución de actuaciones por la igualdad de género (esto sería deseable ya que indicaría que se va asumiendo la perspectiva de género en el resto de las políticas públicas), no tendrían por qué perderlo nunca respecto a su coordinación y, precisamente, al seguimiento y la evaluación de las mismas.

De esta manera, se señala la importancia que podría y debería tener un buen liderazgo de los organismos de igualdad respecto a la evaluación de las políticas de igualdad de género. Precisamente, el hecho de que sean políticas que se plantean como transversales y cuyo objetivo último es conseguir que el resto de las políticas públicas tengan perspectiva de género, podría poner en duda la necesidad o conveniencia de organismos específicos de igualdad. El necesario papel que pueden cumplir los organismos específicos de igualdad parece evidente cuando pensamos en las funciones de análisis e investigación sobre género y de impulso, coordinación, seguimiento y, sobre todo, evaluación de las políticas de

¹³ En algunas comunidades autónomas., especialmente las más pequeñas, se tiende a tomar el papel de ejecutor directo de actuaciones de promoción de la igualdad de género. Este es el caso, por ejemplo, de La Rioja. También en Extremadura, aunque no sea una Comunidad Autónoma tan pequeña, se ha dado esta característica, siendo la Dirección de la Mujer un organismo que ejecuta directamente las actuaciones; en este caso, además, en el momento en el que se realiza la investigación, no tenían un segundo plan de igualdad.

igualdad de género. Además, debido a la falta de cultura evaluativa existente en las administraciones públicas españolas, este liderazgo puede ser aún más importante. Más allá, la función de evaluación, en este caso, es extremadamente importante por el hecho de tratarse de políticas innovadoras que son especialmente susceptibles de revisión, aprendizaje, ajuste y mejora.

Sin embargo, este no ha sido el caso en las experiencias analizadas, salvo el caso de *Emakunde*, ya mencionado, en donde la función de evaluación de sus Planes de Acción Positiva ha tenido un papel prioritario. En la mayoría de los casos, sin embargo, los procesos de evaluación no han sido demasiado protagonistas, dentro de los organismos, ni en su concepción, ni en su ejecución y ni siquiera en su posterior utilización.

Relacionado con esto, nos gustaría señalar que, desde nuestro punto de vista, el organismo de igualdad central, el IM, y debido a su posición respecto al resto de los organismos de igualdad autonómicos, podría haber ejercido un mayor liderazgo en temas de evaluación. Precisamente en unos momentos en los que el IM ha tenido que ir redefiniendo sus funciones, tanto por la transversalidad de las políticas y el cambio de enfoque de promoción de la mujer a cuestiones de género, como por el propio desarrollo del Estado de las autonomías y las políticas de igualdad autonómicas. La cuestión de la evaluación de las políticas de igualdad podría ser una de las cuestiones clave en torno a la cual el IM ejerciera un liderazgo orientando, proponiendo, coordinando y reuniendo a las comunidades autónomas. Incluso hubiera podido recuperar parte de su liderazgo perdido en la generación de los terceros planes autonómicos, en los que, como se ha visto en el tercer capítulo, se da una mayor influencia internacional (del IV Programa Europeo y de la Plataforma para la Acción de Pekín) en detrimento de una menor influencia nacional que había sido mucho más evidente en los primeros y segundos planes. Sin embargo, el IM ha desaprovechado esta oportunidad, al menos en los procesos de evaluación estudiados. Tras muchas peticiones por parte de algunos organismos autonómicos, se realizó, en junio de 1998, un breve curso de dos días y medio, en la Universidad Internacional Menéndez Pelayo, auspiciado por el IM, y que no ha tenido ninguna continuidad posterior¹⁴.

Algunas comunidades autónomas, especialmente el Instituto Vasco de la Mujer (*Emakunde*), podrían ejercer este liderazgo (de hecho, en cierto modo *Emakunde* lo ejerce de forma implícita siendo una referencia para el resto, ya que es el organismo de igualdad que comenzó hace más tiempo a evaluar y que lo ha trabajado de forma más rigurosa y continua). Aunque tienen una posición institucional difícil para hacerlo: en realidad, el instrumento de cooperación inter-gubernamental más importante, las Conferencias Sectoriales, ofrece mucho más protagonismo al nivel central que al autonómico, siendo aquél el que convoca e imprime el ritmo y agenda de las mismas. Es una lástima, por lo tanto, que no existan más vías para la cooperación interinstitucional (y entre las propias comunidades autónomas), y que den lugar a un verdadero intercambio, reflexión y aprendizaje sobre la evaluación de los planes y las políticas. Lo que ha ocurrido hasta este

¹⁴ “Evaluación de las políticas de igualdad entre mujeres y hombres” (Santander, 29 junio-1 julio de 1998; directores: Rocío Fdez. Ballesteros y Mariano Álvaro). Al él se invitan a las comunidades autónomas de Castilla y León, Cataluña, Galicia y País Vasco para que expongan sus experiencias de evaluación.

momento es que, o bien no se han aprovechado las experiencias ya realizadas por otros organismos de igualdad, o bien se ha producido un cierto mimetismo plano sin demasiado cuestionamiento de unas experiencias con otras (por ejemplo, tratar diferenciadamente una evaluación cuantitativa y una cualitativa), al igual que ha ocurrido, en muchos casos, con la propia elaboración de los planes (Bustelo, 1999a: 381-382).

2.5. Visión parcial de la evaluación: una mala gestión y utilización de los diferentes tipos de evaluación

En general, esta segunda generación de evaluaciones, que para muchos organismos de igualdad representan sus primeras experiencias, adolece en cierto modo de una visión estrecha y parcial de la evaluación. Esto consiste, como en muchos casos en los que se carece de cultura y experiencia evaluativa, en realizar evaluaciones con una tendencia a ser parciales y finalistas.

Una buena forma de analizar esta visión parcial es repasando las tipologías básicas de evaluación, comprobando la utilización que de ellas se ha hecho en estas experiencias, que, en realidad, ha sido parcial, “desaprovechando” en cierto modo, la potencialidad de la evaluación. Para esto se utilizan tres de las cuatro tipologías mencionadas en el capítulo dos. Según el papel que cumple la evaluación, según el contenido, y según el momento. La cuarta, según el agente, se trata en el punto siguiente de forma separada.

En lo que se refiere al papel que cumple la evaluación, casi todas las evaluaciones se han concebido de manera *sumativa*. Esto es normal si se considera que las evaluaciones sumativas pretenden hacer recapitulación del programa a evaluar y valorarlo de forma global. Las formativas, sin embargo, pretenden mejorar el programa sobre la marcha. En la mayoría de los casos, las evaluaciones se han empezado tarde, por lo que una visión formativa era casi imposible. En cualquier caso, aunque se concibieran las evaluaciones exclusivamente desde un punto de vista sumativo, estas deberían concebirse desde un principio, para que se pudiera ir recogiendo información que, en un momento determinado, se analizaría y valoraría globalmente. Además, las evaluaciones sumativas, aunque no sirvan para mejorar el propio plan que se está evaluando, sí sirven para mejorar los siguientes, cumpliendo así, y de esta manera, una de las funciones clave de la evaluación. Y si realmente se consiguiera normalizar la función de evaluación y ésta se incorporara al quehacer cotidiano de los organismos de igualdad y restantes unidades administrativas implicadas, habría que ir también hacia una evaluación formativa, al menos de algunos de los aspectos del plan, que ayudara a ir mejorando sobre la marcha los propios planes que se están evaluando.

Respecto al momento, todas las evaluaciones son *post-facto*. Inclusive, como se verá más adelante, una de las enormes dificultades de los procesos de evaluación analizados es, precisamente, su dimensión temporal y los retrasos que se han producido, tanto en su planteamiento, como en el propio proceso, una vez empezado. Sin embargo, podría darse una evaluación *ex-ante* de los planes, que resultaría sumamente útil y por la que no habría

que hacer demasiados esfuerzos o invertir demasiados recursos. Se trataría de realizar una buena y sistemática evaluación *ex-ante*, una vez se elaborara el plan a evaluar o en el mismo proceso de elaboración y como parte del mismo, y antes de que éste fuera aprobado y puesto en marcha. Esto podría contribuir enormemente a reforzar los planes y los compromisos que con éstos adquieren los diferentes actores gubernamentales y no gubernamentales. Este tipo de evaluación no se ha dado en ninguno de los casos.

Finalmente, respecto al contenido, se puede apreciar que muy pocas experiencias realizan evaluación del *diseño o conceptualización* del programa. Sin embargo, resaltamos aquí la importancia de este tipo de evaluación, para tener en cuenta a los planes en su conjunto y reforzar su “teoría de programa”. Este tipo de evaluación es precisamente una de las que más se reforzaría, en caso de que se hiciera énfasis en una perspectiva de evaluación de políticas más que de planes y actuaciones concretas.

En lo que se refiere a la evaluación de *proceso*, se puede concluir que ésta se ha hecho mucho menos de lo que se pretendía o se ha querido creer. En realidad las evaluaciones llamadas de proceso o implementación han sido, en muchos casos, evaluaciones de *outputs*, limitándose a describir las actuaciones o actividades que se habían realizado -o no- como parte de los planes. Sin embargo, prácticamente no se han evaluado los mecanismos de funcionamiento a través de los cuales los planes se ponen en marcha: desde cómo se llega a la gente destinataria de las actuaciones hasta, y sobre todo, cómo llega el organismo de igualdad a las otras unidades administrativas que deben estar implicadas, cómo se coordinan, cómo funcionan los foros de coordinación y decisión en el plan, cómo se siguen las actividades, etc. Es decir, además de evaluar el funcionamiento de cada actividad, hay que hacerlo del plan en su conjunto, ya que se necesita de una visión integral y holística para valorar adecuadamente las políticas. Y para valorar bien las políticas no sólo hay que evaluar su diseño y sus resultados, sino también cómo éstas se ponen en marcha. Si esto resulta vital en cualquier política pública, en las políticas de igualdad de género las cuales se conciben como transversales, lo es aún más: una buena parte de su éxito o fracaso podrá explicarse por su compleja puesta en práctica o implementación, en la que están implicados tantos agentes críticos. La evaluación sistemática de este proceso de puesta en práctica resulta crucial para conocer y valorar bien las políticas públicas de igualdad de género.

Por último, la evaluación de *resultados* es la que más se ha desarrollado en su vertiente de evaluación de productos o *outputs*. Es decir, se han evaluado las actuaciones que han resultado de los planes y, en algunos casos, se ha hecho de forma bastante completa y exhaustiva, analizando también los agentes implicados, el tipo de personas destinatarias, el presupuesto invertido, el tiempo empleado y hasta el tipo de actividad del que se trataba. Y sin embargo se ha hecho realmente muy poca evaluación en su vertiente de efectos o *outcomes*, sobre todo si hablamos de los efectos netos de los planes. Es decir, algunos de los estudios (evaluaciones de los segundos planes central y madrileño y primer plan castellano y leonés) han pretendido que algunas de las variables medidas eran efectos provocados por los planes que supuestamente estaban evaluando, pero no han utilizado ninguna técnica o dispositivo que permitiera hacer una conexión directa entre lo que estaban midiendo y el plan, por lo que a duras penas siquiera se les podría llamar efectos. Ciertamente, se ha de señalar aquí, que la evaluación de efectos netos es muy difícil de

llevar a cabo: muchas veces los efectos que se pretenden conseguir con los planes de igualdad son muy diversos, a largo plazo y, en general, difíciles de medir (por ejemplo, cambio de actitudes en la sociedad). Además, también es difícil identificar o delimitar los destinatarios de dichas actuaciones. Sin embargo, no por su dificultad hay que dejar de intentarlo. Una de las cuestiones que aquí se proponen, y que podría ser muy interesante, sería la medición de efectos intermedios (efectos que, aunque no sean los finales, se sabe que tienen una alta probabilidad de acabar produciendo los efectos finales que se proponen. Por ejemplo, cambio de actitudes ante las futuras profesiones en los y las jóvenes de un Instituto que han asistido a un programa de orientación profesional no sexista -el efecto final sería una selección más igualitaria de profesiones-). Esto podría hacerse si se explicita la “teoría de programa” de los planes, y se pueden establecer efectos intermedios bien fundamentados.

Además de aprovechar mejor las oportunidades que nos brindan los diferentes tipos de evaluación, se trataría de utilizarlos conjunta y sinérgicamente, sobre todo en lo que se refiere a la tipología mencionada aquí en último lugar -según el contenido o fase-. Es decir, cuando se hace evaluación de la implementación o proceso de ejecución del plan y evaluación de resultados, se suelen hacer dos estudios separados con procesos diferenciados. Y esto hace que la evaluación pierda mucha de su posible potencia. Pongamos un hipotético ejemplo: si lo que se quiere evaluar es la coordinación gubernamental que ha habido en un determinado plan de igualdad, se podría realizar:

- una evaluación de diseño, que tendría que ver con la conceptualización de dicha coordinación, la existencia de estructuras de coordinación, si su diseño es adecuado y si estaban previstas y explicitadas en el plan

- una evaluación de proceso (o implementación), que tendría que ver con el funcionamiento real de dichas estructuras de coordinación (periodicidad, calendario, contenido, bajas y estabilidad de los componentes, etc.), porque se sabe que no es condición suficiente que estén diseñadas formalmente y que hay múltiples factores que pueden influir en dicha implementación

- una evaluación de *outputs* o productos, en la que se registren y evalúen las reuniones resultantes, las relaciones establecidas, los acuerdos tomados y cumplidos, los trabajos conjuntos realizados y las actuaciones ejecutadas como consecuencia de dicha coordinación, y

- una evaluación de *outcomes* o efectos, en la que se valore si las actuaciones realizadas por cada uno de los departamentos gubernamentales, como consecuencia de la coordinación propuesta por el plan, han sido realizadas con perspectiva de género o pueden contribuir a que las políticas públicas correspondientes la tengan.

Lógicamente, para hacer una evaluación completa de la coordinación, no tendría sentido que se hicieran estudios diferentes y desconectados entre sí; se trataría de integrar las informaciones y valoraciones de cada uno de los aspectos planteados (su diseño, funcionamiento, resultados y efectos conseguidos), dándole así una mayor potencia a la

evaluación global del plan. En definitiva, para tener una visión holística de los planes y las políticas de igualdad hay que evaluar tanto su diseño, como su proceso de puesta en práctica y sus resultados, analizando conjuntamente dichas fases o elementos de la intervención pública.

2.6. La optimización de agentes externos en la evaluación: las evaluaciones mixtas vs. las evaluaciones externas

En este trabajo se ha utilizado la clasificación de tipos de evaluación según el agente evaluador -evaluación interna, externa o mixta- para exponer los datos ya que ésta era la clasificación básica que implicaba procesos más distintos entre sí¹⁵. Según esta clasificación, dos evaluaciones eran consideradas internas (Aragón y La Rioja), seis externas (Andalucía, Canarias, Castilla-La Mancha, Castilla y León, Galicia y Madrid) y otras tres mixtas (IM, Cataluña y País Vasco). En realidad, ya se había señalado que las evaluaciones mixtas eran en realidad procesos internos en los que se había recurrido a colaboraciones externas en alguna parte del proceso de evaluación. En este punto se establecen algunas razones sobre cuál ha sido el tipo de evaluación más adecuado y por qué.

Los procesos internos puros son poco frecuentes y muy probablemente, cada vez lo sean menos aún, debido principalmente a que los procesos de evaluación requieren un esfuerzo que, normalmente, la plantilla de los organismos de igualdad no está en condiciones de asumir plenamente, además de la tendencia general, cada vez mayor, de las administraciones públicas a concertar servicios externos. De hecho, de las dos experiencias clasificadas como internas en este trabajo, La Rioja ha recurrido finalmente a colaboraciones externas no sólo en la parte que tenía prevista inicialmente, como era la evaluación final, sino también en las últimas evaluaciones anuales (la del año 1999 y el año 2000). Ya se justificó en su momento por qué se decidió seguir clasificando dicha experiencia como interna y teniendo en cuenta los datos de los primeros años de vigencia del plan que corresponden a una evaluación netamente interna. En el caso de la otra experiencia de evaluación interna -Aragón-, se trató de una experiencia de dimensiones pequeñas y que se realizó en un momento en el que había habido un fuerte cambio organizativo y con una idea muy concreta de obtener una información descriptiva simple que diera una idea de lo realizado hasta entonces y lo que quedaba por hacer.

También ya se ha señalado como las evaluaciones mixtas son, en realidad, procesos internos de evaluación en los que se recurre a colaboraciones externas en alguna parte del proceso de evaluación. La cuestión es distinguir, dentro de las evaluaciones mixtas, otros tipos. Y, en este caso, ya se vio como eran diferentes los casos de Cataluña y el País Vasco que, aunque internos, se pide colaboración externa desde el inicio, en la propia definición y concepción del proceso de evaluación. Sin embargo, en el caso del IM central, la ayuda externa es mucho más concreta y delimitada en el tiempo, ya que la empresa externa que

¹⁵ Véase el punto 4.3.3 en el capítulo 4.

colabora lo hace en el proceso de aplicación y análisis del cuestionario en la realización y posterior análisis de las entrevistas en profundidad y los grupos de discusión. Es decir, mientras en las dos primeras la colaboración externa se concibe más como un proceso de asesoramiento para la totalidad del proceso de evaluación, en la última se trata de una participación en unas tareas concretas sin que eso implique una visión global del proceso por parte del equipo externo.

Estableciendo una comparación entre las evaluaciones mixtas y las externas se ve que, en general y con la excepción antes mencionada, las primeras han llevado a procesos más adaptados a los organismos de igualdad, su contexto, su realidad y sus necesidades y parecen haber producido una mayor aprendizaje organizativo y una mayor implicación de los equipos internos en los procesos de evaluación. En definitiva, en los casos de Cataluña y el País Vasco se ha producido un proceso de aprendizaje sobre evaluación en el seno de los respectivos institutos de la mujer, más reciente en el caso de Cataluña y más continuado desde hace tiempo (ya se había realizado una evaluación anterior en términos parecidos) en el País Vasco. En ambos casos se han utilizado las colaboraciones externas para orientar, sugerir, ajustar y contrastar los procesos de evaluación. Los equipos externos han trabajado *con* y no *para* los organismos de igualdad. Esto ha producido un proceso de aprendizaje organizativo mucho mayor que en el resto de los casos.

Este tipo de aprendizaje que se ha producido en estos dos casos de evaluación mixta es lo que nos hace concluir que, desde este punto de vista, los procesos mixtos son preferibles a los externos. Es cierto que las evaluaciones externas tienen más probabilidades de ser coherentes desde un inicio, en el sentido de que pueden contar con especialistas en evaluación que ofrezcan una visión global. Sin embargo, este aprovechamiento de una especialización y autoridad de expertos en evaluación también puede darse en procesos mixtos, si así se concibe desde los organismos de igualdad. También las evaluaciones externas tienen, en principio, mayores probabilidades de ajustarse a un calendario y de hacerse en tiempo, dado que no están tan determinadas por las urgencias y los procesos internos de las instituciones, pero esto no ha sido así ya que los problemas de calendario se han dado tanto en los casos mixtos como en los externos. De hecho, uno de los procesos con problemas de calendario más graves ha sido la evaluación del II Plan Andaluz, concebida como externa desde un principio. Otra de las cuestiones clave de las evaluaciones externas, y que normalmente se atribuye como ventaja sobre las internas o mixtas, es la mayor credibilidad que aquellas tienen. Sin embargo, en los casos estudiados éste no parece haber sido un factor determinante. De hecho, la colaboración y el asesoramiento en los procesos mixtos funcionan y se mencionan como un cierto aval en este sentido.

Esta tendencia, sin embargo, no puede ser generalizable de forma absoluta. Creemos que lo que hace preferible unos procesos sobre otros es el trabajo conjunto de un equipo experto externo y un equipo interno que se implica en los procesos de evaluación. Así, podemos encontrar, entre los casos externos estudiados, procesos en los que se ha dado un mayor esfuerzo por implicar a los equipos internos, como es el caso de Galicia. Es decir, en realidad, el éxito en este sentido no depende tanto de que sean procesos externos o no, sino de cómo sean éstos concebidos. Salvo en el caso mencionado de Galicia y, en parte, en el

de Castilla-La Mancha, en el que se procuró la participación de todos los agentes implicados en el plan (dado que se hacía paralelamente al proceso de elaboración del III Plan), el resto de las experiencias han sido concebidas y llevadas a cabo como procesos totalmente externos -a veces pecando de ajenos- a la realidad concreta del plan y de su principal protagonista, el organismo de igualdad. Es decir, se concebía más como un estudio externo en el que sólo existía interacción con el organismo de igualdad para la recopilación de información y el establecimiento de fechas de entrega, y poco más. Por lo tanto, la tendencia que creemos que sí puede generalizarse es la de a mayor implicación interna y mejor coordinación de equipos externos e internos, mayor ajuste al contexto y a la realidad de los procesos de evaluación y mayor aprendizaje organizativo sobre evaluación en los respectivos organismos de igualdad.

2.7. La falta de claridad en el QUÉ. La dictadura de la metodología y las técnicas en los procesos de evaluación

Uno de los pasos cruciales en evaluación consiste en determinar qué es lo que se quiere evaluar, una vez se ha clarificado el por qué y el para qué queremos evaluarlo. En primer lugar, es importante ya que es lo que determina y guía el estudio, orientando la recopilación y el análisis de la información; en definitiva, se trata de la identificación y definición de las variables que se van a valorar. Pero además, en evaluación, se determinan también los criterios de valor, es decir, a través de la definición de los criterios estamos definiendo la base lógica con la que posteriormente vamos a emitir los juicios sobre el programa o política evaluada. En definitiva, los criterios de evaluación son los que van a definir la perspectiva de valoración a utilizar.

Por otro lado, la definición de los criterios de evaluación puede -y debe servir- para realizar un proceso de revisión del propio programa o política a evaluar, es decir para clarificar y reforzar los objetivos y para repasar o explicitar la “teoría de programas”, en otras palabras, las premisas, principios y mecanismos subyacentes al programa (por ejemplo, ¿qué es lo que hace pensar o por qué una campaña publicitaria sobre el reparto de las responsabilidades domésticas va a contribuir a un cambio de actitudes en la población?). Y esto en la doble vertiente apuntada por Weiss, lo que ella llama las “teorías de cambio” de un programa, que serían la combinación de la “teoría de programas” (¿cuáles son los mecanismos por los que la transmisión de valores a través de los medios de comunicación provoca cambios de actitud? y la “teoría de la implementación” del programa (¿qué secuencia de distribución debe darse o cómo debe hacerse una campaña publicitaria, en definitiva, qué es lo que provoca que una campaña publicitaria sensibilice a la sociedad?) (Weiss, 1998: 57-58).

Sin embargo, el análisis de las experiencias de evaluación metaevaluadas en este trabajo permite detectar, en general, una cierta confusión respecto a la definición de los criterios de evaluación. En parte esto tiene que ver con una cierta confusión respecto a la “teoría de programas” anteriormente mencionada. Esta “teoría de programas” hay que clarificarla y hacerla explícita, y esto requiere tiempo y reflexión previa al establecimiento

de los criterios de evaluación. En general, pensar sobre qué se quiere evaluar -criterios de evaluación- es un paso clave y que, sin embargo, se suele obviar o se suele dar por hecho, pareciendo que ya están preestablecidos unos criterios para la evaluación que son generales, preexistentes y válidos para cualquier situación, pero que además nadie explicita de una forma clara.

Una primera cuestión detectada, pues, es que parece que no ha habido una reflexión seria previa a los procesos de evaluación sobre realmente qué se quiere evaluar. Es más, en muchas ocasiones da la impresión de que el QUÉ evaluar (criterios de evaluación) se ha definido a través del CÓMO evaluarlo (metodología y técnicas utilizadas), dando prioridad a esto último. Un ejemplo ilustrativo de esta circunstancia es la evaluación del II Plan de Igualdad central (y significativo ya que ha marcado pautas para algunas comunidades autónomas) en la que se hace una distinción entre una “evaluación cuantitativa” y una “evaluación cualitativa”. Son dos estudios totalmente diferenciados y el aspecto clave que los diferencia, y por lo que se definen, no es su contenido sino por el tipo de técnicas de recopilación y análisis que se han utilizado en los mismos. El estudio cuantitativo analiza, a través de una encuesta, una serie de variables en la población general (usos del tiempo, valores, actitudes, estereotipos y calidad de vida antes y después del plan), y como se ha comentado en el capítulo 5 de este trabajo, sin ninguna conexión real con el plan supuestamente evaluado. El cualitativo recoge las opiniones sobre el plan de personas implicadas y colectivos destinatarios del mismo: por lo que se define es porque se lleva a cabo a través de entrevistas en profundidad y grupos de discusión (por eso lo denominan cualitativo) y no por qué tipo de opiniones (si son sobre el diseño, el desarrollo o los resultados, sobre el papel de los planes en la sociedad o sobre cómo ha afectado dicho plan a estas personas) ni por qué se ha recogido esa información y no otra. La falta de claridad sobre los criterios de evaluación queda más patente aún cuando se comprueba que el tercer estudio que completa la evaluación del II Plan central es denominado “evaluación de la implementación”. En este caso la categoría utilizada sí tiene que ver con el contenido o fase del plan que se evalúa y no con la metodología utilizada. En general, definir el tipo de evaluación por el tipo de instrumentos que se utilizan en la recopilación de información -cualitativos o cuantitativos- parece una visión equivocada de la evaluación de programas, especialmente si tenemos en cuenta que, tanto la teoría como la propia práctica de la evaluación aconsejan y abogan por la utilización conjunta y dialéctica de diferentes instrumentos, tanto cuantitativos como cualitativos. En el ejemplo que hemos escogido -evaluación del II Plan central- sí se han utilizado ambos tipos de técnicas, pero no se ha hecho conjuntamente, sino de forma independiente, lo que resulta un indicador claro de que el elemento organizador de la evaluación ha sido el tipo de instrumentos de recopilación y análisis de la información y no los criterios de evaluación. Y esto, en nuestro parecer, es confundir el propio sentido de la evaluación: los instrumentos deben estar al servicio de los criterios y no al contrario, es decir los instrumentos sirven para recopilar y analizar una información sistemática que va a ser la que nos va a permitir contestarnos a las preguntas o criterios de evaluación que se han formulado y valorar, entonces, el programa o política evaluada según dichos criterios.

El ejemplo escogido ilustra de forma evidente este factor que determina, desde su propia concepción, la experiencia de evaluación mencionada. Pero en el resto de las experiencias mencionadas también existe, en mayor o menor medida, esta tendencia. Parece que lo más importante son las técnicas o instrumentos para recoger la información y esto, en muchos de los casos, opaca pasos previos (definición de los criterios de evaluación) y posteriores (valoración, emisión de juicios y recomendaciones) de la evaluación. Creemos que esto está influido por factores anteriormente mencionados, especialmente la falta de cultura de evaluación y de experiencia y el hecho de que los profesionales que han intervenido, salvo en algunas excepciones, son más especialistas en metodología de investigación social que en evaluación de programas y políticas públicas.

Otra de las carencias detectadas y que es otro indicador de esta falta de claridad en qué se debe evaluar es, como ya se ha visto en el punto 9.2.5., que se tiende a separar los diferentes tipos de evaluación realizados -en caso de que se hagan diferentes tipos-, sin integrarlos en un necesario análisis y valoración común.

Por otro lado, uno de los criterios más utilizados en prácticamente todas las evaluaciones analizadas es el del “grado de cumplimiento” de las actuaciones del plan. Tal y como se concibe este criterio en la mayoría de las evaluaciones, este “grado de cumplimiento” se suele referir a la ejecución de cada una de las actuaciones propuestas en el plan. Es decir, se trata simplemente de comprobar si se han realizado o no cada una de las acciones o medidas propuestas. En muchos casos, ni siquiera se realiza una valoración global del plan, sino de cada una de las actuaciones. De esta manera, aunque algunas de las experiencias se denominan “evaluación de la implementación” (que sería una evaluación de proceso), en realidad se trata más de una simple evaluación de *outputs*, analizando el producto final -si se ha realizado o no- más que el proceso por el cual se llega al mismo. Esta tendencia hacia la evaluación de los resultados de los planes es también general; se tiende a pensar que lo que hay que evaluar son los productos, salidas o efectos de los planes, es decir lo externo, más que el diseño o los procesos, algo más interno. El hipotético ejemplo que antes hemos señalado en el punto 9.2.5. -evaluación de la coordinación, por otro lado, elemento clave en un plan o política transversal- se señala como hipotético porque no hemos encontrado nada parecido en ninguna de las experiencias analizadas, es decir no se han evaluado prácticamente los procesos de coordinación, a pesar de que las responsables políticas y técnicas coinciden en valorarlo como un elemento clave de las políticas de igualdad de género.

Sin embargo, y a pesar de todas las limitaciones que se han señalado, en algunas comunidades autónomas se han trabajado muy seriamente los criterios de evaluación y se ha avanzado mucho en la valoración y reflexión sobre el tipo de actuaciones que se estaban realizando. Dichas clasificaciones, que se han hecho sobre todo en el País Vasco, y también en Galicia y Cataluña aunque en menor medida, han ayudado, por un lado, a clarificar algunos criterios de evaluación (no sólo se trata de comprobar si se han ejecutado o no las

actuaciones, sino qué tipos de actuaciones se realizan, por parte de quién y en qué medida), a tener información que permita una cierta evaluación del diseño de los planes (qué se está haciendo y por qué) y a reflexionar, en parte, sobre cómo se traduce en la práctica la promoción de la igualdad, y quiénes hacen qué cosas y qué es lo que resulta más fácil o menos para los poderes públicos.

Una vez realizada la reflexión sobre qué se ha evaluado en las evaluaciones estudiadas, se propone aquí un ejercicio de análisis más global sobre QUÉ habría que evaluar de las políticas públicas de igualdad de género. Parece evidente que, en último extremo, se trataría de saber si dichas políticas han modificado la situación social de la mujer, o mejor dicho, las relaciones entre los géneros. Sin embargo, ya se ha visto lo difícil que resulta poder discernir lo que es el efecto neto de un plan o política, es decir lo que se ha modificado a consecuencia del plan y sólo del plan. Por otro lado, cuando se evalúa la actuación de los poderes públicos, no sólo se deben evaluar los resultados o efectos de dicha actuación, sino también qué es lo que hacen dichos poderes públicos (y si lo hacen bien y qué pueden hacer para hacerlo mejor) para abordar el problema público de la desigualdad de género, en este caso concreto (es decir, hacer también evaluación de cómo están diseñadas las políticas y los planes y cómo se implementan o ponen en práctica). En definitiva, se trata de evaluar muchos aspectos intermedios, antes de o además de intentar medir los efectos. Más allá, cuando se proponen objetivos últimos tan abstractos (modificar las relaciones de género y la situación de la mujer y, por lo tanto, también la de los hombres), es necesario establecer objetivos intermedios que normalmente vienen definidos por las estrategias planteadas para abordar el problema público. Así, se deben también definir efectos intermedios, fundamentados en la “teoría de programas” y, en general, en la teoría subyacente a las políticas públicas de igualdad de género.

En esta línea, ya se ha visto que una de las cuestiones teóricas que se ha desarrollado como principal estrategia y fundamento en las políticas de igualdad de género es el *mainstreaming* o transversalidad normalizada de la perspectiva de género. Por lo tanto, si la estrategia planteada es dicha transversalidad de la perspectiva de género y dado que dichas políticas, a través de los planes, son un conjunto de actuaciones coordinadas (es decir, no es un sector o tema de mujeres sino un tema transversal amplio que implica a todos los sectores - CAD, 1998: 38) ¿no sería el efecto final de estas políticas la perspectiva de género en el resto de las políticas públicas? Esto implicaría que, realmente, la forma de evaluar los efectos finales de los planes y políticas de igualdad sería la valoración del resto de las políticas públicas en el sentido de si están formuladas, ejecutadas y evaluadas con perspectiva de género. A este respecto, ha habido alguna experiencia interesante sobre la valoración del impacto de género en otras políticas (*Gender Impact Assessment*) en Holanda, Nueva Zelanda y Bélgica, aunque este tipo de instrumento se ha utilizado más como diagnóstico o evaluación “ex-ante” de las políticas más que como evaluación de las políticas de igualdad¹⁶. Evidentemente, esto también implica que la evaluación de las políticas de igualdad de género debería “cargar las tintas”, además de en la perspectiva de

¹⁶ Véanse Verloo (1998) y Meier (1998).

género de las otras políticas, en su diseño global y como instrumento de política, en su proceso de puesta en práctica, sobre todo en lo referente a los procesos de coordinación intergubernamentales e interinstitucionales y en cuestiones relacionadas con el desempeño organizativo de los planes y políticas, más que en la exclusiva evaluación de la ejecución de las diferentes acciones que se proponen.

2.8. Problemas en la recopilación y los flujos de información: sistemas, cultura de recopilación, estructuras de coordinación

Una de las grandes dificultades asociadas a los procesos de evaluación, y que se menciona reiteradamente por las responsables de los organismos de igualdad, es la dificultad de recopilar la información sistemática necesaria para realizar la evaluación. Al ser transversales los planes, una buena parte de la información que hay que recopilar para la evaluación no se encuentra en los propios organismos de igualdad, que es donde se origina y desde donde se dirige la evaluación, sino en otras unidades administrativas. Desde los organismos de igualdad se debe pedir esa información y el resto de los organismos debe responder y, además, devolver la información. Para esta ardua tarea, los organismos de igualdad necesitan de una buena dosis de autoridad y de persuasión, además de unos sistemas de información bien diseñados, sencillos y comprensibles y un buen protocolo de seguimiento de dicha información.

De hecho, lo que ha ocurrido es que normalmente se ha conseguido información más exhaustiva y completa de las acciones que dependían directamente de los organismos de igualdad y, otras veces, también de las que provenían de unidades administrativas en las que existía ya una cultura de recopilación de información y/o con las que había estrechas relaciones con el organismo de igualdad (y que normalmente han venido dadas por la existencia de profesionales con sensibilidad de género en las unidades en las que se tratara). Pero, en general, uno de los grandes escollos con los que se han encontrado los responsables de las evaluaciones ha sido la falta real de información y de datos sistemáticos. Así, en algunos organismos de igualdad se tomaba una postura más activa respecto a dicha recopilación para asegurarse alguna información: por ejemplo, en el Instituto Aragonés de la Mujer o el propio IM central, en lugar de enviar los registros a los diferentes departamentos o unidades administrativas, se “rellenaban” en el propio organismo de igualdad con la información que proporcionaban cada una de las unidades, a través de memorias (que muchas veces también resultaban difíciles de conseguir) o incluso telefónicamente. En otros casos, se han enviado los registros para que se rellenen en cada unidad administrativa, cumpliendo el organismo de igualdad un papel de seguimiento y orientación (aclaración de dudas, etc.) y, por supuesto, de motivación para dicha recopilación de información (por ejemplo, en *Emakunde*, el Instituto Vasco de la Mujer).

En algunos casos, la dificultad en la recopilación de información es, además un indicador de la falta de concreción o de la ambigüedad del plan. La información se pierde en el camino porque la acción se había desdibujado previamente. Es decir, en algunos casos no se ha podido recoger bien la información de lo que se había hecho correspondiente al

plan en tal Ministerio o Consejería, porque nadie sabía identificar muy bien qué se había hecho al respecto y/o porque las responsabilidades respecto a tales acciones estaban excesivamente diluidas o no identificadas. De hecho, en algún caso se han justificado muchas acciones “a posteriori”, simplemente contabilizando el número de mujeres participantes en una acción que en cualquier caso, existiera o no el plan, habría realizado una determinada unidad administrativa.

Sin embargo, en otros casos, la dificultad en la recopilación sistemática y unificada de información persiste, a pesar de que las acciones del plan se hayan llevado conscientemente a cabo. Y esto tiene que ver no sólo con la ausencia de una cultura de evaluación y, en general, de recopilación sistemática de información, sino también con la ausencia de sistemas de información adecuados y de estructuras de coordinación entre las diferentes unidades administrativas implicadas. Un diseño y un funcionamiento adecuado de estos sistemas de información y de estas estructuras de coordinación son factores clave para el éxito de una evaluación. Es más, diríamos que para que se dé una recopilación de información realmente completa y fiable se tienen que dar estos dos elementos mencionados: un mínimo sistema de información que organice la recopilación de información y una estructura de coordinación que lo sustente, alimente y mantenga.

En realidad, para que un *sistema de información* sea válido (que mida lo que quiere medir, es decir, que la información recogida sea pertinente al objetivo con el que se ha diseñado, en este caso la evaluación) y fiable (que lo mida bien, de forma completa y sin errores) tienen que darse una serie de condiciones. Un sistema de información debe ser:

- Claro y definido: las categorías, items, variables y/o indicadores deben estar bien definidos de forma que sean unívocos. Ya que distintos actores van a recopilar la información, todos deben entender lo mismo para que dicha recopilación sea uniforme, comparable, válida y fiable.
- Comprensible y lo más sencillo posible: debido a que se tiene que implicar a muchos actores en la recopilación y dado que normalmente van a ser profesionales que no van a estar dedicados en exclusiva a ello - y para lo que tampoco tengan probablemente mucho tiempo- , hay que intentar que el sistema no resulte complejo y además que no ocupe mucho tiempo su cumplimentación.
- Consensuado: cuanto más se implique a los actores que tienen que recopilar la información, desde la propia definición y concepción del sistema, mayores probabilidades de éxito tendrá. Con un sistema de información que es percibido como impuesto y ajeno es muy difícil que se identifiquen y se motiven los protagonistas de la recopilación de información.

Pero como se ha señalado antes, un buen sistema de información no basta. Hace falta también una *estructura de coordinación*. En el capítulo quinto de este trabajo (punto 5.2.3. referente a las estructuras institucionales para la evaluación) se describen detalladamente los tipos de estructuras de coordinación en torno a las once experiencias de evaluación metaevaluadas. En general, se ve que dichas estructuras son todavía muy débiles y se

establecen una serie de tendencias, como son la importancia de los organismos autónomos en los procesos de evaluación y sus propias estructuras (Consejos Rectores), las ventajas y desventajas de las estructuras creadas “ad-hoc” para el seguimiento y evaluación de los planes y de las estructuras no específicas, y la poca vinculación real -salvo alguna excepción- de los Consejos de la Mujer y de los Parlamentos y Asambleas Legislativas.

En cualquier caso, sí se puede extraer una conclusión general a este respecto, y es que es bueno que existan vínculos y compromisos, respecto a la evaluación, de:

- Las estructuras vinculadas a los planes y/o a los organismos de igualdad, y que quienes tienen responsabilidades respecto a las políticas de igualdad, las tengan también, aunque sea sólo formalmente, con su evaluación. Esto no obvia que puedan darse, y que esto sea muy positivo, estructuras “ad-hoc” para la evaluación, de carácter más operativo (y que siempre cuenten con el apoyo político necesario).
- Las estructuras vinculadas al poder legislativo: el rendimiento de responsabilidades al poder legislativo, por parte del ejecutivo, funciona como acicate y elemento provocador de la evaluación y la recopilación sistemática de datos. En una de las experiencias analizadas, esta tendencia se comprueba de forma contundente: el hecho de que en la misma ley de creación de *Emakunde* se establezca un rendimiento de su gestión anual al Parlamento Vasco (que se materializa en la presentación de una memoria anual al Parlamento por parte de su Directora) ha generado todo un proceso y una cultura de evaluación en torno a los Planes de Acción Positiva del País Vasco.
- Las estructuras formales de participación del movimiento de mujeres: esta es la gran asignatura pendiente de la evaluación de los planes de segunda generación. La participación del movimiento de mujeres en los procesos de evaluación ha sido prácticamente nula. Entre otras cosas, no se ha establecido prácticamente ninguna estructura de coordinación en la que participe dicho movimiento (si hay alguna, como son los casos canario y madrileño -la “Comisión para la Igualdad Canaria” y la “Comisión para el impulso y seguimiento del plan madrileño”-, la participación de dichas estructuras en la evaluación fue meramente formal y simbólica). Las estructuras formales de participación del movimiento de mujeres, como son los Consejos de la Mujer, no han tenido prácticamente ninguna participación en los procesos de evaluación de esta segunda generación.

Por lo tanto y en general, en esta segunda generación de evaluaciones han existido muchas dificultades vinculadas a las estructuras de coordinación y eso ha ido en detrimento de los procesos de evaluación y, en particular, de los de recopilación de información. Además, al no estar incorporada o normalizada la función de evaluación en los organismos específicos de igualdad, y salvo excepciones como en el País Vasco, las estructuras ya existentes (Consejos Rectores, estructuras no específicas de seguimiento de los planes como la “Comisión de Igualdad de la Mujer” de Canarias, Consejos de la Mujer, comisiones legislativas, etc.) que formalmente podían cumplir un papel importante en dichos procesos, cumplen una mera función simbólica y, en muchos casos, ni siquiera tienen participación alguna en los mismos.

Como se ha señalado en el capítulo anterior, las estructuras creadas “ad-hoc” para el seguimiento de los planes, salvo en el caso de Galicia en el que se crea la comisión específicamente para la evaluación del Plan y por consejo del equipo evaluador externo, son estructuras de tipo formal y que no se crean para evaluar el Plan, sino que poseen unas atribuciones mucho más generales de impulso, coordinación y seguimiento del Plan. Y aquí parece haber una tendencia: a mayor responsabilidad política de sus miembros (por ejemplo en Andalucía, los integrantes deben ser viceconsejeros), mayor forma de arrancar compromisos políticos transversales y mayor compromiso formal, pero menor operatividad en la gestión de los procesos de evaluación y de recopilación de la información. Entre otras cosas, para que una estructura de coordinación apoye adecuadamente un proceso de evaluación y la creación y mantenimiento de un sistema de información que la permita, necesita de reuniones periódicas y de un sistema de trabajo al que hay que dedicarle tiempo por parte de cada uno de sus miembros. Y esto es algo que normalmente no se ha dado, ya que son comisiones que se reúnen pocas veces al año y que normalmente tienen otras tareas y cometidos además de la evaluación. En el caso de Galicia, único caso en el que se crea una comisión específicamente para el proceso de evaluación, dicha comisión creada “ad-hoc” fue clave en el proceso de evaluación -fue la forma en la que se pudo gestionar la recopilación de la información- pero tampoco funcionó de forma óptima, entre otras cosas por falta de apoyo político y por los propios cambios internos durante el proceso de evaluación.

Por último, además de a los problemas asociados directamente a la falta de una cultura de evaluación, se deben añadir las dificultades inherentes a la complejidad de abordar el problema de la igualdad de género. En definitiva, si no se adquiere una profunda perspectiva de género, es decir, si no se asumen como propios los objetivos propuestos por los planes y las políticas de igualdad, es muy difícil que en el resto de las unidades administrativas -diferentes al organismo específico de igualdad- no se viva la recopilación de información que es necesaria para hacer evaluación, como algo que se está haciendo “para otros”. Esto añade mayor dificultad a conseguir un sistema de información fluido que ya, de por sí, es complejo por la cantidad de actores de los que necesita para que funcione. Por eso resulta clave la existencia paralela de una estructura de coordinación que implique, motive y comprometa a los diferentes actores en esa recopilación de información.

Por lo tanto, en el CÓMO de la evaluación, más que preocuparse directamente de los instrumentos y técnicas que se van a utilizar para recoger y analizar la información (que es lo que se ha visto ha sido el principal elemento organizador de las experiencias analizadas), hay que hacerlo previamente de las estructuras de coordinación que van permitir que se recopile información suficiente, sistemática y fiable. Y esto es especialmente importante, al menos, por dos razones:

- Por el carácter transversal y horizontal de las políticas: al ejecutarse las acciones por parte de diferentes entidades y unidades administrativas, la información siempre va a estar “desperdigada” y hay que diseñar unos sistemas que permitan centralizarla de forma continua, estable y homogénea.

- Porque en la evaluación de programas en general, además de la información que se recoja “ad-hoc” para la evaluación (utilizando normalmente fuentes primarias), resulta clave la información producida por la propia puesta en práctica y gestión de los programas a evaluar. Si de partida no está disponible una buena información de este tipo, va a ser muy difícil conocer mínimamente el programa que hay que evaluar y poder construir algo más de conocimiento sobre el mismo.

2.9. La importancia de los tiempos y el calendario de los procesos: la temporalidad como elemento clave de la evaluación

Hay dos cuestiones relacionadas con el tiempo y que resultan cruciales en cualquier proceso de evaluación, y en los que las políticas de igualdad de género no van a ser una excepción. En primer lugar, el **momento en el que se plantee la evaluación**, ya que ésta resulta una herramienta mucho más potente si se plantea de una forma continua. Esto no quiere decir que se deba estar evaluando todo el tiempo sino que la evaluación, para que se aproveche en todas sus posibilidades, debe estar planteada desde el inicio, en la misma fase de diseño y planificación de los planes. Y esto es algo de lo que han adolecido todas las experiencias analizadas, a excepción de la evaluación del País Vasco y La Rioja: se han planteado las evaluaciones hacia finales del período de vigencia del plan que se quería evaluar. Esta circunstancia, además de ser un claro indicador de que los planes no habían sido elaborados en términos de resultados esperados (y, por lo tanto, muchos de los objetivos y actuaciones estaban formulados vaga y ambiguamente), ha acarreado múltiples dificultades: o bien no se había recogido ninguna información previamente y si se había hecho la información recopilada no solía ser útil para la evaluación. De esta forma, una buena cantidad de información necesaria -que se hubiera podido recoger si se hubiera planteado la evaluación con antelación- resultaba imposible ya tenerla, simplemente por criterios temporales.

La segunda cuestión es la del **calendario de la evaluación** o, por decirlo más adecuadamente, la necesaria oportunidad en el tiempo o la temporalidad de los procesos de evaluación. Como se ha señalado en los capítulos iniciales de este trabajo, los criterios de evaluación surgen en un contexto, situación y momento determinado: si se tarda mucho tiempo en contestar las preguntas de evaluación, por muy bien contestadas que estén, es posible que el debate que las generó ya no sea el mismo, que se hayan quedado obsoletas y sean irrelevantes o que los destinatarios a los que se había dirigido la evaluación hayan desaparecido. Por lo tanto, la capacidad de respuesta de la evaluación está claramente ligada al calendario. La utilidad de la información que arroja una evaluación depende, en gran medida, de que ésta llegue a tiempo para la toma de decisiones sobre las mejoras, las posibles modificaciones y el rendimiento de cuentas. La entrega tardía de un informe de evaluación puede significar la práctica inutilidad del mismo. Para Chelimsky, además, cuanto más inestable e incierto políticamente sea el contexto en el que se encuentra el programa o política a evaluar, más acuciante es la necesidad de que la evaluación cumpla los criterios de temporalidad. Así, el incumplimiento de plazos en un proceso de evaluación suele dañar la reputación y credibilidad de los evaluadores a largo plazo; además del

incumplimiento de compromisos, los retrasos suelen entenderse como señal de incompetencia o de que hay algo que ocultar (Chelimsky, 1998: 41).

Las evaluaciones analizadas son muy heterogéneas respecto a los calendarios utilizados en los procesos de evaluación, yendo de los escasos tres o cuatro meses de la evaluación del II Plan castellano-manchego a los dos años largos de la del II Plan andaluz. De todas formas, se han dado problemas al respecto, habiendo retrasos sobre el calendario previsto en prácticamente todas las evaluaciones. En algunas, los retrasos no han sido tan significativos, pero en otras, dichos retrasos pueden poner en duda la utilidad global del estudio (como es el caso de Andalucía).

En cualquier caso, el tiempo ha sido una variable que ha condicionado enormemente los procesos de evaluación estudiados. Por lo tanto, el control de dicha variable (tanto del momento en el que se plantea la evaluación como de un calendario real ajustado) parece esencial para una evaluación de las políticas de igualdad.

2.10. El claro déficit de “elaboración práctica”

En el análisis realizado de las once experiencias metaevaluadas, se concluye que existe un claro déficit de lo que denominamos “traducción o elaboración práctica”, que se refiere a los dos últimos niveles de análisis mencionados en el capítulo 1 de este trabajo¹⁷. Además de recopilar información, analizarla e interpretarla, en evaluación también se debe valorar dicha información (enjuiciamiento sistemático) y, posteriormente, realizar sugerencias para la acción (establecimiento de recomendaciones). La denominación de “traducción o elaboración práctica” viene dada por el hecho de que son estos dos últimos pasos los que confieren el necesario carácter aplicado de la evaluación. En definitiva, la recopilación y análisis sistemático de información y datos se realiza para emitir juicios sobre el programa a evaluar y para sugerir cómo mejorarlo.

Respecto al enjuiciamiento sistemático, y como ya se ha mencionado, no se trata sólo de recopilar y analizar la información sino de valorarla a la luz de los criterios de evaluación previamente establecidos. Esto significa el carácter sistemático de dicho enjuiciamiento: en primer lugar, se hace con una perspectiva de valoración previamente establecida y que nace de una adecuada lectura del contexto y situación de cada programa a evaluar y de la teoría subyacente a cada programa; y, en segundo lugar, se hace de forma sistemática y basándose en información sistemáticamente recopilada y analizada.

Además, la evaluación, al ser una forma especialmente aplicada de indagación ya que se hace con la intención de mejorar y rendir cuentas sobre el programa o plan a evaluar, debe ser práctica y estar dirigida a la acción. En definitiva, las evaluaciones se hacen con la clara y directa intención de ser utilizadas, bien sea para tomar decisiones y modificar

¹⁷ Véase punto 1.2.1. del capítulo primero, en el que se exponen los cuatro niveles de análisis en evaluación, expuestos por Patton: resultados o hallazgos, interpretaciones, valoraciones y recomendaciones.

aspectos concretos del programa a evaluar, bien sea para contribuir al aprendizaje de las organizaciones que tienen como una de sus misiones la función de evaluar sus actuaciones públicas. Por lo tanto parece evidente que uno de los “productos” de los procesos de evaluación debe ser un listado de recomendaciones que “traduzca” en términos prácticos lo que se ha concluido a través de dicho proceso o estudio.

Además, estas dos cuestiones mencionadas -enjuiciamiento sistemático y establecimiento de recomendaciones- que son los dos componentes de esa “traducción o elaboración práctica” a la que nos referimos, son precisamente dos elementos específicos de la evaluación de programas y políticas y que la diferencian de otros tipos de análisis o indagación.

A pesar de la importancia específica de esta elaboración práctica, tanto en lo que se refiere al enjuiciamiento sistemático como al establecimiento de recomendaciones, prácticamente todas las experiencias analizadas son deficitarias al respecto, algunas de forma más evidente que otras. Aunque en el caso de las recomendaciones es unánimemente deficitario: ninguno de los informes de las once experiencias analizadas contiene un apartado específico de recomendaciones o sugerencias, a pesar de la importancia que a este elemento se le da en la teoría de la evaluación¹⁸.

Respecto a la valoración o enjuiciamiento, se debe prestar atención a los apartados de conclusiones, ya que en esta redacción es donde suelen verse o emitirse los juicios resultantes del proceso de evaluación. En este sentido, sí encontramos algunas diferencias: Cataluña, Galicia y el País Vasco son las que tienen un mayor grado de elaboración de dichas conclusiones. De las tres mencionadas, las conclusiones de la única experiencia externa -Galicia- son muy generales, aunque se realizan sobre la base de una síntesis de todos los resultados obtenidos. En el caso de Cataluña, se realiza una valoración por áreas del Plan, en las que además se incluyen algunas sugerencias (y que pertenecerían al nivel de análisis siguiente, el establecimiento de recomendaciones), pero no se realiza ninguna valoración global del Plan. Finalmente, en el País Vasco, cuya experiencia ofrece una mayor elaboración práctica, se extraen conclusiones parciales y globales y además también se establecen conclusiones sobre el propio proceso de evaluación y no sólo sobre el plan evaluado.

En el resto de las experiencias, la elaboración de conclusiones es muy pequeña o prácticamente nula. De hecho, ya se ha comentado como en una de las experiencias (la de la evaluación del II Plan de la Comunidad de Madrid) se menciona expresamente, en el informe de evaluación, que no se hace ninguna valoración para que sean los propios lectores quienes las hagan (*sic*). En el caso de Andalucía y Canarias se establecen algunas conclusiones al final de los informes, pero éstas son de carácter general y ocupan muy poco espacio en dichos informes. En la evaluación de Castilla-La Mancha existen una serie de conclusiones en cada uno de los tipos de evaluación pero ninguna de carácter general ni ningún apartado específico en el informe. Ni en la síntesis de resultados de la evaluación

¹⁸ Véase punto 1.3.3. del capítulo 1, especialmente las referencias realizadas a Torres, Preskill y Piontek (1996) y a Patton (1997).

del II Plan central ni en la de Castilla y León, al igual que ocurre en la de Madrid, se realiza ningún tipo de conclusión.

Desde nuestro punto de vista, existe pues, un clarísimo déficit de elaboración práctica, en algunos casos tan extremo, que esta circunstancia bastaría para poner en duda si realmente nos encontramos ante un proceso de evaluación o ante otro tipo de análisis o investigación: en cualquier caso, podrían ser valorados como procesos de evaluación claramente incompletos. En general, sí parece que puede concluirse que a los procesos de evaluación analizados se les podría haber “sacado más jugo” o, al menos, se hubiera podido profundizar más en los mismos de forma que su utilidad práctica fuera mayor.

Esta circunstancia del déficit de “elaboración práctica” es una de las que más influye, según el punto de vista que aquí se expone, en la falta de utilidad y de impacto de las evaluaciones realizadas. En un contexto en el que no existe cultura de evaluación, probablemente sea muy difícil que los potenciales destinatarios de dicha evaluación sean capaces de hacer dicha “traducción o elaboración” práctica por su cuenta, no por falta de capacidad, sino por falta de costumbre, método y, seguramente, también tiempo. Pretender que los responsables políticos y técnicos de un plan de igualdad -teniendo en cuenta todas las unidades gubernamentales implicadas- hagan por sí solos dicha “traducción” es algo, al menos, ingenuo. Esto se agrava más aún si se piensa que otros potenciales destinatarios de la evaluación de los planes deben ser el propio movimiento de mujeres, otros niveles gubernamentales y hasta la propia ciudadanía.

2.11. Los deficientes procesos de comunicación y difusión de la evaluación. La accesibilidad de los informes.

Una de las primeras cuestiones a tener en cuenta en los procesos de comunicación es qué se comunica, es decir, si se refleja en los informes finales y de una forma completa, los procesos de evaluación que se han llevado a cabo. Lógicamente, el primer paso en ese proceso de comunicación es que se haga explícito lo que va a transmitir y difundir. Y el resultado o producto tangible de un proceso de evaluación son los informes que se redactan, en los que se exponen los resultados de dicho proceso. En este sentido, los informes de evaluación no son, por lo general, completos, aunque existen diferencias entre las once experiencias, que se comentan más adelante.

Un segundo paso en los procesos de comunicación es la difusión que se hace de los informes, cuándo se hace y quiénes son los destinatarios. Tampoco parece haber habido una gran difusión de los informes de evaluación ni tampoco ha quedado claro quiénes eran los destinatarios de los mismos; más bien parece haber sido una cuestión más bien interna y restringida a los propios organismos de igualdad, salvo en el caso del País Vasco y en los casos en los que los informes -resúmenes de los mismos- han sido publicados.

Para que se dé un buen proceso de comunicación y difusión tienen que darse los dos pasos descritos: que el informe sea completo y que se difunda bien. Por lo general, todas las

experiencias analizadas adolecen de una cosa o de la otra. En el caso de las evaluaciones externas, la responsabilidad del primer aspecto (si los informes son completos) recae principalmente sobre los equipos externos. La cuestión de la difusión, sin embargo, es más responsabilidad del organismo de igualdad correspondiente.

En general, los informes tienden a ser incompletos, aunque varían enormemente desde los más completos y exhaustivos, tanto en los procedimientos y la metodología de evaluación como en las valoraciones y conclusiones, como es el caso vasco, a los más “parcos”, como es el caso madrileño, en el que se ofrecen directamente los resultados del estudio sin ofrecer ninguna valoración ni explicar nada sobre los procedimientos empleados en el mismo. Además de no desarrollar normalmente conclusiones generales, valoraciones y, especialmente, recomendaciones, tampoco, salvo en alguna excepción, son muy explícitos respecto a la metodología y al proceso seguido. De las experiencias internas mixtas, ya se ha dicho que el País Vasco es la que produce informes más completos (la única experiencia que ofrece información no sólo de los resultados sino también del propio proceso de evaluación) y que mejor difusión tiene. En el caso catalán el informe es bastante completo y, sin embargo, su difusión no parece haber sido muy alta. Justo al contrario que la evaluación del II Plan central, en la que se publica un resumen del estudio muy incompleto. De las experiencias externas, la andaluza y la canaria son las que parecen estar más equilibradas respecto a informes más o menos completos (en el caso canario las limitaciones reflejadas en el informe son las que se dieron en el proceso de evaluación, y en el caso andaluz, se hace un buen resumen aunque sea pequeño aunque no refleje toda la riqueza del proceso), y un mínimo proceso de difusión (publicación andaluza y jornadas canarias). En los casos restantes existe un desequilibrio entre los dos extremos: hay dos casos en los que los informes son más o menos completos -Castilla-La Mancha y Galicia- pero en los que el proceso de difusión posterior es prácticamente inexistente. En los otros dos (Castilla y León y Madrid), aunque el proceso de difusión es relativamente bueno, los informes que se difunden son muy incompletos. En general, además, no suelen haberse identificado previamente los destinatarios del informe, por lo que los procesos de difusión suelen ser incompletos.

Además de la difusión activa de los informes, también se ha querido valorar la accesibilidad a la información arrojada por la evaluación y existe una cierta ambigüedad sobre el carácter público de los mismos. En general, la accesibilidad a los informes de evaluación no es muy alta, exceptuando de nuevo al País Vasco, en la que no sólo no tuvimos ningún problema para obtenerlos, sino que éstos tienen un claro carácter público. También son accesibles los resúmenes de informes publicados por el IM, el Instituto Andaluz de la Mujer y la Dirección General de la Comunidad de Madrid: en estos casos hay que valorar si la información publicada en los resúmenes es o no completa. En el resto de los casos, sin embargo, la accesibilidad a los informes de evaluación no está clara. En algunos casos, los informes de evaluación quedan en una situación de “uso interno”¹⁹ y, a veces, el acceso a los mismos está determinado por la interpretación que de esta situación se

¹⁹ En algunos casos, se pueden consultar los informes en las oficinas del organismo de igualdad, pero no se pueden reproducir. En otros, como es el caso del IM, las copias están disponibles en la oficina correspondiente, pero no lo están en el Centro de Documentación.

hace por parte de las personas que tienen una responsabilidad técnica o incluso, política. De hecho, en más de una Comunidad Autónoma hemos tenido problemas para obtener copias a efectos de esta investigación (que iba avalada por una universidad, en la que se aseguraba que no se iban a utilizar los datos más que en un análisis general para la investigación, y en la que se decía que el objeto de la investigación eran los procesos de evaluación y no los resultados obtenidos en los mismos). Lo curioso es que no parece existir un criterio común respecto al posible acceso a los informes de evaluación. Como anécdota ilustradora de esta falta de criterios a la que nos referimos, mientras en una comunidad autónoma tuvimos que avalar nuestra petición de información con la garantía de la Universidad Complutense, en otra acabaron enviando la copia del informe a la dirección particular de la investigadora, después de muchos meses de espera y probablemente por temor a que la “vía oficial” no fuera la adecuada.

En definitiva, y con las excepciones mencionadas, en general, se hubieran podido comunicar y difundir mucho mejor los resultados de las evaluaciones. Resulta evidente que si los procesos de comunicación y difusión de los resultados de la evaluación han sido en cierto modo deficientes y la accesibilidad a la información arrojada por las evaluaciones no es muy alta, la utilidad de la evaluación se ve también mermada.

Finalmente, respecto a la accesibilidad de los informes de evaluación, creemos que sería imprescindible discutir y poner en común los criterios en lo que se refiere a dicha accesibilidad. Esto creemos que no es algo específico de la evaluación de las políticas de igualdad de género, sino que el acceso y el carácter público de los informes de evaluación realizados en el seno de o para las instituciones públicas españolas, no está siempre claro.

2.12. La escasez de recursos dedicados

Por último, se señala aquí un factor que afecta de forma muy directa a los procesos de evaluación. En términos generales, se han dedicado pocos recursos a la segunda generación de evaluaciones de los planes de igualdad. En el capítulo quinto se realiza un análisis de los recursos invertidos en cada una de las once experiencias de evaluación, y, teniendo en cuenta dicho análisis, se puede llamar la atención sobre las siguientes circunstancias:

- Ha sido muy difícil obtener información sobre los recursos empleados en la evaluación. Aunque se han facilitado los gastos directos imputables a la evaluación en los casos de estudios externos, ha sido prácticamente imposible obtener datos fiables sobre los gastos indirectos empleados y, en muchos casos, tampoco se tiene el dato de cuál ha sido el presupuesto global del plan objeto de la evaluación.

- La cantidad y tipo de recursos invertidos en las evaluaciones de los planes ofrece un panorama muy heterogéneo. Esto no podía ser de otro modo ya que son también muy heterogéneas entre sí las propias dimensiones de los organismos de igualdad autonómicos, y de éstos con el organismo central. Es decir que no son comparables entre sí el IM central con el Instituto Andaluz de la Mujer con la Dirección General de Bienestar Social de La

Rioja. Al ser muy diferentes entre sí en presupuestos y estructura, también deben serlo respecto a los recursos invertidos en la evaluación.

- Respecto a los datos obtenidos, los gastos directos, los recursos empleados, en términos generales, han sido muy bajos: de las seis evaluaciones externas, sólo hay dos (Andalucía y Madrid) que hayan superado algo más de los cinco millones de pesetas. Esto si lo consideramos en términos absolutos es poco, pero es mucho menos si lo hacemos en términos relativos. Como ya se indicaba en el capítulo quinto aunque los gastos directos puedan parecer altos, como es en el caso de Andalucía, no lo son en términos relativos. Teniendo en cuenta la evaluación andaluza, que es la “más cara”, (estudio para tres años y que incluía un estudio sobre la situación social de la mujer andaluza, y que costó cincuenta y seis millones aproximadamente), con mucha diferencia respecto a las demás, el monto empleado representa sólo el 0,07% del presupuesto total del Plan y el 0,97% del presupuesto del Instituto Andaluz de la Mujer en los tres años de vigencia del Plan.

- Para poder valorar en su justa medida los recursos dedicados a la función de evaluación, habría que tener en cuenta los gastos indirectos, sobre todo, los gastos imputables a los recursos humanos internos dedicados a los procesos de evaluación. Estas cifras no están calculadas ya que, como se comentaba en el capítulo quinto, las estimaciones que se puedan realizar “a posteriori” son muy poco fiables. De hecho, creemos que este cálculo es importante ya que probablemente los gastos indirectos en evaluación son mucho más altos que los directos. Entre otras cosas, incluso aunque haya evaluaciones externas, la implicación de los equipos internos en los procesos de evaluación es una clave para su ajuste y su éxito. Es decir, se tiene que dar un gasto interno importante. De esta manera, y si se valoraran adecuadamente los gastos indirectos de la evaluación, aparecería como evidente que las evaluaciones externas no son comparativamente tan caras respecto a las internas o mixtas, aunque sea el único gasto reflejado en los presupuestos como tal. Aún así, y sobre la base de estimaciones subjetivas, los gastos indirectos en la segunda generación de evaluaciones no han sido tampoco demasiado altos en general, aunque aquí, como en muchas otras dimensiones, la variabilidad es muy alta.

- Teniendo todo lo mencionado en cuenta, creemos que todavía se le podían, y se debían dedicar muchos más recursos a la evaluación de las políticas públicas de igualdad de género. Suponiendo, y no es una cifra muy alta, que se emplea en la evaluación el 5% del total de presupuesto del plan a evaluar y el 15% del presupuesto del organismo de igualdad, obtendríamos unas cifras muchísimo más altas de lo que se ha empleado en esta segunda generación de evaluaciones.

3. FACTORES RELACIONADOS CON EL IMPACTO DE LA EVALUACIÓN

Además de los factores que condicionan previamente la evaluación y de los factores que inciden directamente en dichos procesos, se identifican en este apartado dos factores que se relacionan directamente con el impacto de dichas evaluaciones. En primer lugar, se concluyen una serie de cuestiones relacionadas con la utilidad final y la pertinencia de las evaluaciones, constatando la relativa mayor utilidad de los procesos respecto a los resultados de la evaluación. En segundo lugar, se señalan las dificultades para realizar una síntesis de resultados de las evaluaciones de los planes de igualdad, pudiéndose aprovechar muy poca información de las evaluaciones realizadas para poder dar una visión general sobre las políticas públicas de igualdad de género en España. Sin embargo, en este último punto se concluye sobre la importancia de seguir realizando metaevaluaciones como un instrumento útil para la mejora y el desarrollo de la función de evaluación de programas y políticas públicas.

3.1. La relativa utilización, utilidad y pertinencia de las evaluaciones

La utilización de los procesos evaluativos ha sido analizada en este trabajo respecto a tres grandes puntos: 1) la pertinencia y utilidad de la información producida (y que analizaba la pertinencia de los contenidos del estudio, así como el grado de “elaboración práctica”), 2) la comunicación y difusión de los resultados (respecto a los contenidos del informe, sus destinatarios y su difusión) y, 3) el impacto de los resultados y los procesos de evaluación en las políticas y en los organismos de igualdad.

Los dos puntos primeros han sido analizados como factores inherentes a los procesos de evaluación ya que forman parte de dichos procesos, aunque justamente creemos que son los factores que condicionan de una forma más directa la utilidad y la utilización posterior de las evaluaciones.

En el primer punto, la **pertinencia de la información producida**, ya se ha destacado la falta de claridad en el QUE se evalúa, lo cual se refleja de forma directa en la *pertinencia de los contenidos del estudio*. El aspecto que se destacaba en el punto 9.2.7 de este capítulo es la falta de dedicación y, por lo tanto, la falta de claridad, a la definición previa de los criterios de evaluación que deben “dirigir” los procesos de evaluación. Esta falta de claridad se refleja en la mayoría de los casos, lógicamente, en la inadecuación o impertinencia de los contenidos del estudio. Sin embargo, en este punto ponemos énfasis en las dos cuestiones relacionadas con esta pertinencia de los contenidos de las evaluaciones y que se han destacado en la parte III de esta tesis doctoral: el alcance de la información y la globalidad de la misma.

Respecto a la globalidad de la información, ya se ha visto que en esta segunda generación de evaluaciones se han tendido, sobre todo, a evaluar los “productos” o *outputs* de los planes. En algunos casos se ha hecho una mera descripción de las actuaciones y, en otros, como en el País Vasco, Cataluña o Galicia, se ha hecho (o se ha intentado hacer) una evaluación más completa de dichos *outputs* que ha permitido en parte hacer una evaluación del diseño de los planes y obtener algunos datos interesantes para una evaluación de proceso. En otros casos, se han diseñado procesos de evaluación más globales, como en Andalucía y Castilla-La Mancha -aunque luego haya sido problemática la puesta en práctica de los procesos de evaluación- y, en otros, se ha pretendido hacer una evaluación de unos supuestos efectos de los planes (IM, Madrid y Castilla y León). En estos últimos casos, la no adecuación de los contenidos de la evaluación aparece como evidente: el objeto de estudio de dichas evaluaciones no es el plan a evaluar, sino la situación social de las mujeres supuestamente destinatarias, y sin ningún mecanismo que permita asignar a las variables estudiadas la condición de efectos netos de los planes. Inclusive, en algún caso los supuestos efectos que se pretenden medir son relativamente parciales (por ejemplo, cambios en valores, estereotipos y actitudes) si se tienen en cuenta el abanico de efectos que normalmente se pretenden con un plan integral de igualdad de género.

El alcance de la información recopilada y analizada en muchos casos no ha sido grande debido, sobre todo, a las dificultades encontradas para su recopilación, al no depender dicha recopilación del organismo de igualdad correspondiente sino de múltiples unidades administrativas, al tratarse de planes transversales. Sin embargo, en muchos de los casos, se recoge información de otras fuentes diferentes a las administrativas (de destinatarios de las medidas, por ejemplo) y a través de otras técnicas además de la recopilación documental (por ejemplo, a través de encuestas o cuestionarios). Por lo tanto, el alcance de la información recopilada, es decir su cantidad y calidad, sí es uno de los aspectos que se ha intentado cuidar en prácticamente todos los procesos de evaluación, aunque en muchos casos no se haya conseguido, por factores externos a los organismos de igualdad.

También como parte de la pertinencia y utilidad de la información producida, en la parte III de esta tesis, se analizaba el *grado de “elaboración práctica”*, es decir, el grado en que dicha información se traduce en conclusiones y recomendaciones para la práctica. El déficit de esta elaboración práctica es tan grande, evidente y generalizado, que aparece de forma clara como uno de los factores clave que, según nuestro punto de vista, determinan los procesos de evaluación estudiados (punto 9.2.10. de este capítulo). Por lo tanto, en este punto sólo reiteramos la importancia y necesidad de este “paso” para la posterior utilización de las evaluaciones. Aunque se recopile y analice una información adecuada, pertinente y de calidad, si no se completan los pasos siguientes en el análisis evaluativo (emisión de conclusiones, valoraciones y recomendaciones), a la luz de los criterios previamente establecidos, será muy difícil que los resultados de la evaluación sean útiles y utilizados.

Los procesos de **comunicación y difusión** de los resultados de las evaluaciones son otro paso más hacia la utilización de los mismos. También se analizan los, por lo general, deficientes procesos de comunicación como uno de los factores inherentes a los procesos de evaluación y que dificultan los mismos (punto 9.2.11. de este capítulo).

Por lo tanto, lo que queda por analizar relacionado con la utilización de las evaluaciones es el **impacto de los resultados y los procesos de evaluación en las políticas y en los organismos de igualdad**. Una primera cuestión es que resulta necesario diferenciar entre el impacto de los resultados de la evaluación y el impacto que pueden generar los propios procesos de evaluación. En general, el impacto de los resultados es más bajo que el impacto de los procesos en esta segunda generación de evaluaciones.

La utilidad de los RESULTADOS de la evaluación, por lo general, ha sido muy baja. De hecho, los resultados de los estudios realizados no parecen haber sido demasiado utilizados ni para la mejora de las políticas, ni para rendir responsabilidades ni para iluminar acciones futuras. Si acaso, podemos encontrar en algunos casos, en términos de Weiss (Weiss 1998: 23), un cierto uso conceptual y un cierto uso persuasivo, y mucho menos uso instrumental (utilizar los resultados de la evaluación para tomar decisiones sobre los planes). Un pequeño uso conceptual puede encontrarse en los casos de Cataluña y el País Vasco²⁰, ya que se hace una reflexión sobre el tipo de actuaciones que se realizan al amparo de los planes, realizando, en ambos casos, clasificaciones interesantes que permiten ser más consciente de en qué consisten las políticas de igualdad, a qué tipo de actuación se le da más peso, cuáles son más difíciles de llevar a cabo y por qué. También puede encontrarse un cierto uso que podríamos clasificar como persuasivo (se utiliza la evaluación para reclamar más apoyos para el plan), ya que las evaluaciones sirvieron, en algunos casos, para demostrar o sacar a la luz los diversos niveles de implicación real en los planes (en muchos casos muy baja) de las diferentes unidades administrativas que habían adquirido un compromiso con el mismo, a través de su aprobación por parte de los correspondientes gobiernos. Es decir, en cierto modo las evaluaciones han servido en parte para la función de responsabilidad o rendimiento de cuentas y para recabar más apoyos o, al menos, mayores compromisos.

El poco uso instrumental de los resultados de las evaluaciones (que se haya hecho algo, se haya tomado alguna acción o decisión, como consecuencia de dichos resultados) es lógico si se tiene en cuenta el conjunto de cuestiones que han sido analizadas en puntos anteriores: muchas veces se debe a un mal enfoque (pertinencia del contenido de la evaluación), unido, además, a un mal aprovechamiento de la información recopilada (poca elaboración práctica y pobre comunicación), que hace que la utilidad final para la toma de decisiones sobre los planes haya sido muy baja.

Sin embargo, el impacto o la potencial utilización de los PROCESOS de evaluación ha sido relativamente más alta, a pesar de que no existieran objetivos explícitos sobre ello. Por supuesto, este impacto hubiera podido ser mucho mayor y, de hecho, creemos que una de las mejores inversiones que los organismos de igualdad podrían hacer en un futuro sería empeñarse en sacarle el máximo rendimiento y aprendizaje a los propios procesos de evaluación, de forma que su función fuera consolidándose.

²⁰ En Galicia, aunque existían unas ciertas condiciones para este uso conceptual, dado el contenido de los resultados de la evaluación realizada, el estudio de evaluación fue posteriormente muy poco utilizado, en general, por parte de las responsables del organismo de igualdad que recibieron el informe de evaluación (y que habían cambiado respecto a quienes lo habían encargado).

Recordando la clasificación de usos de la evaluación que hacen Torres, Preskill y Piontek (1996: 2) en las acciones que se llevan a cabo como consecuencia de la evaluación (lo que hemos llamado uso instrumental), una mayor comprensión de un programa u organización (uso conceptual) y, finalmente, los efectos de los propios procesos de intervención, vemos que incluyen, entre estos últimos, una mayor comprensión y/o apreciación de la evaluación, el fortalecimiento o institucionalización de la función de evaluación, tanto en los propios planes, como en las instituciones que los impulsan -los organismos de igualdad-, además del aprendizaje sobre dicha función de evaluación, es decir, que por efecto de la evaluación, se hayan aprendido a evaluar mejor los planes de igualdad.

A pesar de que en esta segunda generación de evaluaciones, en la mayoría de los casos, falla la pertinencia de los procesos de evaluación que se han llevado a cabo, y también la comunicación formal sobre dichos procesos (en los informes de evaluación no se transmite prácticamente nada sobre los procesos y la metodología de la evaluación), en general, el impacto de los procesos estudiados es, en cualquier caso, grande, ya que son los primeros que se llevan a cabo y, por lo tanto, marcan pautas de actuación en un terreno nuevo como es la evaluación e introducen una cierta cultura y “saber hacer” sobre evaluación. Este tipo de impacto al que nos referimos se hubiera producido, fueran como hubieran sido dichos procesos, debido a la novedad, a la falta de referencias y al hecho de que han sido las primeras experiencias de evaluación para la mayoría de los organismos de igualdad.

La realización de estas evaluaciones supone que la evaluación se vaya consolidando como idea y como realidad, y que se aprecie algo más que antes cuando no existía ninguna referencia. Aunque todavía se considere como una función marginal, y no se le otorguen los suficientes recursos ni la centralidad necesaria, la función de evaluación se empieza a fortalecer. Y, sobre todo, a raíz de estas experiencias se da un aprendizaje sobre la función de evaluación, sobre todo en lo que se refiere a las formas y métodos de recopilar información. Por ejemplo, los sistemas de información en torno a los planes mejoran mucho a consecuencia de estas primeras evaluaciones. Este aprendizaje, sin embargo, es irregular según las experiencias, produciéndose mucho aprendizaje de los propios procesos en algunas (por ejemplo, en el País Vasco) y mucho menos en otras (por ejemplo, IM o la Comunidad de Madrid). El mejor indicador de que se produce este aprendizaje es la introducción de cambios o mejoras en los procesos de evaluación en el planteamiento de las evaluaciones de planes posteriores (o como en el caso del País Vasco, en un proceso continuado desde la evaluación del I PAPME). Creemos que este aprendizaje sobre evaluación se da en mayor medida cuanto mayor preocupación, atención y conciencia existe sobre la misma, es decir, cuanto mayor actitud de búsqueda y aprendizaje existe por parte del equipo técnico de los organismos de igualdad. Si se vive la evaluación como algo externo que realiza un equipo especializado, o se tienen ideas fijas sobre cómo ha de realizarse y no se cuestionan, menos probabilidad de aprendizaje existe.

En definitiva, las evaluaciones de la segunda generación han tenido muy poco impacto en lo que se refiere a sus resultados. En general, diríamos que son evaluaciones muy poco aprovechadas, tanto en su uso instrumental como en el conceptual y en el persuasivo. Si el primero es importante en el sentido de que las evaluaciones tienen que servir para tomar decisiones que mejoren los planes, los dos siguientes nos parecen cruciales en el caso de las políticas públicas de igualdad de género. El uso conceptual es especialmente importante en políticas públicas innovadoras, como son las que nos ocupan, ya que la evaluación representa una forma de ser más consciente de dichas políticas, comprender mejor su fundamentación y puesta en práctica y obtener nuevas perspectivas sobre dichas políticas y el marco institucional que las impulsa, es decir los organismos de igualdad y el resto de unidades administrativas que están transversalmente implicadas en las mismas. Por último, el uso persuasivo de la evaluación parece especialmente adecuado en estas políticas de carácter transversal y en las que hemos visto que la estrategia de persuasión “horizontal” resulta clave para conseguir el fin último a conseguir que sería una perspectiva de género normalizada en el resto de las instituciones y políticas públicas.

Aunque el impacto de los procesos de evaluación ha sido algo mayor, se debe, sobre todo, al hecho de haber sido las primeras experiencias de evaluación en la mayoría de los casos. También creemos que la evaluación podría tener mucho más impacto si se desarrollara la función de evaluación en los organismos de igualdad de una forma más consciente y otorgándole la centralidad que requiere.

3.2. La imposibilidad de una síntesis de evaluaciones de las políticas de igualdad centrales y autonómicas: análisis de los resultados

En capítulos anteriores se diferencian los conceptos de “metaevaluación” y de “síntesis de evaluaciones”. La metaevaluación es la evaluación de evaluaciones y la síntesis de evaluaciones, sin embargo, analiza los resultados de dichas evaluaciones. El foco de la metaevaluación se encuentra en los PROCESOS de evaluación y el de la síntesis en los RESULTADOS de dichas evaluaciones. Mientras el objeto de las metaevaluaciones son las evaluaciones, en las síntesis de resultados son las políticas evaluadas. Por lo tanto, el objetivo de una metaevaluación es la valoración de evaluaciones y la construcción de conocimiento sobre la función de evaluación en torno a determinadas políticas o programas. El objetivo de la síntesis de evaluaciones es la valoración de las políticas evaluadas y la construcción de conocimiento sobre las mismas.

Esta tesis doctoral se circunscribe al ámbito de la metaevaluación como evaluación de varios procesos de evaluación. Su objeto, por lo tanto, es la evaluación de las políticas de igualdad de género en España, no las políticas. Por lo tanto, en ningún momento se pretende extraer conclusiones sobre el estado de las políticas de igualdad de género, a través de sus evaluaciones. De hecho, no hemos estudiado los resultados obtenidos en cada una de las once experiencias metaevaluadas más que en su faceta de si éstos han sido reportados, informados, comunicados, difundidos y utilizados. Pero no nos hemos metido en el contenido de dichos resultados.

Sin embargo, creemos que la síntesis de evaluaciones cumple una función muy distinta a la de metaevaluación, pero absolutamente complementaria y conveniente. En los últimos años se ha visto como han aumentado las síntesis de investigaciones, los estudios de meta-análisis y también la síntesis de resultados de estudios de evaluación, especialmente en el mundo anglo-sajón (Cook, 1997: 31). De hecho, existe todo un movimiento de reivindicación de este tipo de estudios, ya que la racionalidad lo aconseja: existen ya una buena cantidad de investigaciones y evaluaciones que pueden ser re-analizadas y re-utilizadas, aprovechando las sinergias de unas y otras e incluso eliminando los posibles sesgos de los estudios individuales. En definitiva, esta tendencia significa reconocer que existe una buena cantidad de información ya generada que puede ser utilizada en un análisis secundario para dar nueva luz a los problemas a afrontar. En el caso de la síntesis de evaluaciones se propone analizar conjuntamente los resultados de varias evaluaciones de programas similares para poder llegar a conclusiones sobre la efectividad de los programas (Cook, 1997: 36-37).

La realización de una síntesis de evaluaciones permitiría tener una información de conjunto sobre las políticas y planes de igualdad de género, tanto de su diseño (sobre los propios planes y sobre cómo es el diseño de su puesta en práctica), como de los procesos de puesta en práctica (sobre todo en lo referido a la coordinación con el resto de departamentos y entidades, cómo se llega a los destinatarios, etc.), como de los resultados (tanto de productos o *outputs* -tipos de actuaciones, destinatarios, gastos, etc.) como de efectos o *outcomes* -qué se está consiguiendo y modificando por efecto de los planes). Pero para poder realizar esta síntesis de evaluaciones se tendrían que dar una serie de condiciones:

- Que se hayan realizado evaluaciones de planes de igualdad y se hayan reportado en un informe que esté disponible. Esta condición es difícil si no existe una cultura evaluativa y, sobre todo, tratándose de políticas innovadoras, con pocos precedentes tanto en los instrumentos de política como en su evaluación.

- Que las evaluaciones realizadas sean más o menos homogéneas, para que sus resultados puedan ser sintetizados y comparados entre sí. Las evaluaciones metaevaluadas son experiencias completamente heterogéneas entre sí y cuyos resultados difícilmente puede ser sintetizados en un cuerpo común de criterios, variables y datos.

Ninguna de estas condiciones se ha producido todavía en España, por lo que resulta muy poco probable la realización de una síntesis de resultados de evaluaciones de planes de igualdad de género.

Esta es otra de las razones por las que creemos que puede ser especialmente útil la realización de metaevaluaciones como la que aquí se presenta. En cierto modo, se trata del posible uso conceptual que de estas metaevaluaciones puede hacerse. Se trata de entender mejor la función de evaluación y el tipo de evaluaciones que pueden realizarse y es conveniente realizar. De esta manera, se puede contribuir a un desarrollo más pleno y racional de dicha función de evaluación, produciéndose más y mejores estudios de evaluación que posteriormente podrán ser sintetizados.

CONCLUSIONES

Sobre la evaluación en el marco del análisis de políticas públicas

1. Desde la perspectiva del análisis de políticas públicas, la evaluación ha sido tradicionalmente la fase menos estudiada del ciclo de vida de una política. A pesar de la poca atención prestada, la evaluación ha ido ampliando su objeto de estudio, abandonando paulatinamente la idea de restringirse a una fase final que valora exclusivamente resultados, y pasando a ser algo presente a lo largo de todo ciclo de vida de la política. Mientras el análisis de políticas públicas se ocupa, en general, de los problemas públicos y de cómo los poderes públicos los abordan, la evaluación de las políticas públicas es una parte de ese análisis general que se ocupa de valorar la actuación de los poderes públicos en su intento de solucionar problemas.
2. De los dos componentes del análisis de políticas públicas descritos por Lasswell -descriptivo y prescriptivo-, y que, aunque complementarios, han producido tradiciones científicas distintas -*policy making* y *policy analysis*, la perspectiva prescriptiva ha sido la más trabajada en el campo concreto de la evaluación. En definitiva, cuando se han realizado estudios de evaluación, se han hecho más bien evaluaciones con una intención prescriptiva relacionada con la mejora la política evaluada, más que análisis descriptivos de la fase de evaluación de las políticas. Esta tesis doctoral partió, precisamente, de dicha perspectiva descriptiva de procesos de evaluación para un caso concreto -las políticas públicas de igualdad de género en España- incluyendo también el componente prescriptivo, ya que la metaevaluación se realiza con la intención de prescribir mejoras en la función de evaluación que es objeto de estudio.
3. En el ámbito de las políticas públicas, se ha considerado tradicionalmente la evaluación como la fase final del ciclo de vida de una política. Esta característica ha provocado que se haya enfatizado un tipo de evaluación finalista, con carácter retrospectivo y exclusivamente centrada en los resultados de la política. Sin embargo, los estudios de evaluación han ido ampliando paulatinamente su objeto de estudio en los últimos años. En los últimos años, se ha reconocido que la evaluación puede y debe tener como objeto de estudio y valoración no sólo los resultados de una política sino también su diseño y conceptualización, así como sus procesos de ejecución y puesta en práctica.

Sobre el concepto de evaluación de políticas públicas

4. Desde aquí se propone una **definición de evaluación** que refleje tanto lo referido al **objeto** como lo referido a los **procedimientos**, además de incluir en el mismo concepto de evaluación también cuestiones referidas a las **funciones** de la misma. Respecto al **objeto**, la evaluación de políticas públicas se ocupa de cualquier intervención social que pretenda abordar un problema público, bien sea globalmente (políticas) o bien de forma particular a través de proyectos, programas, planes, medidas, servicios, materiales, etc. Como además, las intervenciones (políticas, programas, etc.) no son estáticas y tienen un ciclo de vida, el objeto puede ir desde su diseño hasta los impactos a largo plazo que puedan producir. En lo que se refiere a los **procedimientos**, como la evaluación pretende emitir juicios sobre el valor y mérito del objeto (programas, políticas, etc.) evaluado, además de la recopilación y el análisis sistemático de información, se realiza un enjuiciamiento también sistemático. Por último, las **funciones** deben ser incluidas en la definición ya que la evaluación no se hace en abstracto sino con la intención de mejorar los programas y políticas evaluadas, rendir cuentas y/o ejemplificar o ilustrar acciones futuras. Así, el sentido último de la evaluación es su **utilidad** para el proceso de intervención social. Esta utilidad debe ser concebida, además, en dos sentidos. Por una lado, la evaluación debe ser capaz de generar una información útil para el programa o política que esta evaluando, es decir, la información ha de ser pertinente al objeto evaluado y a su contexto determinado. Por otro, debe ser una información con carácter práctico, es decir, debe estar dirigida a la acción, proponiendo cursos de acción futura (que pueden ir desde modificaciones concretas a continuar o terminar el programa o política evaluada).

5. La primera de las **funciones de la evaluación**, la de **mejora o perfeccionamiento** (*improvement*) del programa o política evaluada, significa que la evaluación se concibe como un medio para la retroalimentación y el aprendizaje sobre la propia práctica. En este sentido, la evaluación resulta imprescindible como garante de la calidad de un servicio o programa. La segunda función básica de la evaluación, el **rendimiento de cuentas o responsabilidad** (*accountability*) significa que la evaluación debe servir como instrumento que permita, en los diferentes niveles de responsabilidad, rendir cuentas sobre la gestión y los resultados de un programa. Esto adquiere aún mayor sentido cuando se refiere a programas y políticas públicas, ya que el rendimiento de cuentas debe formar parte indiscutible del sistema democrático. La tercera función básica, dar luz o **iluminar acciones futuras** (*enlightenment*) hace referencia a que aunque las evaluaciones tienen que ver, esencialmente, con el programa concreto evaluado en un contexto, situación y momento determinado de forma que se contribuya a su mejora y se rindan cuentas sobre el mismo, las evaluaciones también ofrecen una información sistemática que contribuye al "acervo científico" no sólo de ese programa, servicio o política concreta sino también del abordaje general de determinados problemas públicos.

Teniendo en cuenta estas funciones, tradicionalmente, se ha ligado a la evaluación al **propósito** principal de **facilitar el proceso de toma de decisiones**

sobre el programa o política evaluada. Sin embargo, en los últimos años ha aumentado la conciencia de que las organizaciones y agentes en torno a un programa o política no aplican de forma directa e inmediata los resultados de una evaluación. Más frecuentemente aprenden de la evaluación y este aprendizaje afecta a las decisiones de forma sutil y gradual. Por esta razón, en los últimos tiempos se ha ido reconociendo el importante papel que puede tener la evaluación como **facilitadora del aprendizaje organizativo**. También hay que estar atentos a posibles **propósitos encubiertos** que pueden aparecer en determinadas situaciones evaluativas y que los evaluadores deben saber detectar y contrarrestar, como el aplazamiento (llevar a cabo una evaluación con el propósito de demorar una decisión), la evitación de responsabilidades (obligando a los oponentes a compartir responsabilidades por decisiones impopulares), la legitimación de una decisión ya tomada, la utilización de la evaluación como medio de hacer visible o darle publicidad al programa o política evaluada, o el ofrecimiento de una imagen determinada del programa (realizando estudios controlados políticamente).

6. Existen una serie de *características específicas* de la evaluación y que la diferencian de otros tipos de análisis o indagación, como son su marcado carácter político, su elemento valorativo y de enjuiciamiento sistemático y su fuerte carácter práctico y aplicado a la acción, el cual le otorga un papel crucial a la utilización y al uso posterior de la evaluación.

- 6.1. En los últimos años de desarrollo de la disciplina de la evaluación de programas y políticas, se ha ido reconociendo, de forma paulatina y creciente, el **carácter político de la evaluación** por parte de la comunidad científica. Al fin y al cabo, las políticas y programas que se evalúan son resultado de decisiones políticas, la evaluación pretende alimentar el proceso de toma de decisiones y, además en sí misma adopta una postura política ya que, por su misma naturaleza, expone implícitamente la necesidad y la legitimidad de los programas evaluados. La consideración de la política como algo central en la evaluación tiene implicaciones especialmente importantes como la necesidad de flexibilidad, temporalidad y credibilidad de los procesos de evaluación, así como de que la evaluación se entienda desde una comprensión profunda de los procesos políticos y de las políticas públicas, más allá de una posición técnica estrictamente metodológica.

- 6.2. El reconocimiento del **elemento valorativo** de la evaluación implica su consideración como elemento sistemático en el proceso de evaluación. Esto significa que no sólo hay que procurar que la información que se recopile sea de calidad sino que se proporcione y clarifique una base lógica que justifique las perspectivas de valoración utilizadas para interpretar y valorar los resultados. Por esta razón es tan importante que se definan bien los criterios de evaluación (qué se va a evaluar): por un lado, dichos criterios establecen las variables que van a ser contempladas en el estudio, pero además, por otro, son criterios de valor, es decir, llevan implícitas las bases de valoración con las que posteriormente se emitirán los juicios. El carácter sistemático del enjuiciamiento se da tanto por esa explicitación consciente y previa como por una recopilación y análisis sistemático de la información en la que posteriormente se basan los juicios.

6.3. Por la propia naturaleza de la evaluación, ésta debe contar con un último nivel de análisis y exposición de los resultados que consiste en la traducción de los resultados de la evaluación en **recomendaciones**. Estas recomendaciones son sugerencias sobre posibles líneas de acción a seguir respecto al programa o política evaluada, cambios propuestos en el programa o cosas a mantener en el mismo o consejos para la mejora dirigidos a los patrocinadores, responsables y/o profesionales del programa. Evidentemente, estas recomendaciones deben seguir la secuencia lógica y sistemática de los niveles de análisis anteriores (recopilación, análisis o interpretación y enjuiciamiento).

6.4. La **preocupación por la utilización y el uso** de la evaluación hace que su éxito no sólo se mida por el grado en que ésta contribuye al aumento o mejora del conocimiento sino, sobre todo, por el grado en el que ha resultado valiosa para cubrir las necesidades específicas de información de un cliente, destinatario o usuario particular. Esto hace que, necesariamente, existan diferentes necesidades de información y la evaluación pueda tener distintos usos (instrumental, conceptual, persuasivo, etc.).

7. La necesidad de caracterizar específicamente a la evaluación, nos hace poner especial énfasis en la **diferenciación entre investigación y evaluación** en el marco de las políticas públicas. En nuestra opinión, es necesario distinguir la evaluación de la investigación para otorgar a aquella un carácter e identidad propia. Las diferencias se pueden establecer según una serie de dimensiones. Según su *naturaleza*, siendo la investigación un modo de indagación más académico y la evaluación uno más práctico. La *finalidad* principal de la investigación es la construcción de conocimiento mientras que las de la evaluación son la mejora las políticas y el rendimiento de responsabilidades sobre las mismas. Esto implica que la evaluación tiene sentido mientras sea útil para dichas finalidades y para esto debe tener capacidad de respuesta y una temporalidad ajustada. La evaluación es dependiente del *contexto* y tiene por objetivo la respuesta a dicho contexto específico, que además es marcadamente político. Además, la evaluación siempre ocurre en un escenario de acción, es decir lo prioritario es el programa o política evaluada y no su evaluación. También existen diferencias según el *objeto*, siendo el de la investigación los problemas sociales a solucionar y el de la evaluación la manera en que los poderes públicos abordan dichos problemas y si son efectivos o no en ello. Mientras que los *destinatarios* primarios de la investigación son grupos abstractos de la comunidad científica, los de la evaluación son grupos y personas concretas, agentes críticos (*stakeholders*) en torno al programa o política evaluada. Finalmente, respecto a los *agentes*, los investigadores y evaluadores requieren de capacidades de alguna forma diferenciadas: la capacidad investigadora requiere tener conocimientos básicos sobre metodología y exhaustivos del sector sobre el que se investiga. Los evaluadores necesitan de conocimientos metodológicos amplios aunque el conocimiento sobre el sector sobre el que se evalúa no debe ser tan exhaustivo y puede irse desarrollando a lo largo del proceso. Sin embargo, el evaluador requiere de otras habilidades que le son necesarias como son el análisis de políticas públicas y el conocimiento de los procesos políticos, el conocimiento y la capacidad de trato con personas y grupos, así como habilidades de comunicación.

Sobre la historia de la evaluación y los diferentes enfoques o modelos propuestos

8. El estudio de los diferentes **enfoques o modelos de evaluación** propuestos a lo largo de la historia de la evaluación, ofrece una idea no sólo de las distintas propuestas de los autores clave, sino también de la propia evolución de la evaluación como actividad científica diferenciada. Realizando una breve **historia** de la evaluación y estudiando los factores que inciden en su evolución, se entienden mejor las respuestas que los especialistas han ido dando, proponiendo diferentes enfoques o modelos para explicar y practicar la evaluación de programas.
9. Uno de los factores que ha influido en el desarrollo de la evaluación como ámbito académico y profesional ha sido el **grado de institucionalización y consolidación profesional** de la misma, que ha sido diferente en cada país. La evaluación se produce principalmente en los países en los que existe un desarrollo importante de políticas de bienestar social. En España dicho grado de institucionalización es muy bajo (aunque ha ido aumentando paulatina y tímidamente en los últimos años), ya que todavía no existe cultura de evaluación ni está normalizada su función en el sistema político-administrativo.
10. El segundo factor que influye en la evolución de la evaluación como actividad científica diferenciada y en la propuesta de modelos y enfoques es la propia **evolución en el estudio y la concepción de los problemas sociales**. La confianza reinante a principios de los años sesenta en la capacidad del gobierno para afrontar los problemas públicos, se torna en una visión más pesimista a partir de los años setenta, cuando los politólogos comenzaron a ser conscientes de que los problemas públicos rebasaban por mucho las políticas públicas. Paralelamente, se empieza a ser consciente de que las grandes evaluaciones de los programas sociales, bajo un paradigma estrictamente positivista-experimentalista no resultaban útiles para la mejora de las políticas reales. ¿Qué había que evaluar pues? Esta pérdida general de la ingenuidad y la fe ciega en la aplicabilidad directa de la ciencia, así como el reconocimiento de la complejidad de los procesos políticos que intervienen en una política pública son también factores importantes para entender la evolución y el desarrollo de las propuestas teórico-prácticas en el campo de la evaluación.
11. Para la exposición de los **modelos o enfoques** seguimos la clasificación propuesta por Stufflebeam (Stufflebeam y Shinkfield, 1987; Stufflebeam, 2001). Las “cuasi-evaluaciones” se determinan por su visión parcial de los programas y políticas y por la ausencia del elemento valorativo y son el enfoque por objetivos y los enfoques experimentales. Entre las “evaluaciones verdaderas” se encuentran el enfoque orientado a la toma de decisiones (Cronbach, Weiss, Stufflebeam), el centrado en el cliente (Stake), el centrado en el consumidor (Scriven), el orientado a la utilización (Patton) y, finalmente, algunos enfoques participativos (evaluación de cuarta generación -Guba y Lincoln-, evaluación para el “empoderamiento” -Fetterman- y evaluación pluralista -Monnier-). Para cada enfoque se presentaron, de forma escueta, sus presupuestos básicos, sus contribuciones más representativas al campo

de la evaluación, así como algunas de sus fortalezas y debilidades para la práctica evaluativa.

Sobre los tipos de evaluación, los actores y agentes y sus implicaciones éticas, en una perspectiva histórica

12. Respecto a los **tipos de evaluación**, seleccionamos las cuatro tipologías básicas que nosotros consideramos clave para distinguir la naturaleza de la evaluación. Estas tipologías básicas se construyen a partir de cuatro dimensiones. Según *el papel que cumple la evaluación*, distinguimos entre evaluación **formativa** -con la intención de ir modificando sobre la marcha el programa evaluado- y evaluación **sumativa** -con la intención de hacer una recapitulación y valoración final al cabo de un período de tiempo-. Según *el contenido o fase* de la política o programa, se puede realizar una evaluación **del diseño** -cuyo objeto es el diseño o conceptualización del programa-, **de proceso** -que se centra en el proceso de puesta en práctica o ejecución, y las cuestiones de gestión y funcionamiento de los programas-, o **de resultados** -valora tanto los *outputs* o productos del programa, su realización, cumplimiento y calidad, como los *outcomes*, efectos o impactos del programa tanto sobre sus destinatarios o beneficiarios como a la comunidad en general-. Según *el agente evaluador* distinguimos entre evaluación **interna** -la cual se lleva a cabo por parte de los responsables o profesionales del programa o por una unidad interna de la entidad que financia, planifica y/o ejecuta el programa, siendo en realidad una (auto)evaluación-, y evaluación **externa** -se realiza por parte de alguien ajeno al programa, y que no tiene nada que ver ni con su financiación, ni con su planificación ni con su ejecución-. Finalmente, según *el momento en el que se evalúa* existe la evaluación **ex-ante** -que se realiza antes de que el programa se ponga en marcha- y la evaluación **ex-postfacto** -que se realiza una vez el programa se ha puesto en marcha o ha terminado-.
13. Una de las características de los contextos evaluativos es la presencia de diferentes personas y colectivos que tienen algo que ver con el programa o política que se pretende evaluar. Aunque el **cliente** o comanditario de la evaluación, es decir, quien encarga la evaluación y decide que hay que realizarla porque tiene unas determinadas necesidades de información, sea sólo uno, los **destinatarios** de la evaluación, es decir a quién va dirigida y quiénes son los posibles usuarios de la evaluación, son muchos más. Esta realidad significa que existe un elenco amplio de potenciales **agentes críticos (stakeholders)** en la evaluación. Todos los agentes críticos en un programa, por el hecho de serlo, serían susceptibles de tener interés en su evaluación y, por lo tanto, ser destinatarios y posibles usuarios de los procesos evaluativos. Se pueden establecer cuatro grandes categorías: 1) Patrocinadores, formuladores de políticas y responsables políticos; 2) Directores de los programas, gestores públicos y profesionales en general; 3) Destinatarios directos e indirectos (por ejemplo, padres de niños en programas educativos), usuarios potenciales del servicio, programa o política evaluada y el público o

comunidad en general; 4) Otros: profesionales y responsables políticos de programas similares, expertos del sector o política que se evalúa, científicos sociales en general. La elección de unos u otros agentes críticos como destinatarios principales de la evaluación depende del enfoque que se adopte. En cualquier caso, se cree conveniente insistir en la pluralidad de actores potenciales de un proceso de evaluación y en su necesaria participación e inclusión en el mismo ya que a mayor participación de diferentes actores en la evaluación, mayor probabilidad de uso de la misma. Por otro lado, la implicación de los principales actores es condición para un cierto pluralismo y para la integración de todas las informaciones, interpretaciones y juicios pronunciados en el marco del proceso evaluador.

14. El **papel del evaluador o evaluadora** varía sustancialmente dependiendo de la orientación, el tipo de evaluación a realizar y el tipo de evaluador del que se trate (evaluador en el *staff* u organigrama de una organización, organizaciones externas especializadas en estudios de evaluación -empresas consultoras privadas, departamentos o institutos universitarios de investigación, expertos o consultores individuales-). En cualquier caso, caben dos posibilidades que se podrían situar en los extremos de un continuo:

- a. El evaluador distanciado o despegado: este es el papel que se le ha otorgado tradicionalmente al evaluador. Aquí el evaluador procura situarse por encima de la situación y mantener la máxima autonomía y objetividad posible. Este evaluador busca los hechos y pone toda su confianza en la metodología.
- b. El evaluador comprometido o apegado: el evaluador trata de responder de forma sensible e interactiva a las preocupaciones e intereses de los grupos de agentes críticos. En este caso, el evaluador interactúa de forma cotidiana durante el proceso de evaluación, emplea un tiempo considerable en el lugar en el que se ejecuta el programa a evaluar y se implica con el grupo de profesionales.

El segundo tipo de papel se ha reforzado mucho con los últimos enfoques en evaluación y a raíz del fracaso de muchas experiencias de evaluación que defendían a ultranza un papel absolutamente distanciado para el evaluador en aras de la objetividad.

15. En los últimos años se le ha prestado mucha atención a las **cuestiones éticas** relacionadas con la evaluación. Esta preocupación es paralela al reconocimiento del carácter político de la evaluación y, por lo tanto, al poder que ésta puede tener y a la necesidad de que el evaluador interactúe en el contexto en el que debe intervenir. Muchos de los más recientes manuales en evaluación de programas y políticas contienen un capítulo específico en cuestiones éticas, y se han elaborado algunos códigos deontológicos (*Joint Committee* y asociaciones americana, canadiense y australasiática). Estos códigos, además de ser, a la vez, un indicador y un elemento generador o provocador de una mayor consolidación institucional y profesional de la evaluación y fomentar un sentido de pertenencia e identificación entre los profesionales, acotan y regulan la práctica profesional. Pero el aspecto más característico de estos códigos deontológicos es que permiten la introducción y desarrollo de los temas éticos, que resultan clave en la práctica de la evaluación. En

definitiva, se trata de orientar ante los conflictos y dilemas éticos que se van a producir, inevitablemente, en cualquier situación evaluativa.

Sobre las políticas de igualdad de género en España

16. Las **políticas públicas de igualdad de género se definen**, en este trabajo, como el conjunto de intenciones y decisiones, objetivos y medidas adoptadas por los poderes públicos en torno a la promoción de la situación de la mujer y de la igualdad de género entre mujeres y hombres. Así, se incluyen todas las acciones impulsadas por los poderes públicos y que tienen por objetivo conseguir una mayor igualdad de oportunidades entre mujeres y hombres, incidiendo sobre las condiciones socio-económicas y culturales que impiden y obstaculizan dicha igualdad. Su origen se encuentra en la idea o el reconocimiento de que el Estado o los poderes públicos pueden y deben hacer algo para paliar las desigualdades de género y la discriminación de la mujer. Aunque las **relaciones de las mujeres con el Estado** han sido casi siempre difíciles y, de alguna forma, contradictorias, el impulso esencial para que el tema de la desigualdad de género se convirtiera en una cuestión pública, entrara en la agenda de los poderes públicos y se comenzaran a formular medidas al respecto, se debe al **movimiento feminista** de segunda ola de la década de los años sesenta y la de los setenta.
17. El desarrollo de las políticas públicas de igualdad en España está estrechamente conectado, al igual que en otros países occidentales, al desarrollo del llamado **feminismo institucional**, que se generaliza a partir de la década de los años setenta y la de los ochenta en la mayoría de los países occidentales, y que consiste en la institucionalización de organismos públicos que promueven una agenda de igualdad de la mujer. También se utiliza el término **“femócratas”** o “feministas de Estado” para designar a las personas que dirigen y trabajan en dichos organismos de igualdad. El feminismo institucional, que nace como una respuesta a la presión de los movimientos de mujeres y a la buena acogida que tienen estas demandas en los organismos internacionales, ha sido y sigue siendo, en la mayoría de los casos, el principal protagonista e impulsor de las políticas de igualdad de género.
18. En España es especialmente importante reflejar en los estudios de las políticas públicas de igualdad de género el **nivel autonómico**, además del central, no sólo por la escasez de estudios específicos sobre las políticas de igualdad en dicho nivel gubernamental, sino por el crucial desarrollo del Estado de las autonomías en los últimos quince años y la importancia que las políticas sociales han adquirido en las comunidades autónomas.
19. Las **políticas de igualdad de género en España** han tenido un rápido crecimiento desde la creación del Instituto de la Mujer (IM) central, a finales del año 1983. El rápido desarrollo del **feminismo institucional en España** cristaliza en la creación generalizada de **organismos de igualdad** autonómicos, en los últimos años de la

década de los ochenta y los primeros de los noventa. Entre los factores que configuran ese feminismo institucional y que son el marco institucional de las políticas de igualdad, se identifican el tipo de organismo y la posición en el organigrama gubernamental de dicho organismo, concluyendo que el marco institucional más es el que consiste en:

organismos autónomos, creados por ley y con una autonomía presupuestaria y de gestión importantes, normalmente denominados “Institutos de la Mujer”,

adscritos al área gubernamental de presidencia, ya que esta adscripción permite una posición más cercana a los centros de poder que permite hacer mejor la tarea transversal recomendada de “persuasión”, y una visión más global de los temas de género no escorada hacia los problemas de los que se ocupan determinadas áreas gubernamentales.

20. La **definición del problema de la desigualdad de género** ha ido evolucionando en los últimos años, aportándose nuevas definiciones y principios subyacentes al concepto de igualdad de género (igualdad de hecho –e oportunidades y resultados– y no sólo de derecho, igualdad “entre” y no “a”, foco en las relaciones de género y no en las diferencias biológicas, etc.), así como nuevas estrategias de intervención, como es el planteamiento de la transversalidad y normalización de la perspectiva de género (*gender mainstreaming*) en todas las políticas públicas. Esto implica diferentes respuestas por parte de los poderes públicos y, por lo tanto, una evolución también en las políticas públicas.
21. En la entrada del problema de la desigualdad en la **agenda** de actuación de los poderes públicos y en la formulación de las políticas de igualdad de género en España contribuyen una serie de **actores** entre los que destacan el propio feminismo institucional, los partidos políticos –especialmente importante fue la contribución del PSOE en los inicios del feminismo institucional en España–, las organizaciones internacionales y supranacionales –especialmente la ONU y la UE–, el movimiento de mujeres, y el feminismo académico.
22. El principal instrumento con el que se articulan las políticas objeto de estudio son los **planes de igualdad**, que consisten en un conjunto de medidas y objetivos que son aprobados por consejo de gobierno. Los planes fomentan potencialmente una visión estratégica y de conjunto y tienen como principio subyacente la transversalidad de la perspectiva de género, al menos teóricamente; también se identifican y analizan algunas de las debilidades de estos planes de igualdad, como son el posible mimetismo en su elaboración, su gran dependencia del poder de persuasión de las femócratas que están al frente de los organismos de igualdad que los impulsan, y el peligro de quedarse en un plano meramente simbólico y formal.
23. La articulación de las políticas de igualdad en planes se ha generalizado en España, también en el nivel autonómico además de en el central, a lo largo de la primera mitad de la década de los noventa. Esto hace que se haya comenzado también a pensar en términos de evaluación de planes y políticas de igualdad y, por lo tanto, la evaluación de dichos planes se convierta en un interesante objeto de estudio, el cual se aborda en este trabajo y se desarrolla en los capítulos siguientes.

Sobre la metaevaluación como marco para el análisis y la valoración de las evaluaciones de los planes de igualdad de género

24. La fase de evaluación es la menos estudiada desde el análisis de políticas públicas. En general, no existen muchos **estudios sobre los procesos de evaluación** de las políticas públicas, en los cuáles se estudie el desarrollo de la función de evaluación en torno a una determinada política y que describan, analicen y valoren las evaluaciones que se llevan a cabo. La metaevaluación es un enfoque adecuado y novedoso para ello.
25. La **metaevaluación** se define como evaluación de evaluaciones y consiste en analizar y valorar procesos de evaluación. Se debe distinguir de la síntesis de evaluaciones, que consiste en realizar análisis comparados de los resultados de varias evaluaciones y del meta-análisis, que es una síntesis de evaluaciones o investigaciones que utiliza específicamente una técnica cuantitativa definida para dicho tipo de estudios.
26. La metaevaluación tiene un primer significado, vinculado al análisis de políticas públicas, como **descripción, análisis y valoración de los procesos de evaluación**, y que es la perspectiva con la que se realiza este trabajo de investigación, y que se inscribiría en el componente, descrito por Lasswell, del conocimiento del proceso de la política (*knowledge of*) cuyo objetivo es conocer el proceso de la política así como de hecho sucede. De esta manera, la metaevaluación sirve para analizar y valorar la fase de evaluación del ciclo de vida de la política. El segundo significado, más desarrollado en el campo de la evaluación de programas, de la metaevaluación es su función como **control de calidad de las evaluaciones**, tratándose de un seguimiento, fomento y comprobación de la calidad de un servicio profesional como es, en este caso, la evaluación. Con esta idea de fondo, se le asigna a la metaevaluación una función de control del sesgo del evaluador siendo también una forma de asegurar la credibilidad de una evaluación a través de la acreditación.
27. La metaevaluación es, en realidad, una evaluación que tiene por objeto de estudio, en lugar de un programa o política pública, la *evaluación* de un programa o política. Esto quiere decir que puede ser clasificada, como cualquier otra evaluación y, por lo tanto, es posible establecer una **tipología de (meta)evaluaciones** basándose en las mismas clasificaciones básicas ya mencionadas: 1) Formativa o sumativa, según se trate de mejorar la evaluación sobre la marcha o si se realiza una recapitulación de la globalidad de un proceso de evaluación para rendir cuentas sobre el mismo o mejorar otras evaluaciones subsiguientes. 2) De diseño, proceso y/o resultados, según la metaevaluación tenga por objeto de estudio el diseño de la evaluación, su proceso puesta en práctica (cómo se ha llevado la evaluación a cabo), o sus resultados (principalmente el informe, la comunicación de resultados y su posterior utilización). 3) Interna o externa, según se haga por los propios agentes que han realizado la evaluación o por agentes ajenos a la misma, y 4) *Exante* o *expostfacto* según se realice antes o después de que el estudio de evaluación haya concluido.
-

28. Los **criterios** que se proponen en la literatura especializada de evaluación de programas para realizar metaevaluaciones tienen que ver con la relevancia y adecuación (utilidad) de la evaluación, con la utilización de una metodología adecuada y con la obtención de unos resultados de calidad, así como con su viabilidad y su propiedad ética. Nosotros tomamos estos criterios como referencia y, dando prioridad a la metaevaluación como descripción, análisis y valoración de los procesos de evaluación en torno a una determinada política pública, establecimos los criterios para nuestra metaevaluación, organizándolos en tres grandes dimensiones:

- 1) El planteamiento de las evaluaciones: ¿Cómo se han planteado las evaluaciones y cuáles han sido las estrategias evaluativas planteadas? Para esto analizamos 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 realizados y los recursos empleados en los procesos de evaluación.
- 2) Elementos clave de las evaluaciones: ¿Cuál ha sido el papel de algunos elementos clave en la evaluación?, centrándose el estudio en los agentes y los participantes de la evaluación, sus momentos y sus tiempos, los criterios e indicadores, así como procedimientos e instrumentos de recopilación y análisis de la información utilizados.
- 3) La utilización de las evaluaciones: ¿Cuál ha sido la utilización de las evaluaciones?, analizando 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.

29. Centrándonos en las políticas de igualdad de género, en España, el feminismo de Estado ha sido eficaz en la definición de problemas y en la introducción de los mismos en la agenda política, pero no lo ha sido tanto ni en la fase de formulación y ejecución de dichas políticas ni, sobre todo, en la función de control de dicha ejecución. Las “femócratas”, conscientes de la importancia de dicha función de control y el papel que los organismos de igualdad pueden cumplir respecto a la misma, comienzan a dirigir e impulsar, de forma creciente, procesos de evaluación de los planes de igualdad. Por lo tanto, resulta interesante un análisis de dichos procesos de evaluación, valorando la respuesta por parte del feminismo de Estado para el cumplimiento de la función de control de las políticas públicas de igualdad.

Sobre las experiencias de evaluación de planes de igualdad de género en España y los criterios de selección para la metaevaluación

30. De todo el conjunto de evaluaciones realizadas en España de planes de igualdad en los niveles central y autonómico, establecemos una clasificación en primera, segunda y tercera

generación, La primera generación de evaluaciones, realizadas en la primera mitad de la década de los años noventa, consistió únicamente en cuatro experiencias de evaluación (la del primer plan central y las de los primeros planes andaluz, catalán y vasco) que, en general, fueron muy limitadas. La tercera generación de evaluaciones está todavía por definir de forma exacta, ya que consiste en los proyectos –la mayoría de ellos todavía sin concretar en el momento en el que se realiza esta investigación– de procesos de evaluación que tenían previsto concluir a partir del año 2001. Por lo tanto, el objeto de la metaevaluación que aquí se presenta es la segunda generación de evaluaciones que integra once experiencias de evaluación, todas ellas realizadas y concluidas entre los años 1995 y 1999. Una vez aplicado el criterio cronológico, decidimos incluir todas las experiencias existentes de evaluación de los planes en los niveles central y autonómico.

Sobre el planteamiento de las evaluaciones de la segunda generación

31. Teniendo en cuenta el conjunto de experiencias metaevaluadas, los tipos de evaluación predominantes son las evaluaciones sumativas, de resultados (en lo referente a los productos o *outputs*), *expost facto* y externas. En concreto:
- Según el *papel que cumple la evaluación*: Salvo en el caso de la evaluación del plan aragonés (que se realiza expresamente para re-dirigir el plan evaluado, antes de su finalización) y el del vasco que se concibe como un proceso continuado con la intención de ser utilizado sobre la marcha, el resto tiene un claro carácter sumativo, dejando para la evaluación el papel de hacer recapitulación global de lo que ha sido el plan evaluado. En el caso riojano, aunque se concibe como una evaluación continua, ésta se hace más con la intención de seguimiento y se deja para el final la valoración. Este es un buen ejemplo de cómo el hecho de recopilar información desde un principio y con carácter continuado no tiene por qué implicar necesariamente una evaluación formativa, ya que ésta viene determinada por el papel que se le otorga a la misma para ir modificando el plan sobre la marcha.
 - Según el *contenido o fase* evaluada del plan: Con la excepción de la evaluación andaluza y la castellano-manchega, ningún estudio de evaluación se plantea una

evaluación de *diseño* del plan. En los casos vasco y catalán, se recopiló información que, por su naturaleza, pudo dar lugar a algunas conclusiones relacionadas con el diseño de los planes (tipos de actuaciones), pero de una forma parcial. Lo que, en muchos casos, se denomina evaluación de *proceso* se refiere a una evaluación del “grado de cumplimiento” de las acciones del plan: en la mayoría de los casos, esto se refiere a la simple comprobación de que las acciones propuestas se han ejecutado. Si esto se queda sólo en eso, sería más bien una muy simple evaluación de producto (resultados). Sin embargo, la evaluación de proceso es la que valora el proceso de puesta en práctica del plan, incluido el proceso de progresión en la ejecución de las acciones, el ajuste al calendario, si se llega a las personas destinatarias previstas, las dificultades que se presentan en dicho proceso (problemas de información, tiempo o desplazamientos) y todo lo referente al funcionamiento de dicha puesta en práctica incluyendo, sobre todo, los procesos de coordinación con otros departamentos y entidades. Por último, respecto a la evaluación de *resultados*, la parte de los productos o *outputs* queda bastante más cubierta que la de efectos o *outcomes*, que sólo pretende ser cubierta de forma más completa por la evaluación andaluza (y que precisamente, al final del proceso, tuvo muchos problemas para concluir esta parte).

En general, otro problema añadido además de las carencias concretas, es que las evaluaciones realizadas no han tenido un carácter global, desde el punto de vista de su contenido. En muchos casos se ha escogido un tipo de evaluación y, en otros, aunque se haya hecho más de uno, no se aprovechan las sinergias de un análisis más global, combinando los resultados de una y otra evaluación (el caso más ilustrativo es la evaluación del II plan central).

Más allá, podemos concluir que, en las experiencias analizadas, se han practicado muy poco los tipos de evaluación que, en nuestra opinión, serían más interesantes para una evaluación de las políticas de igualdad de género que tuviera relevancia en la actualidad. Esos tipos de evaluación son, por un lado, una evaluación del diseño de los planes que permitiera valorar si los planes son el instrumento de política adecuado, si están bien diseñados y si responden a la definición de (des)igualdad de género que se pretende. Por otro, una verdadera evaluación de proceso, que analizara de forma especial los mecanismos de coordinación transversales necesarios para la aplicación del *mainstreaming*.

- Según el *momento* en el que se evalúa, todas las evaluaciones son *expost facto*.
- Según el *agente* se distinguen seis evaluaciones externas, tres mixtas (que en realidad son procesos internos en los que se recurre, en algún momento, a colaboraciones externas) y dos internas. Esta clasificación, al ser la más

ilustrativa, fue la que se escogió para la exposición de los datos. Esto sirvió también para poder concluir que el modelo mixto es el que parece producir evaluaciones más enraizadas en (y, por lo tanto, más relevantes para) el organismo de igualdad que las impulsa. En definitiva, este modelo mixto permite potencialmente colaboraciones externas adecuadas, pero dirigidas y contextualizadas por los equipos internos que, además, adquieren una fuerte implicación en la evaluación. De hecho, de las tres experiencias mixtas, la que utiliza una colaboración externa más concreta y “mecánica” (la intervención externa en la evaluación del plan central del IM consiste exclusivamente en ejecutar la recopilación y el análisis de los instrumentos diseñados internamente), es decir la más “interna” de las tres, es la que peores resultados obtiene a este respecto, haciendo un claro contraste con las otras dos.

32. En general, se puede afirmar que el **contexto** en el que se generaron las experiencias de evaluación era favorable a las mismas, aunque, como ocurre con todas las dimensiones estudiadas, los contextos son bastante heterogéneos entre sí. Hay variables que sí parecen afectar algo más a las situaciones de evaluación como es el cambio de responsables de los organismos, dándose procesos de evaluación más completos y coherentes en los contextos más estables; también el hecho de haber tenido una experiencia previa en evaluación. Otras, como es que sean impulsadas desde un marco institucional más fuerte (desde organismos autónomos adscritos al área de presidencia) afectan de una forma más indirecta.
33. Los **propósitos** de la evaluación no suelen estar claramente explicitados. En esto hay una diferencia, siendo las evaluaciones externas las que han producido más declaraciones escritas sobre los mismos, dado que suele existir un convenio o un pliego de condiciones técnicas que rigen la colaboración del equipo externo con el organismo de igualdad. Sin embargo, en algunos casos, los propósitos iniciales escritos no concuerdan de forma clara con los procesos de evaluación que posteriormente se llevan a cabo. En cualquier caso, los propósitos que se explicitan, bien en los proyectos iniciales o bien en las justificaciones posteriores (intervenciones parlamentarias o presentaciones en foros públicos), son bastante heterogéneos entre sí.
34. Respecto a las **estructuras institucionales para la evaluación**, existe una diferencia clara entre esta segunda generación de evaluaciones y la siguiente ya que los compromisos explícitos (*mención explícita* a la evaluación en el propio plan aprobado) aumentan sustancialmente en la tercera generación de planes. Con esto podemos concluir que esta segunda generación de evaluaciones sirve, al menos, para que la realidad de la evaluación esté presente y los poderes públicos se sientan comprometidos con ella. Aunque, como se ha visto al analizar esta segunda generación, tampoco la existencia de una mención explícita asegura necesariamente que exista un compromiso real ni que se cumpla dicho compromiso. En general, también es importante la estructura del *personal* de los organismos de igualdad:

en muchos casos, los procesos de evaluación se han llevado a cabo con poca inversión de recursos humanos en los mismos, en algunos casos porque la plantilla del organismo es objetivamente muy pequeña. En general, además de una mayor dedicación, se necesita una mayor formación técnica en evaluación de programas. En lo que se refiere a las *estructuras para la coordinación* de las evaluaciones, parece haber estructuras más adecuadas que otras para asegurar una evaluación. Sin ánimo de establecer unas conclusiones definitivas, creemos que se pueden observar las siguientes tendencias:

- 1) Los fundamentales garantes de los procesos de evaluación son los propios organismos de igualdad. Cuanto más fuerte sea la estructura de dichos organismos, mayores posibilidades de que la estructura institucional para la evaluación sea también más fuerte.
 - 2) Las estructuras creadas “ad-hoc” para coordinar, impulsar, seguir y evaluar los planes son ventajosas porque implican formalmente al resto de los departamentos en una tarea en la que su aportación de información va a ser clave. Pero se corre el peligro de que queden en una mera estructura formal además de ser menos estables institucionalmente.
 - 3) Las estructuras no específicas para los planes tienen, a veces, una fortaleza institucional mayor que las estructuras “ad-hoc” (son incluso, más permanentes), pero también tienen otras funciones que cumplir, lo que puede hacer que la función de evaluación de los planes se diluya.
 - 4) Por ahora, la presencia de la sociedad civil y el movimiento de mujeres (a través de los Consejos de la Mujer u otras estructuras) en los procesos de seguimiento y evaluación de los Planes es meramente simbólica.
 - 5) Los vínculos con el poder legislativo en los procesos de seguimiento y evaluación son siempre impulsores de los mismos. Si el grado de compromiso es importante y está vinculado directamente al organismo de igualdad, como es en el caso del País Vasco, se convierte en un factor crucial para impulsar, desarrollar y mantener procesos de evaluación.
35. La medición de los **recursos económicos** dedicados a las experiencias de evaluación es muy difícil de obtener, dado que habría que obtener datos fiables de los costes indirectos de las mismas y éstos son muy difíciles de ponderar si no se han ido recogiendo datos desde el principio de las evaluaciones y de forma sistemática. Sin embargo, podemos concluir que los recursos invertidos en las evaluaciones analizadas son, a todas luces, escasos, especialmente si se analizan los costes directos de las evaluaciones externas. Lo que en términos absolutos pueden parecer altas sumas, en términos relativos son muy bajas. La evaluación andaluza, que fue la que más dinero costó representa sólo el 0,97% respecto al presupuesto del IAM y 0,07% respecto al presupuesto del plan evaluado.

Sobre los elementos clave de las evaluaciones

36. Respecto a los **agentes y participantes** en las evaluaciones, concluimos la importancia de los *equipos técnicos*, tanto los externos como los internos, así como la coordinación existente entre ellos. De la sensibilidad y empeño de los equipos técnicos internos (que en muchos casos se componen de una sola persona) depende, en muchas ocasiones, que la evaluación vaya delante de forma adecuada. Por eso es especialmente importante la formación específica de este personal porque, en esta segunda generación de evaluaciones, se ha dependido más de la buena voluntad y el sentido común de dicho personal técnico interno, que de su formación técnica en cuestiones de evaluación. En general, aunque la voluntad política es absolutamente necesaria para el éxito de los procesos de evaluación, se necesitan también equipos técnicos preparados, sensibilizados y entusiastas respecto a la evaluación. En lo que se refiere a los *participantes* en las evaluaciones, el gran “caballo de batalla” ha sido la participación de los *gestores de otras unidades administrativas* implicadas en la ejecución del plan evaluado, ya que ésta ha sido, normalmente, muy escasa y ha sido, precisamente, una de las cuestiones que más ha dificultado la recopilación de información. Por último, la participación de *agentes no gubernamentales* ha sido inexistente en la evaluación pero, de alguna manera, acorde con la también escasa participación en la fase de ejecución de los planes evaluados.
37. El **calendario y los momentos** de las evaluaciones estudiadas han supuesto también algunas dificultades importantes, sobre todo en lo referente al *momento del periodo de vigencia del plan* en el que se plantea la evaluación que, con las excepciones del País Vasco y La Rioja, es siempre muy tardío. No haberse planteado la evaluación ni la recopilación de información desde un principio ha acarreado dificultades posteriores insalvables. La *duración* de los procesos de evaluación ha sido muy heterogénea y los *retrasos* han supuesto también dificultades en algunos casos (especialmente Andalucía).
38. El contenido de las evaluaciones se concreta a través de los **criterios e indicadores** de evaluación que representan dos cuestiones fundamentales: por un lado, las variables estudiadas y, por otro, la base lógica en la que se fundamentan los juicios de valor que se establecen posteriormente en el proceso de evaluación. En general, en la práctica totalidad de las experiencias de evaluación metaevaluadas era muy débil la fundamentación de la concreta elección y definición de los criterios e indicadores utilizados. En muchos casos, tuvimos que recurrir a los instrumentos de recopilación utilizados para poder extraer (suponer) las variables estudiadas, ya que en los informes de evaluación no se habían especificado los criterios de evaluación. En general, uno de los indicadores de la poca importancia concedida a la definición previa de los **criterios de evaluación (qué se va a evaluar)** es cómo, en muchos de los informes de evaluación, en los capítulos correspondientes al diseño y metodología de la evaluación, no suele aparecer un apartado específico para ello.
39. De forma similar, los **procedimientos** utilizados tampoco estaban profusamente fundamentados en los informes; es decir, prácticamente, no encontramos una justificación de por qué se eligió un determinado instrumento y no otro. Sí encontramos, en algún caso, especificaciones muy detalladas sobre los instrumentos en sí: estas especificaciones son siempre de carácter técnico y versan sobre las
-

características intrínsecas de dichos instrumentos, no tanto sobre su valía o adecuación para la situación concreta en la que se aplica. En algunos de los casos estudiados, se entiende la metodología y los procedimientos empleados de una forma más amplia (algo más que los instrumentos para recopilar y analizar la información) y se especifican en los informes los pasos dados y las estructuras de coordinación utilizadas para la recopilación de información, como parte de dichos procedimientos. Los casos en los que parece existir esta concepción metodológica más amplia (o al menos se refleja en los documentos e informes resultantes de los procesos de evaluación), son la evaluación del I Plan canario, y el II castellano-mancheño en cierta medida, y del II Plan catalán, el II PAPME vasco y el II Plan gallego, en mayor medida. En estos casos, cuando se describen los procedimientos, se explica de alguna manera y de forma más o menos sistematizada, a través de qué mecanismos se ha recopilado la información, además de explicar las técnicas concretas utilizadas.

Sobre la utilización de los procesos de evaluación

40. Una de los criterios que se analizaron, referidos a la utilización de los procesos de evaluación fue el de la **pertinencia y utilidad de la información producida** por la evaluación. Para ello, se analizó y valoró la globalidad y el alcance de la información, así como el grado de elaboración práctica, es decir, si la información analizada se traduce en conclusiones, valoraciones y recomendaciones para la mejora de los planes. En general, la información producida por las evaluaciones no es demasiado completa o, por decirlo de otro modo, no contempla los planes de forma *global*. La excepción en este caso es el caso andaluz, aunque la globalidad era mayor en el proyecto de evaluación inicial que lo que luego realmente fue. Hay tres casos –Cataluña, Galicia y País Vasco– en el que la información es completa dentro de un tipo de evaluación (*outputs*). El *alcance* de la información es algo mejor, aunque es muy heterogéneo según los casos. Finalmente, el *grado de elaboración práctica* es muy bajo, siendo prácticamente inexistente en muchos de los casos. Como indicador de este bajo grado de elaboración práctica, podemos mencionar el hecho de que ni un solo informe de los consultados contiene unas recomendaciones claras que orienten la mejora de los planes evaluados. Aquí, de nuevo, la excepción es la evaluación del plan vasco ya que, aunque sin recomendaciones, en las memorias anuales hay una parte importante de conclusiones que, además, incluyen conclusiones sobre el propio proceso de evaluación y no exclusivamente sobre los resultados.
41. Respecto a la **comunicación** de los resultados, los *informes* suelen ser bastante incompletos, aunque varían enormemente desde los más exhaustivos, tanto en los
-

procedimientos y la metodología de evaluación como en las valoraciones y conclusiones, como es el caso vasco, a los más “parcos”, como es el caso madrileño, en el que se ofrecen directamente los resultados del estudio sin ofrecer ninguna valoración ni explicar nada sobre los procedimientos empleados en el mismo. En general, los informes tienden a ser incompletos, sobre todo en lo que se refiere a la exposición de conclusiones generales, valoraciones y, especialmente, recomendaciones. Tampoco, salvo alguna excepción, son muy explícitos respecto a la metodología y al proceso seguido. En lo que se refiere a la *accesibilidad* a los informes, también existe una gran variabilidad, habiendo algún caso en el que resultó muy difícil tener acceso a los mismos, incluso a efectos de esta investigación. Esta heterogeneidad va desde la prácticamente nula *difusión* del informe gallego (a pesar de que las intenciones iniciales eran su publicación y su carácter público) a la total accesibilidad de los múltiples informes en los que se reflejan los procesos de evaluación llevados a cabo por el Instituto Vasco de la Mujer (*Emakunde*). Con esta excepción de transparencia respecto al proceso de evaluación del País Vasco, y como es normal que ocurra en un país con escasa cultura y experiencia evaluativa, no suelen identificarse previamente los destinatarios del informe, y existe, por lo general, una cierta ambigüedad sobre el carácter público de los informes de evaluación. También es interesante analizar el *calendario de entrega y difusión de informes*. En algunos casos, el calendario ha dificultado mucho la utilización de la evaluación, al menos su uso instrumental para la mejora del plan siguiente. Se han dado casos en los que la difusión del mismo ha coincidido en el tiempo o ha sido incluso posterior a la aprobación del siguiente plan (las publicaciones de los resultados de la evaluación son simultáneas e incluso posteriores a las aprobaciones de los III sendos planes a nivel central y en la Comunidad de Madrid). En uno de los casos, como es el de la evaluación del II Plan castellano-manchego, la propia entrega del informe final se hace después de que el III Plan haya sido aprobado.

42. El **impacto de las evaluaciones** no sólo se analizó respecto al uso que se ha hecho de los *resultados* de la evaluación, sino también a los usos de los **procesos de evaluación**, sirviendo éstos para introducir una cultura y una metodología de evaluación y para generar una mayor comprensión de la intervención pública en general además de una mayor conciencia sobre cómo se abordan los problemas públicos. De esta manera, concluimos sobre los siguientes tres aspectos:

1) *Efectos sobre los propios procesos de evaluación*. Estas experiencias iniciales de evaluación han hecho que se empiece a hablar de la evaluación como algo normalizado, que haya alguien que asuma las responsabilidades de evaluación en los organismos de igualdad, que se dedique un apartado sobre evaluación en los planes siguientes y que nadie se cuestione la necesidad de evaluar. En algunos casos, como por ejemplo el vasco, se ha generado una cultura evaluativa que ha trascendido, además, el caso concreto y se ha convertido en “modelo” para otras evaluaciones.

2) *Una mayor comprensión o nuevas perspectivas en una situación, programa u organización*. Es evidente que la reflexión sobre los planes que, en algunos casos, proporcionan las evaluaciones ayudan a tener una visión más global, consciente y profunda de lo que se está haciendo y cuáles son las principales dificultades y

fortalezas de las políticas de igualdad. Aunque creemos que, en este sentido, el impacto no ha sido tan grande o, al menos, no ha sido generalizado. Existen muchas diferencias entre las experiencias analizadas, habiendo procesos de evaluación que realmente han producido aprendizaje organizativo (como es el caso del País Vasco y Cataluña) y otros en los que se han obtenido visiones mucho más parciales que no han tenido tanta influencia (como es el caso de la Comunidad de Madrid y el propio Instituto de la Mujer).

3) *Las acciones que se llevan a cabo como resultado de la evaluación.* Este tipo de uso instrumental se ha dado, creemos, en menor medida de lo que habría podido darse si los estudios de evaluación hubieran estado orientados a una verdadera utilización de los mismos. En muchos casos, las acciones tomadas como consecuencia de la evaluación, tienen más que ver con la propia estructuración del plan (concretar más las acciones, asignarles una delimitación temporal y unos responsables concretos para su ejecución) y con su posterior evaluación (adquirir compromisos, aunque generales, respecto a la misma en los subsiguientes planes), que con la modificación de contenidos o estrategias, tanto en un plano global (por ejemplo, dar más peso a un tipo de acciones que a otras) como en uno concreto de cada actuación (modificando cualquier aspecto de cada actuación o actividad a realizar).

Sobre el discurso de las responsables sobre la evaluación y las políticas de igualdad de género

43. En general, se aprecia que el discurso sobre la evaluación por parte de las responsables políticas de los organismos de igualdad **no es un discurso previamente construido**. Esto es lógico si se tiene en cuenta que la práctica evaluativa es muy novedosa y que, además, no existe cultura de evaluación ni referencias teóricas o prácticas en torno a las políticas de igualdad y, por lo general, ni siquiera tampoco en el entorno institucional de las administraciones públicas españolas. En este caso, el discurso sobre su evaluación **está inextricablemente ligado al de las políticas de igualdad de género** y éste, a través de los objetivos y estrategias de dichas políticas, a los conceptos de acciones positivas, perspectiva de género, transversalidad de las políticas, etc. De hecho, en muchos casos, la falta de discurso elaborado sobre evaluación se suple con uno sobre las políticas que se pretenden evaluar. Otra conclusión general del análisis realizado es que **no se puede hablar de un discurso común o monolítico ante la evaluación**. Sin embargo sí se pueden encontrar *algunas ideas comunes* y tendencias en los posicionamientos y, también, *algunas diferencias discursivas* que pueden ser categorizadas, de forma general, en cinco polos o factores: 1) Las responsables en cuyo organismo se ha hecho evaluación y ellas han tenido implicación en ello vs. las responsables en cuyo organismo no se ha hecho evaluación o ellas no han tenido implicación en ello, 2) Las responsables con más tiempo en el cargo vs. las que llevan menos tiempo en él, 3) Las responsables que tienen un pasado y/o feminista

vs. las que no lo tienen, 4) Las responsables en gobiernos de los partidos conservadores vs. otros partidos, y 5) El bagaje profesional de las responsables políticas. De esta manera, el discurso sobre las políticas de igualdad de género, en general y de su evaluación, en particular, aparece como más elaborado el de las responsables cuyo organismo ha hecho evaluación y ellas han estado implicadas en ello, que llevan más tiempo en el cargo, que tienen un pasado de militancia feminista, que están en partidos no conservadores (aunque esto creemos que está más relacionado con haber llegado “de nuevas” a la cuestión de género, ya que el *cleavage* derecha-izquierda, o la ideología, no parece influir tanto en los temas de género) y que tienen un bagaje profesional relacionado con las ciencias sociales (sociología, trabajo social, ...).

44. Entre las **ideas asociadas a la evaluación** y las actitudes ante la misma nos encontramos algunas ideas generalizables al discurso de todas las responsables como, por ejemplo, que la evaluación de las políticas de igualdad de género es algo propio de los organismos de igualdad. Existe también una cierta tendencia general a considerar la evaluación como algo técnico, que debe centrarse en los resultados y que debe ser sobre todo cuantitativa. Entre las ideas asociadas que generan mayores diferencias discursivas destacamos la existencia de ciertas expectativas mágicas sobre evaluación (la evaluación lo soluciona todo), especialmente si en los organismos de igualdad no se tiene experiencia evaluadora (a mayor experiencia y a mayor implicación de la responsable política en la evaluación, visión más realista y ajustada de los procesos evaluativos).
45. Además de las **funciones** clásicas de mejora, responsabilidad e iluminación para acciones futuras, se identifican otros propósitos y fines que asignan a la evaluación las responsables de los organismos. Entre ellas destacamos la de obtener datos e información que justifique la existencia y la necesidad de las acciones públicas en pro de la igualdad, así como la de documentar los planes, difundirlos y hacerlos más visibles. Respecto al para qué de la evaluación, querríamos mencionar lo que hemos denominado la **“ilusión de utilidad”**, que consiste en que las responsables tienen una clara tendencia a decir que utilizan la evaluación más de lo que en realidad parecen hacerlo. Esto tiene que ver también con la **poca capacidad de sorpresa** que las responsables otorgan a la evaluación.
46. El discurso de las responsables denota muy poca claridad respecto a **QUÉ** se quiere evaluar. En general, las responsables apelan a la falta y a la necesidad de **“indicadores”** preestablecidos, pero no señalan cuáles. Al no existir, en la mayoría de los casos, una reflexión previa sobre qué se debe evaluar, se suple esta carencia apelando a criterios “metodológicos” o técnicos, pasando a hablar de instrumentos para recopilar la información. Y otro fenómeno que ocurre al referirse a los contenidos de la evaluación es que existe la tendencia a confundir la necesidad de datos sobre el *problema* de la igualdad con la información sobre los *planes* y cómo se aborda dicho problema. En algunos casos, las responsables más experimentadas sí parecen haber hecho una reflexión algo más seria sobre el tipo de acciones que se realizan en los planes, preguntándose sobre qué es y qué no es una acción positiva. Pero, en general, cuestiones como el diseño de los planes, si éstos son o no el instrumento adecuado y si se consigue la perspectiva de género en el resto de las políticas públicas, no aparecen realmente como objeto de la evaluación en el

discurso de las responsables. Ni siquiera la implicación del resto de las unidades gubernamentales, y su propia capacidad de persuasión y convencimiento, que es algo que les preocupa enormemente, se concibe como algo a evaluar, algo que pueda ser convertido en criterio de evaluación.

47. Las responsables señalan una serie de dificultades a la hora de evaluar que clasificamos en tres factores. Dificultades relacionadas 1) con la falta de cultura de la evaluación, 2) con la propia complejidad de los planes de igualdad y, 3) con la transversalidad o concepción de horizontalidad de las políticas de igualdad de género. Curiosamente, casi todas **las dificultades mencionadas no tienen que ver con los procesos de evaluación en sí, sino con factores previos o en torno a la misma** (o bien falta de cultura o costumbre de hacer evaluación, o bien dificultades que provienen por los propios planes y políticas y no por su evaluación).

- Entre las dificultades relacionadas con la *inexistencia de una cultura de la evaluación*, las responsables mencionan la falta de referencias y de experiencias previas de evaluación, de equipos técnicos preparados y, en general, de una cultura de sistematización de la información y de la carencia de buenos sistemas para ello en las administraciones públicas.
- También detectan la dificultad de evaluar *planes complejos* y que tienen un alto grado de abstracción. Algunas responsables reclaman que especialmente difícil medir los impactos y los costes económicos de las acciones planteadas.
- El hecho de que sean planes transversales que requieren de la implicación de otros actores en la administración pública, hace especialmente costosa la recopilación de información necesaria y la necesaria participación de esos otros agentes en la evaluación. La *perspectiva de género* como fin y el *mainstreaming* como estrategia son conceptos difíciles de explicar y de aplicar. Las responsables se valen de múltiples estrategias formales e informales para esa aplicación. La evaluación de esas estrategias, a veces poco convencionales y dependientes de factores de tipo personal, resulta todo un reto sobre el que, creemos, hay que seguir investigando y discutiendo.

Sobre los factores que facilitan y dificultan la evaluación y como conseguir evaluaciones de las políticas de igualdad de género más eficaces

48. Identificamos, en primer lugar, una serie de **factores previos** a los propios procesos de evaluación, pero que tienen influencia sobre los mismos, tanto negativa como positiva. Estos factores tienen que ver, por ejemplo, con las características específicas de las políticas de igualdad de género (por ejemplo, la existencia de los planes como instrumento de política o la idea de políticas transversales y horizontales) o con características de tipo institucional (por ejemplo, falta de cultura y experiencia previa en evaluación de programas o la fortaleza institucional relativa

de los organismos de igualdad). En nuestro análisis, se señalan cinco factores que *facilitan* la evaluación: la concepción y estructuración de las políticas de igualdad en planes, la creciente institucionalización de compromisos de los poderes públicos respecto a los planes y a su evaluación, el aval y el grado de acuerdo internacional tanto de la necesidad de las políticas como de la estrategia a utilizar (*mainstreaming*), la fortaleza institucional de los organismos de igualdad, así como el reconocimiento simbólico que ha ido adquiriendo la evaluación en los últimos años. Entre los factores previos que dificultan se encuentran el alto grado de generalidad y abstracción de los planes, su carácter horizontal e integral, su carácter no sancionador que hace que la estrategia principal sea de persuasión, la falta de cultura y experiencias previas de evaluación, así como de equipos preparados y, por último, el hecho de que la función de evaluación no esté normalizada en las administraciones públicas españolas y su repercusión en que la evaluación no se piensa a tiempo.

49. En segundo lugar, se analizan los principales problemas identificados en los procesos de evaluación metaevaluados y que tienen que ver con la propia concepción, diseño y ejecución de las respectivas evaluaciones **-factores inherentes** a la misma. Se identificaron doce factores que han influido negativamente en el desarrollo y el éxito final de las experiencias de evaluación. Lógicamente, estos factores no influyen de igual manera en todos los casos e incluso, la identificación de los mismos, a veces se ha verificado por el contraste positivo en alguna de las experiencias.

- 1) La falta de definición y de claridad de los propósitos de la evaluación se traduce en una alta probabilidad de no “acertar” con las necesidades del contexto concreto en el que debe producirse. Especialmente, en algunos casos, no se distingue bien entre lo que significa investigar y hacer seguimiento sobre la desigualdad de género, su evolución y sus múltiples manifestaciones, y lo que es evaluar las acciones que los poderes públicos impulsan para paliar dicho problema. Es necesario distinguir estas dos funciones –investigación y evaluación- dado que ambas son imprescindibles para una mejor formulación de las políticas de igualdad. El hecho, por ejemplo, de que existan “observatorios de igualdad” que investiguen de forma periódica y continuada el problema de la desigualdad de género sobre el que hay que intervenir, facilita no sólo una mejor formulación de los planes y políticas (se conoce mejor el problema que hay que solucionar) sino también la propia evaluación, ofreciendo una información sistematizada sumamente útil para la misma (y que haga más fácil también identificar qué influencia tiene la intervención pública sobre la evolución del problema). Pero su existencia por sí sola no asegura la evaluación de las políticas y los planes. Por otro lado, la evaluación sistemática de planes y políticas puede ayudar a identificar nuevas formas de concebir el problema, nuevas manifestaciones de la desigualdad de género que puedan ser posteriormente investigadas. Por lo tanto, investigación y evaluación son dos funciones necesarias y complementarias.
- 2) En general, no se ha tenido una visión global de las políticas de igualdad de género a la hora de realizar las evaluaciones. Casi todas las experiencias han hecho más énfasis en la evaluación de las acciones que componían los planes

evaluados que en una visión de conjunto del plan y de la política a la que responde dicho plan. Es necesario, por lo tanto, “globalizar” las evaluaciones de forma que haya una mayor conciencia –tanto por parte de los poderes públicos como del resto de actores críticos- de lo que son y significan y el papel que pueden cumplir las políticas públicas de igualdad de género.

- 3) La falta de reconocimiento del carácter político de la evaluación repercute directamente en una inadecuada lectura y respuesta del y al contexto determinado en el que se desarrollan los planes y las políticas a evaluar. Y esto, a su vez, afecta de lleno a la utilidad y eficacia de los procesos de evaluación. Pretender que la evaluación es algo exclusivamente técnico significa abocarla al fracaso.
- 4) La percepción de la evaluación como una función secundaria en las administraciones públicas es una realidad hoy en día. Esta circunstancia ha provocado que, en la mayoría de los organismos de igualdad no se le haya dado la suficiente centralidad a la evaluación. Sin embargo, en nuestra opinión, y dado el carácter de órganos impulsores y coordinadores, más que ejecutores, en los organismos de igualdad se debería otorgar un papel crucial a la evaluación de los planes y las políticas de igualdad de género. Creemos que las responsables de dichos organismos deberían dar prioridad alta a los procesos de evaluación, ya que ésta debería ser una de las funciones principales. En este sentido, siendo una conclusión válida igualmente para el nivel gubernamental central y el autonómico, creemos que el IM central podría tener potencialmente un papel específico sobre la evaluación de las políticas, y que podría consistir en coordinación e intercambio de experiencias, apoyo técnico y documental, apoyo a la coordinación de los sistemas de información, coordinación de foros y grupos especializados, etc. Sin embargo, no ha ejercido apenas liderazgo en este sentido.
- 5) Al no existir referentes ni apenas experiencias previas en evaluación, ni equipos preparados específicamente en evaluación, en las evaluaciones metaevaluadas no se ha hecho, por lo general, una buena gestión de los recursos evaluativos disponibles. Una de las cuestiones que más destaca al respecto es la mala utilización que se ha hecho de los diferentes tipos de evaluación, separando de forma casi absoluta cuando se realizaba más de un tipo de evaluación (de proceso y de resultados, por ejemplo), y sin aprovechar las sinergias de un análisis conjunto.
- 6) Las evaluaciones mixtas parecen, por lo general, más eficaces que las puramente externas o internas. En este caso, no creemos que se deba a la limitación de la colaboración externa sino, sobre todo, a la implicación de los equipos internos en los procesos de evaluación. En las situaciones de evaluación externa, se tendía a delegar en exceso en el equipo externo, entre otras por concebirse la evaluación como algo técnico y perteneciente a los expertos. Sin embargo, en las evaluaciones mixtas, los equipos internos estaban fuertemente implicados y las colaboraciones externas les servían no sólo como guía y vía de legitimación “técnica”, sino también como acicate para el cumplimiento de las tareas y los calendarios. En definitiva, también las

evaluaciones externas son potencialmente eficaces, siempre y cuando cuenten con la implicación y participación activa en los procesos de evaluación de los equipos internos. Y, dada la falta de preparación específica en evaluación y los pocos recursos humanos internos asignados a las funciones de evaluación en los organismos de igualdad, creemos muy conveniente recurrir a colaboraciones externas.

- 7) La valoración de las experiencias denota la falta de reflexión y trabajo previo –técnico y político- para establecer y definir los criterios de evaluación. Esto se ha suplido, en muchos casos, por una metodología de investigación social que se ha centrado más en los instrumentos de recopilación y análisis de la información necesaria. Es necesario saber qué se quiere evaluar y por qué, y luego decidir cuál es la mejor manera de contestarse a las preguntas que, de forma consciente, se hayan establecido previamente.
- 8) Uno de los mayores problemas detectados han sido las dificultades en la recopilación de información, debidas, especialmente a la transversalidad de los planes lo que provoca que la información necesaria esté diluida en múltiples unidades administrativas y existan diversas responsabilidades en la gestión de dicha información. Por lo tanto, más que centrarse en los instrumentos concretos, se debe tener como prioridad el establecimiento de unas buenas estructuras de coordinación –con la participación necesaria de los agentes implicados- y unos buenos sistemas de información que permitan que la información fluya y pueda ser recogida de forma fiable.
- 9) La temporalidad es un aspecto clave en cualquier proceso de evaluación y, como tal, hay que tenerlo en cuenta. El aspecto temporal que más ha dificultado las evaluaciones analizadas ha sido el hecho de que la mayoría se plantearon muy tarde en el tiempo y esa circunstancia limita enormemente las posibilidades de recopilar la información necesaria para evaluar. También hay que evitar los retrasos en los procesos de evaluación porque estos aumentan mucho las probabilidades de que la evaluación llegue tarde o se quede obsoleta.
- 10) El hecho de que las evaluaciones se hayan centrado especialmente en la recopilación y análisis de la información producida por los instrumentos elegidos, también ha provocado que los pasos posteriores (valoraciones y recomendaciones) prácticamente no se hayan resuelto. El déficit de elaboración práctica de la información sistemáticamente recopilada es evidente en prácticamente todos los casos. Incluso en el caso en el que ha existido un mayor grado de elaboración práctica –País Vasco-, éste hubiera podido ser mucho mayor. Raramente se ha entendido que la evaluación incluye también el análisis global, la extracción de conclusiones, la emisión sistemática de juicios de valor y las sugerencias para la acción posterior. Es una lástima que no se hayan “aprovechado” más los procesos de evaluación, “traduciéndolos” sistemáticamente a unos resultados más aplicables.
- 11) También los procesos de comunicación de los resultados de la evaluación han sido normalmente deficitarios. En este sentido destacamos la variabilidad

encontrada respecto al acceso a los informes de evaluación. El carácter público de los informes de evaluación no está claro en muchos casos. Esto nos sugiere una línea de investigación sobre la difusión y acceso a la información producida por las evaluaciones. ¿Qué criterios se están siguiendo en las administraciones públicas españolas? ¿Y cuáles se deberían seguir?

12) Por último, un factor importante es la escasez de recursos dedicados a la evaluación, tanto en términos directos –partidas presupuestarias específicas para la evaluación- como indirectos, especialmente a través de la dedicación de recursos humanos internos de los organismos de igualdad a la función de evaluación. En nuestra opinión, sería lógico que los organismos de igualdad dedicaran al menos un 15% de sus presupuestos a la función de evaluación, y que se dedicara un mínimo del 5% del presupuesto total de cada plan a su evaluación. Esto significaría una inversión mucho mayor de lo que se ha dedicado en esta segunda generación de evaluaciones.

50. Respecto a los **factores relacionados con el impacto** de la evaluación, valoramos en su conjunto la *relativa utilidad y pertinencia* de las evaluaciones. La utilidad de los RESULTADOS de las evaluaciones, por lo general, ha sido muy baja (en el sentido de uso instrumental que consiste en utilizar los resultados de la evaluación para tomar decisiones sobre los planes). De hecho, no parece haber indicios claros de utilización directa para la mejora de las políticas, el rendimiento de responsabilidades o la ilustración para acciones futuras. Si acaso, podemos encontrar, en algunos casos, un cierto uso conceptual (la evaluación sirve para establecer tipologías de acciones y reflexionar sobre las mismas en Cataluña y el País Vasco) y un cierto uso persuasivo (también sirve para demostrar o sacar a la luz los diversos niveles de implicación real en los planes de las diferentes unidades administrativas que habían adquirido un compromiso). Sin embargo, el impacto o la potencial utilización de los PROCESOS de evaluación ha sido relativamente más alta, a pesar de que no existieran objetivos explícitos sobre ello. A pesar de que este impacto hubiera podido ser mucho mayor, estas primeras experiencias marcan pautas de actuación e introducen una cierta cultura y “saber hacer” sobre evaluación. Este tipo de impacto al que nos referimos se hubiera producido, fueran como hubieran sido dichos procesos, debido a la novedad, a la falta de referencias y al hecho de que han sido las primeras experiencias de evaluación para la mayoría de los organismos de igualdad. En general, la realización de estas evaluaciones supone que la evaluación se vaya consolidando como idea y como realidad. Aunque todavía se considere como una función marginal, y no se le otorguen los suficientes recursos ni la centralidad necesaria, la función de evaluación se empieza a fortalecer. Y, sobre todo, a raíz de estas experiencias se da un aprendizaje sobre la función de evaluación, sobre todo en lo que se refiere a las formas y métodos de recopilar información. Por ejemplo, los sistemas de información en torno a los planes mejoran mucho a consecuencia de estas primeras evaluaciones. Este aprendizaje sobre la evaluación, sin embargo, es irregular según las experiencias y creemos que se da en mayor medida cuanto mayor preocupación, atención y conciencia existe sobre la misma, y cuanto mayor actitud de búsqueda abierta existe por parte del equipo técnico de los organismos de igualdad.

Sobre el grado de verificación de las hipótesis iniciales

51. La primera hipótesis (existen una serie de factores que facilitan y dificultan la evaluación de las políticas de igualdad de género) se verifica, diferenciándose además entre los factores previos que condicionan la evaluación, los factores inherentes a los propios procesos de evaluación y que inciden de forma determinante en el éxito o fracaso de la misma, y algunos factores relativos al impacto global de las evaluaciones, como son la relativa pertinencia y utilidad de los procesos o la imposibilidad de realizar una síntesis de los resultados de las evaluaciones.
52. La segunda hipótesis se confirma igualmente en todos los extremos planteados: las evaluaciones más eficaces son las que se plantean de una forma más global, planificándose desde un principio, con ayuda de agentes externos y fuerte implicación interna, con una mayor elaboración o “traducción” práctica, y con una mejor comunicación y publicidad de sus resultados.
53. Respecto a la tercera hipótesis, el grado de verificación es menor que en las dos anteriores, entre otras cosas, por las dificultades inherentes a la recopilación de información sobre el impacto de las evaluaciones. En cualquier caso, y dada la información obtenida, la relación entre las variables planteadas en la hipótesis admite algunos matices. Siguiendo la conclusión nº 50, podemos decir que se confirma el hecho de que la propia existencia de procesos de evaluación ha contribuido a introducir una cierta cultura de evaluación en los organismos de igualdad, pero esta cultura no parece haberse extendido a otras unidades administrativas. También parece que las evaluaciones realizadas han generado una relativa mejora de los sistemas de información en los organismos de igualdad y, en algunos casos concretos, pero no en todos, una mayor conciencia y comprensión de las políticas. Sin embargo, salvo algunas excepciones, las evaluaciones estudiadas no han sido realmente útiles para los propios planes (para introducir mejoras en ellos y rendir responsabilidades sobre los mismos).

Sobre líneas abiertas para futuras investigaciones

Respecto a la evaluación

53. La evaluación debe ser evaluada y los evaluadores evaluados. Por eso es necesario que se desarrolle el enfoque de la metaevaluación en su acepción más concreta de control de calidad de las evaluaciones. Aquí se abre una discusión sobre quiénes deben o pueden ser los metaevaluadores y cuáles deben ser las instancias que acrediten la calidad de las evaluaciones realizadas en las administraciones públicas.

54. La práctica de la evaluación en España es todavía incipiente pero está claramente en alza. Ya que se están empezando a realizar evaluaciones de muy diverso carácter, es conveniente su estudio, análisis y valoración. En este sentido, queda abierta la posibilidad de realizar otras metaevaluaciones de otras políticas públicas que permitan, además, la mejora de la función de evaluación en diferentes sectores.
55. A pesar de que la práctica de la evaluación va aumentando todavía tiene que crecer sustancialmente. Pero no sólo hay que hacer más evaluaciones sino que hay que investigar más sobre las mismas. Hay que saber traducir la práctica profesional de la evaluación en términos científicos y realizar análisis e investigaciones que tengan por objeto la función de la evaluación en el sector público.

Respecto a las políticas de igualdad de género

56. Hay que seguir profundizando en lo que se debe evaluar de las políticas de igualdad de género. En nuestra opinión, en la actualidad sería prioritaria una evaluación que abordara los planes de igualdad como instrumento de política y su adecuación a los objetivos y estrategias de las políticas de igualdad de género. También resulta necesario la evaluación de las estrategias planteadas (*mainstreaming*) y que incluya una visión global sobre su diseño, su implantación y puesta en práctica y sus resultados.
57. Las políticas de igualdad de género y el feminismo institucional en España han sido estudiadas en el nivel central, pero menos en el autonómico y, sobre todo, en el local. Es necesario realizar estudios de caso sobre cada una de las políticas de igualdad autonómicas (sólo se ha hecho de la política andaluza y, parcialmente, de algunos aspectos de la vasca) y también, sobre todo, de las políticas y las instituciones locales cuyo objetivo sea conseguir una mayor igualdad de género.

BIBLIOGRAFIA

- Aguilar, M. J. y E. Ander-Egg (1992). *Evaluación de servicios y programas sociales*. Madrid, Siglo XXI.
- Aguilar Villanueva, L. F., (Ed.). (1996). *El estudio de las Políticas Públicas*. Colección Antologías de Política Pública. Mexico, Miguel Angel Porrúa. Grupo Editorial.
- Aguilar Villanueva, L. F. (1996). "Estudio Introdutorio 1ª". En L. F. A. Villanueva, (Ed.). *El estudio de las Políticas Públicas*. México, Miguel Angel Porrúa. Grupo Editorial. 15-74.
- Aguilar Villanueva, L. F. (1996). "Estudio introductorio 2ª". En L. F. A. Villanueva, (Ed.). *La hechura de las Políticas Públicas*. México, Miguel Angel Porrúa. Grupo Editorial. 15-84.
- Aguilar Villanueva, L. F. (1996). "Estudio introductorio 3ª". En L. F. A. Villanueva, (Ed.). *Problemas Públicos y Agenda de Gobierno*. México, Miguel Angel Porrúa. Grupo Editorial. 15-72.
- Aguilar Villanueva, L. F. (1996). "Estudio introductorio 4ª". En L. F. A. Villanueva, (Ed.). *La implementación de las políticas* México, Miguel Angel Porrúa. Grupo editorial. 15-92.
- Aguilar Villanueva, L. F., (Ed.). (1996). *La Hechura de las Políticas Públicas*. Colección Antologías de Política Pública. México, Miguel Angel Porrúa. Grupo Editorial.
- Aguilar Villanueva, L. F., (Ed.). (1996). *La implementación de las Políticas*. Colección Antologías de Política Pública. México, Miguel Angel Porrúa. Grupo Editorial.
- Aguilar Villanueva, L. F., (Ed.). (1996). *Problemas Públicos y Agenda de Gobierno*. Colección Antologías de Política Pública. México, Miguel Angel Porrúa. Grupo Editorial.
- Aja, E. (1999). *El Estado autonómico. Federalismo y hechos diferenciales*. Madrid, Alianza Editorial.
- Alba, C. y F. J. Vanaclocha (1997). *El sistema político local : un nuevo escenario de gobierno*. Madrid, Universidad Carlos III de Madrid: Boletín Oficial del Estado.
- Alberdi, C. (1997). "Estrategias para conseguir la igualdad". En E. U. y. A. Elizondo, (Ed.). *Mujeres en Política. Análisis y práctica*. Barcelona, Ariel. 303-311.
- Alberdi, I., P. Escario, et al. (1996). *Lo personal es político. El movimiento feminista en la transición*. Madrid, Ministerio de Asuntos Sociales. Instituto de la Mujer.

-
- Alker, H. R. (1996). "Political Methodology, Old and New". En R. Goodin y H.-D. Klingemann, (Ed.). *A new Handbook of Political Science*. Nueva York, Oxford University Press. 787-799.
- Alkin, M. C. (1985). *A Guide for Evaluation Decision Makers*. Beverly Hills, CA, Sage.
- Alkin, M. C., (Ed.). (1990). *Debates on Evaluation*. Newbury Park, CA, Sage.
- Alkin, M. C., R. Daillak, et al. (1979). *Using Evaluations: Does Evaluation Make a Difference?* Beverly Hills, CA, Sage.
- Allison, G. T. ([1969] 1996). "Modelos conceptuales y la crisis de los misiles cubanos. Apéndice: el bloqueo norteamericano a Cuba en ocasión de la crisis de los misiles cubanos. Un estudio de caso.". En L. Aguilar, (Ed.). *La hechura de las políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 119-200.
- Altschuld, J. W. y M. Engle, (Ed.). (1994). *The Preparation of Professional Evaluators: Issues, Perspectives, and Programs*. New Directions for Program Evaluation. San Francisco, Jossey-Bass.
- Alvaro, M. (1996). *Los usos del tiempo como indicadores de la discriminación entre géneros*. Madrid, Ministerio de Asuntos Sociales., Instituto de la Mujer.
- Alvira, F. (1991). *Metodología de la evaluación de programas*. Madrid, CIS.
- Alvira, F. (1997). *Metodología de la evaluación de programas: un enfoque práctico*. Buenos Aires, Lumen/Hvmanitas.
- Amorós, C., (Ed.). (1995). *10 Palabras Clave sobre Mujer*. Estella, Verbo Divino.
- Anduiza, E., I. Crespo, et al. (1999). *Metodología de la Ciencia Política*. Madrid, Centro de Investigaciones Sociológicas.
- Arenillas, M. y J. M. Canales, (Ed.). (1999). *Gobierno y pacto local*. Madrid, BOE.
- Arranz, F., (Ed.). (2000). *Las Políticas Públicas en favor de las mujeres*. Madrid, Instituto de Investigaciones Feministas. Universidad Complutense de Madrid.
- Arteaga, T. (1997). "Planes de Acción Positiva en la Comunidad Autónoma del País Vasco". En E. U. y. A. Elizondo, (Ed.). *Mujeres en Política. Análisis y Práctica*. Barcelona, Ariel. 319-324.
- Ascher, W. ([1986]1996). "La evolución de las ciencias de políticas". En L. Aguilar, (Ed.). *El estudio de las Políticas Públicas*. México, Miguel Angel Porrúa. Grupo Editorial. 181-196.
- Astelarra, J., (Ed.). (1990). *Participación política de las mujeres*. Madrid, Centro de Investigaciones Sociológicas.
-

-
- Astelarra, J. (1998). "Alcance y limitaciones de las políticas de género". En P. d. Villota, (Ed.). *Las mujeres y la ciudadanía en el umbral del siglo XXI*. Madrid, Editorial Complutense. 213-224.
- Bachi, C. L. (1996). *The Politics of Affirmative Action*. London, Sage.
- Baena, M. (1985). *Curso de Ciencia de la Administración*. Madrid, Tecnos.
- Ballart, X. (1992). *¿Cómo evaluar programas y servicios públicos? Aproximación sistemática y estudios de caso*. Madrid, Ministerio para las Administraciones Públicas.
- Ballart, X. (1993). "Evaluación de políticas. Marco conceptual y organización institucional." **Revista de Estudios Políticos** 80(Abril-Junio 1993): 199-224.
- Ballart, X. (1996). "Modelos teóricos para la práctica de la evaluación de programas". En Q. Brugé y J. Subirats, (Ed.). *Lecturas de Gestión Pública*. Madrid, Ministerio para las Administraciones Públicas. 323-351.
- Ballart, X. (2000). "La Administración de un Presidente: Cataluña 1980-1997". En J. L. Paniagua Soto, (Ed.). *Gobierno y Administración en las Comunidades Autónomas. Andalucía, Canarias, Cataluña, Galicia y País Vasco*. Madrid, Tecnos. 169-203.
- Ballart, X. y C. Ramió (2000). *Ciencia de la administración*. Valencia, Tirant lo Blanch.
- Bamberger, M. (1999). "Ethical Issues in Conducting Evaluation in International Settings". En J. L. Fitzpatrick y M. Morris, (Ed.). *Current and Emerging Ethical Challenges in Evaluation*, Jossey-Bass. 89-98.
- Bamberger, M. (2000). "The Evaluation of International Development Programs: A View from the Front." **The American Journal of Evaluation** 21(1): 95-101.
- Bañón, R. (1993). "La modernización de la administración pública española: balance y perspectivas." **Política y Sociedad** 13.
- Bañón, R. (1997). "Los enfoques para el estudio de la administración pública: orígenes y tendencias actuales". En R. Bañón y E. Carrillo, (Ed.). *La nueva Administración Pública*. La nueva Administración Pública Madrid, Alianza Universidad. 17-50.
- Bañón, R. y E. Carrillo (1995). "Evaluación de la calidad de los servicios públicos". En R. Bañón, A. Halachmi y G. Bouckaert, (Ed.). *La productividad y la calidad en la gestión pública*. La productividad y la calidad en la gestión pública Santiago de Compostela, EGAP.
- Bañón, R. y E. Carrillo, (Ed.). (1997). *La nueva Administración Pública*. Madrid, Alianza Universidad.
-

-
- Bañón, R. y J. L. Paniagua (2000). "Gobierno y Administración en el Estado de las Autonomías y en las Comunidades Autónomas". En J. L. Paniagua Soto, (Ed.). *Gobierno y Administración en las Comunidades Autónomas. Andalucía, Canarias, Cataluña, Galicia y País Vasco*. Madrid, Tecnos. 265-284.
- Barbier, J.-C. (1999). "Inter-governmental Evaluation: Balancing Stakeholders' Expectations with Enlightenment Objectives." **Evaluation** 5(4): 373-386.
- Bardach, E. ([1981] 1996). "Problemas de la definición de problemas en el análisis de políticas". En L. Aguilar, (Ed.). *Problemas públicos y agenda de gobierno*. México, Miguel Angel Porrúa. Grupo Editorial. 219-234.
- Barzelay, M. y B. J. Armajani (1992). *Breaking Through Bureaucracy. A New Vision for Managing in Government*. Berkeley, University of California Press.
- Barzelay, M. y J. M. O'Kean (1992). *Gestión Pública Estratégica. Conceptos, análisis y experiencias: el caso IPIA*. Madrid, Instituto de Estudios Fiscales.
- Behn, R. D. ([1981]1996). "El análisis de políticas y la política". En L. Aguilar, (Ed.). *El estudio de las Políticas Públicas*. México, Miguel Angel Porrúa. Grupo Editorial. 239-274.
- Berk, R. y P. Rossi (1990). *Thinking About Program Evaluation*. Newbury Park, Sage.
- Berman, P. ([1978] 1996). "El estudio de la macro y la micro-implementación". En L. Aguilar, (Ed.). *La implementación de las Políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 281-322.
- Bickman, L., (Ed.). (1990). *Advances in Program Theory. New Directions for Program Evaluation, 47*. San Francisco, Jossey-Bass.
- Bickman, L. (1994). "An Optimistic View of Evaluation." **Evaluation Practice** 15(3): 255-59.
- Birckmayer, J. D. y C. H. Weiss (2000). "Theory-Based Evaluation in Practice. What Do We Learn?" **Evaluation Review** 24(4. August 2000): 407-431.
- Bologh, R. (1991). *Love or Greatness: Max Weber and Masculine Thinking*. Londres, Unwin Hyman.
- Borchost, A. y B. Siim (1987). "Women and the advanced welfare state -a new kind of patriarchal power?". En A. Showstack Sassoon, (Ed.). *Women an te State: The Shifting Boundaries of public and private*. Women an te State: The Shifting Boundaries of public and private Londres, Unwin Hyman. 128-157.
- Bustelo García del Real, C. (1990). "Estrategias futuras en la intervención de la administración". En J. Astelarra, (Ed.). *Participación política de las mujeres*. Madrid, Centro de Investigaciones Sociológicas. 269-280.
-

-
- Bustelo Ruesta, M. (1998). "Deontología de la evaluación: el modelo de los códigos éticos anglosajones." **Gestión y Análisis de Políticas Públicas** 11-12(Enero/Agosto 1998): 141-156.
- Bustelo Ruesta, 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(Primavera-verano 1999): 9-29.
- Bustelo Ruesta, M. (1999). "Políticas Públicas de Igualdad de Género en España: evolución y evaluación". En M. Ortega, C. Sánchez and C. Valiente, (Ed.). *Género y ciudadanía. Revisiones desde el ámbito privado*. Madrid, Ediciones de la Universidad Autónoma de Madrid. 367-389.
- Bustelo Ruesta, M. (2000). "El reto de la evaluación de las políticas de igualdad: Hacia una evaluación de la perspectiva de género." **Cuadernos de Cátedra Abierta** 13(Noviembre 2000): 7-50.
- CAD (1997). *Evaluating programs promoting participatory development and good governance. Synthesis Report*. Grupo de Expertos en Evaluación de la Ayuda del CAD-OCDE. París.
- CAD (1998). *Directrices y guía de conceptos del CAD (OCDE) sobre la igualdad entre mujeres y hombres*. Madrid, Ministerio de Asuntos Exteriores. Secretaría de Estado para la Cooperación Internacional y para Iberoamérica.
- Campbell, D. T. (1969). "Reforms as Experiments." **American Psychologist** 24: 409-29.
- Campbell, D. T. (1991). "Methods for the Experimenting Society." **Evaluation Practice** 12(3): 223-60.
- Campbell, D. T. y J. C. Stanley (1963). *Experimental and Quasi-Experimental Designs for Research*. Chicago, Rand McNally.
- Camps, V. (1998). *El siglo de las mujeres*. Madrid, Cátedra.
- Canales Aliende, J. M. (1987). *Panorama actual de la Ciencia de la Administración*. Madrid, INAP.
- Caracelli, V. J. y H. Preskill, (Ed.). (2000). *The Expanding Scope of Evaluation Use. New Directions for Program Evaluation*, n. 88. San Francisco, Jossey-Bass.
- Cardenal, V. e. a. (1994). *Evaluación de programas de reinserción social*. Málaga, Miguel Gómez Ediciones.
- Carter, N. y G., P. (1993). "Evaluating Agencies: Next Steps and Performance Indicators." **Public Administration** 71(Autumn 1993): 407-416.
Experiencias en Gran Bretaña con la Reforma del sector público
-

-
- Castells, C. (1996). "Introducción". En C. C. (comp.), (Ed.). *Perspectivas feministas en teoría política*. Barcelona, Paidós. 9-30.
- Castells, C. c., (Ed.). (1996). *Perspectivas feministas en teoría política*. Paidós Estado y Sociedad. Barcelona, Paidós.
- Caudle, S. L. (1989). "Evaluating Program Results and Success". En R. E. C. y N. Henry, (Ed.). *Managing Public Programs. Balancing Politics, Administration and Public Needs*. San Francisco, Jossey-Bass. 231-254.
- Cavana, M. L. (1995). "Diferencia". En C. Amorós, (Ed.). *10 Palabras Clave sobre Mujer*. Estella, Verbo Divino. 85-118.
- Chelimsky, E., (Ed.). (1985). *Program Evaluation: Patterns and Directions*. Washington, DC, American Society for Public Administration.
- Chelimsky, E. (1987). "What We Have Learned About the Politics of Program Evaluation." **Educational Evaluation and Policy Analysis** 9: 199-213.
- Chelimsky, E. (1989). "Evaluating Public Programs". En J. L. Perry, (Ed.). *Handbook of Public Administration*. Handbook of Public Administration San Francisco, Jossey-Bass. 259-273.
- Chelimsky, E. (1992). "Expanding Evaluation Capabilities in the General Accounting Office". En C. Wye y R. C. Sonnichsen, (Ed.). *Evaluation in the Federal Government: Changes, Trends and Opportunities*. San Francisco, Jossey-Bass.
- Chelimsky, E. (1995a). "Comments on the AEA Guiding Principles for Evaluators". En W. Shadish, D.L. Newman, M.A. Scherirer y C. Wye, (Ed.). *Guiding Principles for Evaluators*. Guiding Principles for Evaluators Washington, DC, American Society for Public Administration. 53-54.
- Chelimsky, E. (1995b). "The Political Environment of Evaluation and What it Means for the Development of the Field." **Evaluation Practice** 16(3): 215-225.
- Chelimsky, E. (1997). "The Coming Transformations in Evaluation". En E. Chelimsky y W. R. Shadish, (Ed.). *Evaluation for the 21st Century. A Handbook*. Evaluation for the 21st Century. A Handbook Thousand Oaks, Sage. 1-26.
- Chelimsky, E. (1998). "The role of Experience in Formulating Theories of Evaluation Practice." **American Journal of Evaluation** 19(1): 35-55.
- Chelimsky, E. y W. R. Shadish, (Ed.). (1997). *Evaluation for the 21st Century. A Handbook*. Thousand Oaks, Sage.
- Chen, H. T., Rossi, P.H. (1983). "Evaluating with Sense: The Theory-Driven Approach." **Evaluation Review** 7: 283-302.
-

-
- Chen, H. T. (1990). *Theory-Driven Evaluation: A Comprehensive Perspective*. Newbury Park, Sage.
- Chen, H. T. (1994). "Theory-Driven Evaluations: Needs, Difficulties, and Options." **Evaluation Practice** 15(1): 79-82.
- Chen, H. T. (1996). "A Comprehensive Typology for Program Evaluation." **Evaluation Practice** 17(2): 121-130.
- Cirujano, P. (2000). "Beijing + 5: Hacia el cumplimiento de los compromisos con las mujeres." **Revista Española de Desarrollo y Cooperación** n° 6 Primavera/Verano 2000: 113-128.
- Cobo Bedia, R. (1995). "Género". En C. Amorós, (Ed.). *10 Palabras Clave sobre Mujer*. Estella, Verbo Divino. 55-84.
- Cohen, E. y R. Franco (1988). *Evaluación de proyectos sociales*. Buenos Aires, Grupo Editor Latinoamericano.
- Cohen, E. y F., R. (1988). *Evaluación de Proyectos Sociales*. Buenos Aires, Grupo Editor Latinoamericano.
- Conrad, K. J., (Ed.). (1994). *Critically Evaluating the Role of Experiments*. New Directions for Program Evaluation, n.63. San Francisco, Jossey-Bass.
- Cook, T. D. (1997). "Lessons Learned in Evaluation Over the Past 25 Years". En E. Chelimsky y W. R. Shadish, (Ed.). *Evaluation for the 21st Century. A Handbook*. Evaluation for the 21st Century. A Handbook Thousand Oaks, Sage. 30-52.
- Cook, T. D. y D. T. Campbell (1979). *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Chicago, Rand McNally.
- Cook, T. D. y C. Reichart (1986). *Métodos cualitativos y cuantitativos en investigación evaluativa*. Madrid, Morata.
- Cooper, P. J. e. a. (1998). *Public Administration for the Twenty-First Century*. Fort Worth, Harcourt Brace College Publishers.
- Cordray, D. S., Lipsey, M.W. (1987). "Evaluation Studies for 1986: Program Evaluation and Program Research." **Evaluation Studies Review Annual** 11(1986): 17-44.
- Cordray, D. S. (1993). "Strengthening Causal Interpretations of Nonexperimental Data: The Role of Meta-analysis". En L. Sechrest, (Ed.). *Programa Evaluation: A pluralistic Enterprise*. Programa Evaluation: A pluralistic Enterprise San Francisco, Jossey-Bass. 59-97.
- Cotarelo, R. (1987). *Del Estado del Bienestar al Estado del Malestar*. Madrid, Centro de Estudios Constitucionales.
-

-
- Cotarelo, R. (1990). *En torno a la teoría de la democracia*. Madrid, Centro de Estudios Constitucionales.
- Council, of, et al. (1998). *Gender Mainstreaming. Conceptual framework, methodology and presentation of good practice. Final Report of Activities of the Group of Specialists on Mainstreaming (EG-S-MS)*. Estrasburgo, Council of Europe Publishing.
- Covert, R. e. a., (Ed.). (1988). *Evaluation Utilization*. New Directions for Program Evaluation, 39. San Francisco, Jossey-Bass.
- Cronbach, L. J. (1982). *Designing Evaluations of Educational and Social Programs*. San Francisco, Jossey-Bass.
- Cronbach, L. J. y Associates (1980). *Toward the Reform of Program Evaluation*. San Francisco, Jossey-Bass.
- Dahlerup, D. (1987). "Confusing concepts- confusing reality: A theoretical discussion of the patriarchal state". En A. S. Sassoon, (Ed.). *Women and the state: The shifting boundaries of public and private*, Unwin Hyman. 93-127.
- Dancausa, C. (1997). "Las políticas de igualdad en los últimos veinte años: logros y límites". En E. U. y. A. Elizondo, (Ed.). *Mujeres en política. Análisis y Práctica*, Ariel. 312-318.
- Datta, L.-e. (1999). "The Ethics of Evaluation Neutrality and Advocacy". En J. L. Fitzpatrick y M. Morris, (Ed.). *Current and Emerging Ethical Challenges in Evaluation*. San Francisco, Jossey-Bass. 77-88.
- Datta, L.-e. (2000). "Seriously Seeking Fairness: Strategies for Crafting Non-partisan Evaluation in a Partisan World." **The American Journal of Evaluation** 21(1): 1-14.
- Denzin, N. K. y Y. S. Lincoln (1994). *Handbook of Qualitative Research*. Thousand Oaks, CA, Sage.
- Díez López, M. A. (2000). *La evaluación de la política regional: propuestas para evaluar las nuevas políticas regionales*. Bilbao, Universidad del País Vasco.
- Downs, A. ([1972] 1996). "El ciclo de atención a los problemas sociales. Los altibajos de la ecología". En L. Aguilar, (Ed.). *Problemas públicos y agenda de gobierno*. México, Miguel Angel Porrúa. Grupo Editorial. 141-160.
- Dror, Y. ([1964] 1996). "Salir del paso ¿"Ciencia" o inercia?". En L. Aguilar, (Ed.). *La hechura de las políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 255-264.
-

-
- Dror, Y. ([1970]1996). "Prolegómenos para las ciencias de políticas". En L. Aguilar, (Ed.). *El estudio de las Políticas Públicas*. México, Miguel Angel Porrúa. Grupo Editorial. 119-148.
- Dunn, W. N. (1981). *Public Policy Analysis: An Introduction*. Englewood Cliffs, N.J., Prentice-Hall.
- Dunn, W. N., (Ed.). (1986). *Policy Analysis: Perspectives, Concepts, and Methods*. Public Policy Studies: A Multi-Volume Treatise. Greenwich, CT., Jai Press Inc.
- Dye, T. R. (1975). *Understanding Public Policy*. Englewood Cliffs, Prentice Hall.
- Elder, C. D. y R. W. Cobb ([1984] 1996). "Formación de la agenda. El caso de la política de acianos". En L. Aguilar, (Ed.). *Problemas públicos y agenda de gobierno*. México, Miguel Angel Porrúa. Grupo Editorial. 77-104.
- Elizondo, A. y. y E. Martínez (1995). "Presencia de mujeres y políticas para la igualdad entre los sexos: el caso de las instituciones políticas vascas (1980-1994)." **Revista de Estudios Políticos** 89: 345-368.
- Elmore, R. ([1978] 1996). "Modelos organizacionales para el análisis de la implementación de programas sociales". En L. Aguilar, (Ed.). *La implementación de las Políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 185-250.
- Elmore, R. ([1979-80] 1996). "Diseño retrospectivo: la investigación de la implementación y las decisiones políticas". En L. Aguilar, (Ed.). *La implementación de las Políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 251-280.
- Espinoza, M. (1983). *Evaluación de Proyectos Sociales*. Buenos Aires, Hvmantas.
- Etzioni, A. ([1967] 1996). "La exploración combinada: un tercer enfoque de la toma de decisiones". En L. Aguilar, (Ed.). México, Miguel Angel Porrúa. Grupo Editorial. 265-286.
- Evans, G. (1993). "Is gender on the "new agenda"? A comparative analysis of the politicization of inequality between men and women." **European Journal of Political Research** 24: 135-158.
- Ferguson, K. (1984). *The feminist Case Against Bureaucracy*. Philadelphia, Temple University Press.
- Fernández Ballesteros, R., (Ed.). (1995). *Evaluación de programas. Una guía práctica en ámbitos sociales, educativos y de salud*. Madrid, Síntesis.
- Fernández Sierra, J. y M. A. Santos Guerra (1992). *Evaluación cualitativa de programas de educación para la salud. Una experiencia hospitalaria*. Málaga, Aljibe.
-

-
- Ferree, M. M. (1995). "Making Equality. The Women's Affairs Offices in the Federal Republic of Germany". En S. y. Mazur, (Ed.). *Comparative State Feminism*. Comparative State Feminism Londres, Sage. 95-113.
- Fetterman, D. M. (1994). "Empowerment Evaluation. American Evaluation Assotiation Presidential Address." **Evaluation Practice** 15(1): 1-15.
- Fetterman, D. M. e. a., (Ed.). (1996). *Empowerment Evaluation. Knowledge and Tools for Self-Assessment and Accountability*. Thousand Oaks, Sage.
- Fitz-Gibon, C. T. y L. L. Morris (1978). *How To Design a Program Evaluation*. Beverly Hills, Sage.
- Fitzpatrick, J. L. (1999). "Ethics in Disciplines and Professions Related to Evaluation". En J. L. Fitzpatrick and M. Morris, (Ed.). *Current and Emerging Ethical Challenges in Evaluation*. San Francisco, Jossey-Bass. 5-14.
- Fitzpatrick, J. L. y M. Morris, (Ed.). (1999). *Current and Emerging Ethical Challenges in Evaluation*. New Directions for Program Evaluation, n. 82. San Francisco, Jossey-Bass.
- Forbes, I. (1991). "Equal Opportunity: Radical, Liberal and Conservative Critiques". En E. Meehan y S. Sevenhuijsen, (Ed.). *Equality Politics and Gender*. London, Sage. 17-35.
- Forester, J. ([1984] 1996). "La racionalidad limitada y la política de salir del paso". En L. Aguilar, (Ed.). *La hechura de las políticas*, Miguel Angel Porrúa. Grupo Editorial. 315-340.
- Fournier, D. M., (Ed.). (1995). *Reasoning in Evaluation: Inferential Links and Leaps*. New Directions for Program Evaluation, n. 68. San Francisco, Jossey-Bass.
- Gallego Méndez, M. T. (1998). "Visión del Estado y ciudadanía: los derechos políticos". En P. d. Villota, (Ed.). *Las mujeres y la ciudadanía en el umbral del siglo XXI*. Madrid, Editorial Complutense. 83-90.
- Gardiner, F. (1997). "La Unión Europea: ¿hada madrina para las mujeres?". En E. U. y. A. Elizondo, (Ed.). Barcelona, Ariel. 246-270.
- Garson, G. D. ([1986]1996). "De la ciencia de políticas al análisis de políticas: veinticinco años de progreso". En L. Aguilar, (Ed.). *El estudio de las Políticas Públicas*. México, Miguel Angel Porrúa. Grupo Editorial. 149-180.
- Gil Ruiz, J. M. (1996). *Las políticas de igualdad en España: Avances y retrocesos*. Granada, Universidad de Granada.
- Gilligan, C. (1982). *In a Different Voice, Psychological Theory and Women's Development*. Cambridge, Mass., Harvard University Press.
-

-
- Gispert, C. (1990). "Políticas públicas a favor de las mujeres". En J. Astelarra, (Ed.). *Participación política de las mujeres*. Madrid, Centro de Investigaciones Sociológicas. 241-251.
- Gomá, R. and J. Subirats, (Ed.). (1998). *Políticas públicas en España : contenidos, redes de actores y niveles de gobierno*. Barcelona, Ariel.
- Goodin, R. and I. Waldner ([1979] 1996). "Pensar en grande, pensar en pequeño y simplemente no pensar". En L. Aguilar, (Ed.). *La hechura de las políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 283-314.
- Goodin, R. E. and H.-D. Klingemann, (Ed.). (1996). *A New Handbook of Political Science*. Nueva York, Oxford University Press.
- Granados Vaquero, E. (1997). *El Instituto Andaluz de la Mujer: organización, presupuesto y actuaciones*. III Congreso Español de Ciencia Política y de la Administración, Salamanca.
- Granados Vaquero, E. (1999). "El Instituto Andaluz de la Mujer: La creación de un órgano autonómico para la igualdad entre los géneros". En M. Ortega, C. Sánchez y C. Valiente, (Eds.). *Género y ciudadanía. Revisiones desde el ámbito privado*. Madrid, Ediciones de la Universidad Autónoma de Madrid. 391-405.
- Granados Vaquero, E. (1999). *La política Pública de Igualdad entre los Géneros en la Comunidad Autónoma de Andalucía (1989-1998)*. Granada, Universidad de Granada.
- Greene, J. y V. Caracelli, (Ed.). (1997). *Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms*. New Directions for Program Evaluation. San Francisco, Jossey-Bass.
- Greene, J. C. (1988). "Stakeholder participation and utilization in program evaluation." **Evaluation Review** 12(2): 96-116.
- Greene, J. C. (1990). "Technical Quality Versus User Responsiveness in Evaluation Practice." **Evaluation and Program Planning** 13(3): 267-74.
- Greene, J. C. (1997). "Evaluation as Advocacy." **Evaluation Practice** 18(1): 22-35.
- Gregory, A. (2000). "Problematizing Participation: A Critical Review of Approaches to Participation in Evaluation Theory." **Evaluation** 6(2): 179-200.
- Guba, E. (1990). *The Paradigm Dialog*. Newbury Park, Sage.
- Guba, E. y Y. Lincoln (1981). *Effective Evaluation: Improving the Usefulness of Evaluation Results Through Responsive and Naturalistic Approaches*. San Francisco, Jossey-Bass.
-

-
- Guba, E. y Y. Lincoln (1987). "The Countenances of Fourth-Generation Evaluation: Description, Judgment, and Negotiation." **Evaluation Studies Review Annual 11**(1986): 70-88.
- Guba, E. y Y. Lincoln (1989). *Fourth Generation Evaluation*. Newbury Park, Sage.
- Gudiño, F. (1999). "El sistema de evaluación de la cooperación española." **Revista Española de Desarrollo y Cooperación 4**(Primavera-verano 1999): 31-49.
- Guy Peters, B. y V. Wright (1996). "Public Policy and Administration, Old and New". En R. Goodin and H.-D. Klingemann, (Ed.). *A new Handbook of Political Science* Nueva York, Oxford University Press. 628-641.
- Hecló, H. (1972). "Policy Analysis." **British Journal of Political Science 2**: 83-108.
- Hecló, H. ([1978] 1996). "Las redes de asuntos y el poder ejecutivo". En L. Aguilar, (Ed.). *Problemas públicos y agenda de gobierno*. México, Miguel Angel Porrúa. Grupo Editorial. 257-284.
- Henry, G. T. (2000). "Why Not Use?". En V. J. Caracelli y H. Preskill, (Ed.). *The Expanding Scope of Evaluation Use*. San Francisco, Jossey-Bass. 85-99.
- Hernández Bravo, J. (2000). "Gobierno y Administración en la Comunidad Autónoma de Canarias". En J. L. Paniagua Soto, (Ed.). *Gobierno y Administración en las Comunidades Autónomas. Andalucía, Canarias, Cataluña, Galicia y País Vasco*. Madrid, Tecnos. 205-263.
- Hernes, H. M. (1990). *El poder de las mujeres y el Estado de bienestar*. Madrid, Vindicación Feminista. *Original en inglés. Oslo: Universitetsforlaget (1987)*.
- Hofferbert, R. I. y D. L. Cingranelli (1996). "Public Policy and Administration: Comparative Policy Analysis". En R. Goodin y H.-D. Klingemann, (Ed.). *A new Handbook of Political Science*. Nueva York, Oxford University Press. 593- 609.
- Hogwood, B. W. y L. W. Gunn (1984). *Policy Analysis for the Real World*. Oxford, Oxford University Press.
- House, E. (1990). "Trends in Evaluation." **Educational Researcher 20**(6): 2-9.
- House, E. (1994). *Evaluación, ética y poder*. Madrid, Morata.
- House, E. (1996). "In-House Reflection: Internal Evaluation." **Evaluation Practice 7**(1): 63-64.
- House, E. R. (1993). *Professional Evaluation. Social Impact and Political Consequences*. Newbury Park, CA, Sage.
- Howlett, M. y M. Ramesh (1995). *Studying Public Policy. Policy Cycles and Policy Subsystems*. Oxford, Oxford University Press.
-

-
- Jackson, J. E. (1996). "Political Methodology: An Overview". En R. Goodin y H.-D. Klingemann, (Ed.). *A new Handbook of Political Science*. Nueva York, Oxford University Press. 717-748.
- Jaggar, A. (1988). *Feminist Politics and Human Nature*. Totowa, Rowman & Littlefield Publishers Ltd.
- Jaggar, A. M. (1996). "Ética feminista: algunos temas para los años noventa". En C. C. (comp.), (Ed.). *Perspectivas feministas en teoría política*. Barcelona, Paidós. 167-184.
- Jerez Mir, M. (1992). *Business and politics in Spain: from francoism to democracy*. Barcelona, Institut de Ciències Polítiques i Socials, Working Papers, 52.
- Jerez Mir, M. (1999). *Ciencia Política, un balance de fin de siglo*. Madrid, Centro de Estudios Constitucionales.
- Jiménez Perona, A. (1995). "Igualdad". En C. Amorós, (Ed.). *10 Palabras Clave sobre Mujer*. Estella, Verbo Divino. 119-150.
- Joint Committee on Standards for Educational Evaluation (1994). *The Program Evaluation Standards. 2nd ed. How to Assess Evaluations of Educational Programs*. Thousand Oaks, Sage.
- Joint Committee on Standards for Educational Evaluation (1988). *Normas de Evaluación para programas, proyectos y material educativo*. Mexico, Trillas.
- Jones, K. (1993). *Compassionate Authority: Democracy and the Representation of Women*. Londres, Routledge.
- King, J. (1988). "Research on Evaluation Use and Its Implications for the Improvement of Evaluation Research and Practice." **Studies in Educational Evaluation** 14: 285-99.
- King, J., L. L. Morris, et al. (1987). *How To Assess Program Implementation*. Newbury Park, CA, Sage.
- Knott, T. (1998). "A Whiz of a Way to Remember the Five Guiding Principles for Evaluators." **American Journal of Evaluation** 19(1 Winter 1998): 135-139.
- Knox, C. y. M., D. (1995). "Policy Evaluation: Incorporating Users' Views." **Public Administration** 73(Autumn 1995): 413-436.
Dificultad e interés de incorporar la perspectiva de los usuarios. Condiciones para ello.
- Lagarde, M. (1997). *Género y feminismo. Desarrollo humano y democracia*. Madrid, Horas y horas.
-

-
- Landau, M. ([1977]1996). "El ámbito propio del análisis de políticas". En L. Aguilar, (Ed.). *El estudio de las Políticas Públicas*. México, Miguel Angel Porrúa. Grupo Editorial. 275-280.
- Lasswell, H. ([1951]1996). "La orientación hacia las políticas". En L. Aguilar, (Ed.). *El estudio de las Políticas Públicas*. México, Miguel Angel Porrúa. Grupo Editorial. 79-104.
- Lasswell, H. ([1971]1996). "La concepción emergente de las ciencias de políticas". En L. Aguilar, (Ed.). *El estudio de las Políticas Públicas*. México, Miguel Angel Porrúa. Grupo Editorial. 105-118.
- Leeuw, F., R. Rist, et al., (Ed.). (1993). *Comparative Perspectives on Evaluation and Organizational Learning*. New Brunswick, NJ, Transaction.
- Lijphart, A. (1971). "Comparative Politics and the Comparative Method." **American Science Review** 65: 682-93.
- Lijphart, A. (1975). "The comparables cases strategy in comparative research." **Comparative Political Studies** 8: 157-75.
- Lincoln, Y. (1991). "The Arts and Sciences of Program Evaluation." **Evaluation Practice** 12(1): 1-7.
- Lincoln, Y. y E. Guba (1985). *Naturalistic Inquiry*. Beverly Hills, Sage.
- Lincoln, Y. y E. Guba (1986). "Research, Evaluation, and Policy Analysis: Heuristics for Disciplined Inquiry." **Policy Studies Review** 5 (February 86)(3): 546-565.
- Lindblom, C. (1991). *El proceso de elaboración de políticas públicas*. Madrid, MAP.
- Lindblom, C. ([1959] 1996). "La ciencia de "salir del paso"". En L. Aguilar, (Ed.). *La hechura de las políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 201-226.
- Lindblom, C. ([1979] 1996). "Todavía tratando de salir del paso". En L. Aguilar, (Ed.). *La hechura de las políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 227-254.
- Linz, J. y J. R. Montero (1999). *The party systems of Spain: old cleveages and new challenges*. Madrid, Instituto Juan March de Estudios e Investigaciones. Centro de Estudios Sociales Avanzados. Estudios/Working Papers, 1999/138.
- Lipsey, M. (1988). "Practice and Malpractice in Evaluation Research." **Evaluation Practice** 9(4): 5-24.
- Llera, F. (1994). *Los vascos y la política: el proceso político vasco, elecciones, partidos, opinión pública y legitimación en el País Vasco, 1977-1992*. Bilbao, Universidad del País Vasco.
-

-
- Llera, F. (1998). "Los rendimientos de los sistemas electorales de las Comunidades Autónomas: el predominio del partidismo imperfecto." **Revista Española de Investigaciones Sociológicas** 82: 127-158.
- Llera Ramo, F. (2000). "Gobierno y Administración en Euskadi". En J. L. Paniagua Soto, (Ed.). *Gobierno y Administración en las Comunidades Autónomas. Andalucía, Canarias, Cataluña, Galicia y País Vasco*. Madrid, Tecnos. 11-66.
- López, I. (2000). "Empoderamiento y *mainstreaming*: estrategias para la igualdad entre los géneros." **Revista Española de Desarrollo y Cooperación** n° 6 Primavera/Verano 2000: 59-84.
- Love, A. (1991). *Internal Evaluation: Building Organizations From Within*. Newbury Park, Sage.
- Love, A. (1994). "Should evaluators be certified?". En J. Altschuld y M. Engle, (Ed.). *The Preparation of Professional Evaluators: Issues, Perspectives, and Programs*. San Francisco, Jossey-Bass. 29-40.
- Lovenduski, J. (1997). "Feminismo institucional: Género y Estado". En E. U. y. A. Elizondo, (Ed.). *Mujeres en política. Análisis y práctica*. Barcelona, Ariel. 201-210.
- Lowi, T. ([1964] 1996). "Políticas públicas, estudios de caso y teoría política". En L. Aguilar, (Ed.). *La hechura de las políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 89-118.
- MacKinnon, C. (1995). *Hacia una Teoría feminista del Estado*. Madrid, Cátedra-Universidad de Valencia-Instituto de la Mujer. *Original en inglés en Harvard Univ. Press (1989)*.
- Madaus, G. F., M. Scriven, et al., (Ed.). (1983). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. Boston, Kluwer-Nijhoff.
- Majone, G. (1989). *Evidence, Argument, and Persuasion in the Policy Process*. New Haven, Yale University Press.
- Majone, G. (1996). "Public Policy and Administration: Ideas, Interests and Institutions". En R. Goodin and H.-D. Klingemann, (Ed.). *A new Handbook of Political Science*. Nueva York, Oxford University Press. 610- 627.
- Majone, G. ([1975] 1996). "La factibilidad de las políticas sociales". En L. Aguilar, (Ed.). *La hechura de las políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 393-432.
- Majone, G. ([1978] 1996). "Los usos del análisis de políticas". En L. Aguilar, (Ed.). *La hechura de las políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 341-366.
-

-
- Mark, M. y L. Shotland, (Ed.). (1987). *Multiple Methods in Program Evaluation*. New Directions for Program Evaluation, n.35. San Francisco, Jossey-Bass.
- Martínez Fernández, C. y M. J. Montero Corominas (1996). *¿Qué ha supuesto la Conferencia de Beijing para las Mujeres? Aproximación a los acuerdos adoptados en la IV Conferencia Mundial de Naciones Unidas sobre las Mujeres. Beijing (China), 4-15 de septiembre de 1995*. Madrid, Ministerio de Asuntos Sociales, Instituto de la Mujer.
- Martínez Hernández, E. (1997). "Políticas Públicas para la igualdad entre los sexos: reflexiones sobre el caso español (1975-1997)". En E. U. y. A. Elizondo, (Ed.). *Mujeres en política. Análisis y práctica*. Barcelona, Ariel. 211-232.
- Martínez Hernández, E. (1998). "Construyendo la agenda: reflexiones en torno a la formación de políticas públicas para la igualdad entre los sexos." **Inguarak** 22(Diciembre 1998): 149-162.
- Mas Sabaté, J. y C. Ramió Matas (1997). *La auditoría operativa en la práctica : Técnicas de mejora organizativa*. Barcelona, Marcombo.
- Mathison, S. (1999). "Rights, Responsibilities, and Duties: A Comparison of Ethics for Internal and External Evaluators". En J. L. Fitzpatrick and M. Morris, (Ed.). *Current and Emerging Ethical Challenges in Evaluation*. San Francisco, Jossey-Bass. 25-34.
- May, P. J. ([1981] 1996). "Claves para diseñar opciones de políticas". En L. Aguilar, (Ed.). *Problemas públicos y agenda de gobierno*. México, Miguel Angel Porrúa. Grupo Editorial. 235-256.
- McCraen, G. (1988). *The Long Interview*. Newbury Park, CA, Sage.
- McGraw, K. (1996). "Political Methodology: Research Design and Experimental Methods". En R. Goodin y H.-D. Klingemann, (Ed.). *A new Handbook of Political Science*. Nueva York, Oxford University Press. 769-786.
- McIntosh, M. (1978). "The State and the oppression of women". En A. Kuhn y A. Wolpe, (Ed.). *Feminism and materialism: Women and the modes of production*. Londres, Routledge and Kegan Paul. 254-289.
- Meehan, E. (1993). *Citizenship in the European Community*. Londres, Sage.
- Meehan, E. y S. Sevenhuijsen, (Ed.). (1991). *Equality Politics and Gender*. London, Sage.
- Meer, F. B. v. d. (1999). "Evaluation and the Social Construction of Impacts." **Evaluation** 5(4): 387-406.
- Meier, P. (1998). *Policy innovation in terms of mainstreaming? The development of Gender Impact Assessment in Flanders, Belgium*. ECPR 26th Joint Sessions of
-

-
- Workshops. Workshop on "Gender Innovations in Public Policy: Generalizability, Impact and Effectiveness", Warwick (UK).
- Meltsner, A. ([1972] 1996). "La factibilidad política y el análisis de políticas". En L. Aguilar, (Ed.). *La hechura de las políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 367-392.
- Meny, I. y J. C. Thoenig (1992). *Las políticas públicas*. Barcelona, Ariel.
- Mertens, D. M. (1994). "Training Evaluators: Unique Skills and Knowledge". En J. Altschuld y M. Engle, (Ed.). *The Preparation of Professional Evaluators: Issues, Perspectives, and Programs*. San Francisco, Jossey-Bass. 17-28.
- Michavila, F., (Ed.). (1998). *Experiencias y consecuencias de la evaluación universitaria*. Madrid, Fundación Universidad-Empresa. Consejería de Educación y Cultura. Embajada de los EE.UU. de América.
- Miguel, A. D. (1995). "Feminismos". En C. Amorós, (Ed.). *10 Palabras Clave sobre Mujer*. Estella, Verbo Divino. 217-255.
- Miles, M. B. y A. M. Huberman (1994). *Qualitative Data Analysis*. Thousand Oaks, Sage.
- MinisterioAsuntosExteriores (1998). *Metodología de evaluación de la Cooperación Española*. Madrid, Ministerio de Asuntos Exteriores. SECIPI.
- MinisterioAsuntosExteriores (2001). *Metodología de evaluación de la Cooperación Española II*. Madrid, Ministerio de Asuntos Exteriores. SECIPI.
- Monnier, E. (1990-91). "Objetivos y destinatarios de las evaluaciones." **Documentación Administrativa** 224-225: 131-152.
- Monnier, E. (1995). *Evaluación de la acción de los poderes públicos*. Madrid, Instituto de Estudios Fiscales. Ministerio de Economía y Hacienda.
- Montabes Pereira, J. y F. Olvera Porcel (2000). "Gobierno y Administración en la Comunidad Autónoma de Andalucía". En J. L. Paniagua Soto, (Ed.). *Gobierno y Administración en las Comunidades Autónomas. Andalucía, Canarias, Cataluña, Galicia y País Vasco*. Madrid, Tecnos. 67-127.
- Moore, M. H. ([1976] 1996). "Anatomía del problema de la heroína: un ejercicio de definición de problemas". En L. Aguilar, (Ed.). *Problemas públicos y agenda de gobierno*. México, Miguel Angel Porrúa. Grupo Editorial. 189-218.
- Morata, F. (1991). "Políticas públicas y relaciones intergubernamentales." **Documentación Administrativa** 224-225: 153-166.
- Moreno, L. (1997). *La federalización de España*. Madrid, Siglo XXI.
-

- Morris, L., (Ed.). (1978). *Program Evaluation Kit*. Beverly Hills, Sage.
- Morris, L. L. y C. T. Fitz-Gibbon (1978). *How To Deal With Goals and Objectives*. Newbury Park.
- Morris, M. (1999). "Research on Evaluation Ethics: What Have We Learned and Why Is It Important?". En J. L. Fitzpatrick y M. Morris, (Ed.). *Current and Emerging Ethical Challenges in Evaluation*. San Francisco, Jossey-Bass. 15-24.
- Morris, M. y L. R. Jacobs (2000). "You Got A Problem With That? Exploring Evaluators' Disagreements About Ethics." **Evaluation Review** 24(4. August 2000): 384-406.
- Navarro, A. (1990). "Eficacia de las terapias psicológicas: Meta-análisis de estudios en países de habla hispana." **Revista de Psicología Social y Personalidad** 6: 87-89.
- Nelson, B. ([1978] 1996). "La formación de una agenda. El caso del maltrato a los niños". En L. Aguilar, (Ed.). *Problemas públicos y agenda de gobierno*. México, Miguel Angel Porrúa. Grupo Editorial. 105-140.
- Nelson, B. J. (1996). "Public Policy and Administration: An Overview". En R. Goodin y H.-D. Klingemann, (Ed.). *A new Handbook of Political Science*. Nueva York, Oxford University Press. 551-592.
- Newman, D. (1999). "Education and Training in Evaluation Ethics". En J. L. Fitzpatrick y M. Morris, (Ed.). *Current and Emerging Ethical Challenges in Evaluation*. San Francisco, Jossey-Bass. 67-76.
- Nigro, F. A. y N., L.G. (1989). "Program evaluation". En F. A. y N. Nigro, L.G., (Ed.). *Modern Public Administration*. Modern Public Administration Nueva York, Harper Collins Pub. 194-210.
Muy interesante las propuestas para la mejora de la evaluación
- NORAD (1997). *Evaluación de proyectos de ayuda al desarrollo. Manual para evaluadores y gestores*. Madrid, IUDC-UCM y CEDEAL.
- O'Toole, L. J. ([1989] 1996). "Recomendaciones prácticas para la implementación de las políticas que involucran a múltiples actores: una evaluación de campo". En L. Aguilar, (Ed.). *La implementación de las Políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 413-470.
- Olias de Lima Gete, B., (Ed.). (2001). *La nueva gestión pública*. Madrid, Prentice Hall.
- ONU (1996). *Declaración de Beijing y Plataforma para la Acción*. Madrid, Ministerio de Trabajo y Asuntos Sociales. Instituto de la Mujer.
- ONU (1999). *Convención sobre la eliminación de todas las formas de discriminación contra la mujer y el protocolo opcional a la convención*. Madrid, Ministerio de Trabajo y Asuntos Sociales. Instituto de la Mujer.

-
- Osborne, D. y T. Gaebler (1994). *La reinención del gobierno*. Barcelona, Paidós.
- Osborne, R. (1995). "Acción Positiva". En C. Amorós, (Ed.). *10 Palabras Clave sobre Mujer*. Estella, Verbo Divino. 297-330.
- Osuna, J. L. y C. Márquez (2000). *Guía para la evaluación de políticas públicas*. Sevilla, Instituto de Desarrollo Regional. Fundación Universitaria.
- Outshoorn (1991). "Is This What We Wanted? Positive Action as Issue Perversion". En E. Meehan y S. Sevenhuijsen, (Ed.). *Equality Politics and Gender* London, Sage. 104-121.
- Palumbo, D. J., (Ed.). (1987). *The Politics of Program Evaluation*. Newbury Park, Sage.
- Paniagua Soto, J. L., (Ed.). (2000). *Gobierno y Administración en las Comunidades Autónomas. Andalucía, Canarias, Cataluña, Galicia y País Vasco*. Madrid, Tecnos.
- Parsons, W. (1995). *Public Policy. An Introduction to the Theory and Practice of Policy Analysis*. Cheltenham, UK, Edward Elgar.
- Parvikko, T. (1991). "Conceptions of Gender Equality: Similarity and Difference". En E. Meehan and S. Sevenhuijsen, (Ed.). *Equality Politics and Gender*. London, Sage. 36-51.
- Pateman, C. (1996). "Críticas feministas a la dicotomía público/privado". En C. C. (comp.), (Ed.). *Perspectivas feministas en teoría política*. Barcelona, Paidós. 31-52.
- Patton, M. Q. (1980). *Qualitative Evaluation Methods*. Beverly Hills, Sage.
- Patton, M. Q. (1981). *Creative Evaluation*. Beverly Hills, Sage.
- Patton, M. Q. (1982). *Practical Evaluation*. Beverly Hills, Sage.
- Patton, M. Q. (1985). *Culture and Evaluation*. San Francisco, Jossey-Bass.
- Patton, M. Q. (1990). *Qualitative Evaluation and Research Methods*. Newbury Park, Sage.
- Patton, M. Q. (1994a). "Developmental Evaluation." **Evaluation Practice** 15(3): 311-20.
- Patton, M. Q. (1994b). "The Program Evaluation Standards Reviewed." **Evaluation Practice** 15(2): 193-99.
-

-
- Patton, M. Q. (1996). *A world Larger Than Formative and Summative*. San Francisco, Jossey-Bass.
- Patton, M. Q. (1997). *Utilization-Focused Evaluation. The New Century Text. 3rd. ed.* Thousand Oaks, Sage.
- Payne, D. A. (1994). *Designing Educational Project and Program Evaluations*. Boston, Kluwer Academic Publishers.
- Pérez del Río, T. (1996). "La acción posistiva. Justificación jurídico-consitucional." **Jóvenes** 4-5: 82-91.
- Pérez Sánchez, M. (1997). *Formación y empleo en España : análisis y evaluación de la política de formación ocupacional*. Granada, Universidad de Granada.
- Pollit, C. (1993). *El gerencialismo y los servicios públicos. La experiencia anglo-americana*. Madrid, Instituto de Estudios Fiscales.
- Pollit, C. y G. Bouckaert (1995). *Quality Improvement in European Public Services*. Londres, Sage.
- Posavac, E. J. y R. G. Carey (1980). *Program Evaluation. Methods and Case Studies*. Englewood Cliffs, Prentice-Hall.
- Preskill, H. y R. T. Torres (1999). "Bulding Capacity for Organizational Learning Through Evaluative Inquiry." **Evaluation** 5(1): 42-60.
- Preskill, H. y R. T. Torres (2000). "The Learning Dimension of Evaluation Use". En V. J. Caracelli y H. Preskill, (Ed.). *The Expanding Scope of Evaluation Use*. The Expanding Scope of Evaluation Use San Francisco, Jossey-Bass. 25-38.
- Pressman, J. y A. Wildavsky (1984). *Implementation*. Berkeley, CA., University of California Press.
- Putt, A. D. y J. F. Springer (1989). *Policy Research. Concepts, Methods, and Applications*. Englewood Cliffs, Prentice Hall.
- Quade, E. J. (1989). *Análisis de formación de decisiones políticas*. Madrid, Instituto de Estudios Fiscales.
- Quade, E. S. (1989). *Análisis de formación de las decisiones políticas*. Madrid, Instituto de Estudios Fiscales.
- Ragin, C. C., D. Berg-Schlosser, et al. (1996). "Political Methodology: Qualitative Methods". En R. Goodin y H.-D. Klingemann, (Ed.). *A new Handbook of Political Science*. Nueva York, Oxford University Press. 749-768.
- Ramió, C. (1999). *Teoría de la organización y administración pública*. Madrid, Tecnos, D.L.
-

-
- Regonini, G. (1991). "El estudio de las políticas públicas." **Documentación Administrativa** 224-225(Oct.90-Mar.91): 59-88.
- Reichart, C. S. y S. F. Rallis, (Ed.). (1994). *The Qualitative-Quantitative Debate: New Perspectives*. New Directions for Program Evaluation, n.61. San Francisco, Jossey-Bass.
- Rein, M. y F. Rabinovitz ([1978] 1996). "La implementación: una perspectiva teórica. Entre la intención y la acción." En L. Aguilar, (Ed.). *La implementación de las Políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 147-184.
- Restrepo, P. (1990). *Evaluación del trabajo y resultados (Aproximación a los Servicios Sociales)*. Madrid, Comunidad de Madrid. Consejería de Integración Social. Dirección Gnal. de Planificación de los Servicios Sociales.
- Rist, R. C., (Ed.). (1990). *Program Evaluation and the Management of Government. Patterns and Prospects across Eight Nations*. New Brunswick, Transaction Publishers.
- Rittel, H. W. y M. M. Webber ([1973] 1996). "Dilemas de una teoría general de planeación". En L. Aguilar, (Ed.). *Problemas públicos y agenda de gobierno*. México, Miguel Angel Porrúa. Grupo Editorial. 161-188.
- Rog, D. y D. Fournier, (Ed.). (1997). *Progress and Future Directions in Evaluation: Perspectives on Theory, Practice, and Methods*. New Directions for Program Evaluation. San Francisco, Jossey-Bass.
- Rogers, P., T. Hacsí, et al., (Ed.). (2000). *Program Theory in Evaluation: Challenges and Opportunities*. New Directions for Program Evaluation, n. 87. San Francisco, Jossey-Bass.
- Rose, R. (1984). *Understanding Big Government. The Programme Approach*. Londres, Sage.
- Rossi, P. H. F. H. E. (1989). *Evaluación. Un enfoque sistemático para programas sociales*. México, Trillas.
- Ruiz Miguel, A. (1996). "Discriminación inversa, acciones positivas e igualdad." **Jóvenes** 4-5: 92-108.
- Sabatier, P. y D. Mazmanian ([1981] 1996). "La implementación de la política pública: un marco de análisis". En L. Aguilar, (Ed.). *La implementación de las Políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 323-372.
- Saltzman, J. (1989). *Equidad y Género. Una teoría integrada de estabilidad y cambio*. Madrid, Cátedra.
-

-
- Sassoon, A. S. (1996). *Las mujeres y el Estado: los límites cambiantes de lo público y lo privado*. Madrid, Vindicación feminista. *Original en inglés en Londres: Unwin Hyman. 1987.*
- Sassoon, A. S. (1998). "Igualdad, Diferencia, Ciudadanía". En P. d. Villota, (Ed.). *Las mujeres y la ciudadanía en el umbral del siglo XXI*. Madrid, Editorial Complutense. 33-45.
- Sawer, M. (1995). "'Femocrats en Glass Towers?': The Office of the Status of Women in Australia". En D. M. Stetson y A. G. Mazur, (Ed.). *Comparative State Feminism*. Comparative State Feminism Londres, Sage. 22-39.
- Scanlon, G. M. (1990). "El movimiento feminista en España, 1900-1985: Logros y dificultades". En J. Astelarra, (Ed.). *Participación política de las mujeres*. Madrid, Centro de Investigaciones Sociológicas. 83- 100.
- Schwandt, T. A. (1989a). "The Politics of Varying Trustworthiness in Evaluation Auditing." **Evaluation Practice** 10(4): 33-40.
- Schwandt, T. A. (1989b). "Recapturing Moral Discourse in Evaluation." **Educational Researcher** 19(8): 11-16.
- Schwandt, T. A. y E. S. Halpern (1988). *Linking Auditing and Metaevaluation. Enhancing Quality in Applied Research*. Nerwbury Park, Sage.
- Scriven, M. (1967). "The methodology of evaluation." **Perspectives on Curriculum Evaluation (AERA Monograph Series in Curriculum Evaluation No.1)**: 39-83.
- Scriven, M. (1991). "Beyond formative and summative evaluation". En M. M. D. C. Philips, (Ed.). *Evaluation and education: At quarter century*. Chicago. IL, University of Chicago Press.
- Scriven, M. (1991). *Evaluation Thesaurus*. Newbury Park, Sage.
- Scriven, M. (1993). *Hard-Won Lessons in Program Evaluation*. San Francisco, Jossey-Bass.
- Scriven, M. (1994). "The Final Synthesis." **Evaluation Practice** 15(3): 367-82.
- Scriven, M. (1995). "The Logic of Evaluation and Evaluation Practice". En D. M. Fournier, (Ed.). *Reasoning in Evaluation: Inferential Links and Leaps*. San Francisco, Jossey-Bass.
- Scriven, M. (1998). "Minimalist Theory: The Least Theory That Practice Requires." **The American Journal of Evaluation** 19(1): 57-70.
- Sechrest, L., (Ed.). (1993). *Program Evaluation: A Pluralistic Enterprise*. New Directions for Program Evaluation, n. 60. San Francisco, Jossey-Bass.
-

-
- Sechrest, L. B. y A. G. Scott, (Ed.). (1993). *Understanding Causes and Generalizing About Them*. New Directions for Program Evaluation, n. 57. San Francisco, Jossey-Bass.
- SECIPI (1998). *Metodología de evaluación de la Cooperación Española*. Madrid, Ministerio de Asuntos Exteriores. Secretaría de Estado para la Cooperación Internacional y Para Iberoamérica (SECIPI).
- Shadish, W., T. Cook, et al. (1991). *Foundations of Program Evaluation. Theories of Practice*. Newbury Park, Sage.
- Shadish, W. e. a., (Ed.). (1995). *Guiding Principles for Evaluators*. New Directions for Program Evaluation. San Francisco, Jossey-Bass.
- Sierra, B. (2000). "Criterios para la evaluación con perspectiva de género." **Revista Española de Desarrollo y Cooperación** n° 6 Primavera/Verano 2000: 95-112.
- Siim, B. (1991). "Welfare State, Gender Politics and Equality Policies Women's Citizenship in the Scandinavian Welfare States". En E. Meehan y S. Sevenhuijsen, (Ed.). *Equality Politics and Gender*. London, Sage. 175-192.
- Sjoberg, G. (1983). "Politics, Ethics and Evaluation Research". En M. B. Struening y Brewer, (Ed.). *Handbook of Evaluation Research*. Handbook of Evaluation Research Beverly Hills, Sage. 65-88.
- Smith, E. R. y R. W. Tyler (1942). *Appraising and recording student progress*. Nueva York, Harper & Row.
- Smith, N. L., (Ed.). (1981). *Metaphors for Evaluation*. Beverly Hills, CA, Sage.
- Solana, Y., R. Rodríguez, et al. (1997). *La larga marcha hacia la Igualdad. IV Conferencia Mundial sobre las mujeres. Beijing 95. Evaluación del proceso en España*. Madrid, Ministerio de Trabajo y Asuntos Sociales. Instituto de la Mujer.
- Sonnichsen, R. C. (1988). "Advocacy Evaluation: A Model for Internal Evaluation Offices." **Evaluation and Program Planning** 11(2): 141-48.
- Stake, R. (1995). *The Art of Case Research*. Newbury Park, Sage.
- Stake, R. E. (1967). "The Countenance of Educational Evaluation." **Teachers College Record** 68: 523-40.
- Starling, G. (1988). *Strategies for Policymaking*. Chicago, The Dorsey Press.
- Stevens, C. J. y M. Dial, (Ed.). (1994). *Preventing the Misuse of Evaluation*. New Directions for Program Evaluation, n. 64. San Francisco, Jossey-Bass.
-

-
- Stoker, R. ([1989] 1996). "Un marco de análisis para el régimen de implementación: cooperación y reconciliación entre los imperativos federalistas". En L. Aguilar, (Ed.). *La implementación de las Políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 373-412.
- Struening, E. y M. B. Brewer, (Ed.). (1983). *Handbook of Evaluation Research*. Beverly Hills, Sage.
- Stufflebeam, D. (1994). "Empowerment evaluation, Objectivist Evaluation, and Evaluation Standards: Where the Future of Evaluation Sholud Not Go and Where It Needs to Go." **Evaluation Practice** 15(3): 321-38.
- Stufflebeam, D. (2000). "Lessons in Contracting for Evaluations." **The American Journal of Evaluation** 21(3): 293-314.
- Stufflebeam, D. y A. Shinkfield (1987). *Evaluación sistemática. Guía teórica y práctica*. Madrid, Paidós/MEC.
- Stufflebem, D. (2001). *Evaluation Models*. San Francisco, Jossey-Bass.
- Subirats, J. (1990). *Modernizing the spanish public administration or reform in disguise*. Barcelona, Institut de Ciències Polítiques i Socials.
- Subirats, J. (1992). *Análisis de políticas públicas y eficacia de la administración*. Madrid, Ministerio para las Administraciones Públicas, Instituto Nacional de Administración Pública.
- Subirats, J. (1992). *Un problema de estilo. La formación de políticas públicas en España. Exploración sobre los elementos distintivos de nuestro "policy style" (una aportación al análisis del sistema político español)*. Madrid, Centro de Estudios Constitucionales.
- Subirats, J. (1994). *Policy instruments, public deliberation and evaluation processes*. Madrid, Instituto Juan March de Estudios e Investigaciones.
- Subirats, J., (Ed.). (1999). *¿Existe sociedad civil en España? : responsabilidades colectivas y valores públicos*. Coleccion Estudios de la Fundación Encuentro ; 5. Madrid, Fundación Encuentro.
- Suchman, E. (1967). *Evaluative Research: Principles and Practice in Public Service and Social Action Programs*. New York, Russell Sage.
- Threlfall, M. (1990). "¿Patriarca, palanca, paraguas? Planteamientos feministas en torno al Estado asistencial". En J. Astelarra, (Ed.). *Participación política de las mujeres*. Madrid, Centro de Investigaciones Sociológicas. 215-233.
- Threlfall, M. (1994). "Políticas Públicas para la igualdad de sexos: luces, sombras y retos". En I. d. I. Mujer, (Ed.). *Foro Internacional "Mujer, poder político y*
-

-
- desarrollo*". Foro Internacional "Mujer, poder político y desarrollo" Madrid, Ministerio de Asuntos Sociales. 29-40.
- Torgerson, D. ([1986]1996). "Entre el conocimiento y la política: tres caras del análisis de políticas". En L. Aguilar, (Ed.). *El estudio de las Políticas Públicas*. México, Miguel Angel Porrúa. Grupo Editorial. 197-238.
- Torres, R., H. Preskill, et al. (1996). *Evaluation Strategies for Communicating and Reporting*. Thousand Oaks, Sage.
- Torres, R. T. y H. Preskill (1999). "Ethical Dimensions of Stakeholder Participation and Evaluation Use". En J. L. Fitzpatrick y M. Morris, (Ed.). *Current and Emerging Ethical Challenges in Evaluation*. San Francisco, Jossey-Bass. 57-66.
- Ungo Montenegro, U. A. (2000). *Para cambiar la vida: Política y pensamiento del feminismo en América Latina*. Panamá, Instituto de la Mujer.
- Uriarte, E. (1997). "Estudios de mujeres y política en España". En E. U. y. A. Elizondo, (Ed.). *Mujeres en Política. Análisis y Práctica*. Barcelona, Ariel. 15-32.
- Uriarte, E. y E., A., (Ed.). (1997). *Mujeres en Política. Análisis y Práctica*. Barcelona, Ariel.
- U.S. General Accounting Office (1987). *Federal Evaluations: fewer Units, reduced resources, different studies from 1980*. Washington, DC, U.S. General Accounting Office.
- Valcárcel, A., (Ed.). (1994). *El concepto de igualdad*. Madrid, Editorial Pablo Iglesias.
- Valcárcel, A. (1997). *La política de las mujeres*. Madrid, Cátedra.
- Valiente, C. (1995). "The Power of Persuasion. The *Instituto de la Mujer* in Spain". En D. M. Stetson y A. Mazur, (Ed.). *Comparative State Feminism*. Comparative State Feminism Thousand Oaks, Sage. 221-236.
- Valiente, C. (1996). "El feminismo institucional en España: El Instituto de la Mujer, 1983-1994." **Revista Internacional de Sociología** 13(Enero-Abril 1996): 163-204.
- Valiente, C. (1997). "State feminism and gender equality policies. The case of Spain (1983-95)". En F. Gardiner, (Ed.). *Sex Equality Policy in Western Europe*. Sex Equality Policy in Western Europe London, Routledge. 127- 141.
- Valiente, C. (1998a). "Género, mercado de trabajo y Estado de bienestar: el caso de España." **Sociología del Tarabajo Nueva época, núm. 32**(Invierno 1997/1998): 53-79.
- Valiente, C. (1998b). "An overview of the state of research on women and politics in Spain." **European Journal of Political Research** 33(1998): 459-474.
-

-
- Valiente, C. (1999a). "La importancia de la política simbólica: la prohibición del acoso sexual en el trabajo en España". En M. Ortega, C. Sánchez y C. Valiente, (Ed.). *Género y ciudadanía. Revisiones desde el ámbito privado*. Madrid, Ediciones de la Universidad Autónoma de Madrid. 485-508.
- Valiente, C. (1999b). "El feminismo de Estado en los ayuntamientos de la Comunidad Autónoma de Madrid: creación y actuaciones." **Gestión y Análisis de Políticas Públicas 13-14**(Septiembre 1998/Abril 1999).
- Valiente, C. (1999c). "El feminismo de Estado y los debates políticos: la formación ocupacional en España (1983-1998)." **Revista Española de Ciencia Política 1**(1): 127-147.
- Valiente, C. (2000). "Género y ciudadanía: los organismos de igualdad y el Estado del Bienestar en España". En M. Pérez Ledesma, (Ed.). *Ciudadanía y Democracia*. Ciudadanía y Democracia Madrid, Editorial Pablo Iglesias. 199-230.
- Van Meter, D. y C. Van Horn ([1975] 1996). "El proceso de implementación de las políticas. Un marco conceptual". En L. Aguilar, (Ed.). *La implementación de las Políticas*. México, Miguel Angel Porrúa. Grupo Editorial. 97-146.
- Vaux, A., M. S. Stockdale, et al., (Ed.). (1992). *Independent Consulting for Evaluators*. Newbury Park, CA, Sage.
- Vedung, E. (1997). *Evaluación de Políticas Públicas y Programas*. Madrid, Ministerio de Trabajo y Asuntos Sociales. Instituto Nacional de Servicios Sociales.
- Verloo, M. (1998). *Policy design and frame alignment: the case of mainstreaming*. ECPR 26th Joint Sessions of Workshops. Workshop on "Gender Innovations in Public Policy: Generalizability, Impact and Effectiveness", Warwick (UK).
- Vilas Nogueira, J. (2000). "Gobierno y Administración en la Comunidad Autónoma de Galicia". En J. L. Paniagua Soto, (Ed.). *Gobierno y Administración en las Comunidades Autónomas. Andalucía, Canarias, Cataluña, Galicia y País Vasco*. Madrid, Tecnos. 129-167.
- Villota, P. d., (Ed.). (1998). *Las mujeres y la ciudadanía en el umbral del siglo XXI*. Madrid, Editorial Complutense.
- Weiss, C. H. (1983). "Evaluation Research in the Political Context". En M. B. Struening y Brewer, (Ed.). *Handbook of Evaluation Research*. Handbook of Evaluation Research Beverly Hills, Sage. 31-45.
- Weiss, C. H. (1987). "Where Politics and Evaluation Research Meet". En D. J. Palumbo, (Ed.). *The Politics of Program Evaluation*. The Politics of Program Evaluation Newbury Park, Sage. 47-69.
-

-
- Weiss, C. H. (1988). "Evaluation for Decisions: Is Anybody There? Does Anybody Care?" **Evaluation Practice** 9(1): 5-19.
- Weiss, C. H. (1997). "Theory-Based Evaluation: Past, Present, and Future". En D. RogD.yFournier, (Ed.). *Progress and Future Directions in Evaluation: Perspectives on Theory, Practice, and Methods*. New Directions for Program Evaluation, 76 San Francisco, Jossey-Bass.
- Weiss, C. H. (1998). *Evaluation. Methods for Studying Programs and Policies. Second Edition*. Upper Saddle River (NJ), Prentice Hall.
- Weiss, C. H. (1998b). "Have We Learned Anything new About the Use of Evaluation?" **American Journal of Evaluation** 19(1): 21-33.
- Whitemore, E., (Ed.). (1998). *Understanding and Practicing Participatory Evaluation*. New Directions for Program Evaluation, n. 80. San Francisco, Jossey-Bass.
- Wholey, J., (Ed.). (1983). *Evaluation and Effective Public Management*. Boston, Little, Brown and Company.
- Wholey, J., H. Hatry, et al., (Ed.). (1994). *Handbook of Practical Program Evaluation*. San Francisco, Jossey-Bass.
- Wildavsky, A. (1979). *Speaking Truth to Power*. Boston, Little Brown.
- Wildavsky, A. (1985). "The Self-Evaluating Organization". En E. Chelimsky, (Ed.). *Program Evaluation: Patterns and Directions*. Washington, DC, American Society for Public Administration.
- Wilson, E. (1977). *Women and the Welfare State*. Londres, Tavistock.
- Wisler, C., (Ed.). (1996). *Evaluating and Auditing: Prospects for Convergence*. New Directions for Program Evaluation, n. 71. San Francisco, Jossey-Bass.
- Woodward, A. E. (1998). "El Estado y la ciudadanía ¿Quién constituye el Estado? ¿Qué lugar ocupa la mujer?". En P. d. Villota, (Ed.). *Las mujeres y la ciudadanía en el umbral del siglo XXI*. Madrid, Editorial Complutense. 47-63.
- Worthen, B. R. (1994). "Is Evaluationa Mature Profession That Warrants the Preparation of Evaluation Professionals?". En J. Altschuld y M. Engle, (Ed.). *The Preparation of Professional Evaluators: Issues, Perspectives, and Programs*. San Francisco, Jossey-Bass. 3-16.
- Worthen, B. R. y K. R. White (1987). *Evaluating Educational and Social Programs. Guidelines for Proposal Review, Onsite Evaluation, Evaluation Contracts and Technical Assistance*. Boston, Kluwer-Nijhoff Publishing.
- Yin, R. K. (1997). "Case Study Evaluations: A Decade of Progress?". En D. Rog, y Fournier, (Ed.). *Progress and Future Directions in Evaluation: Perspectives on*
-

Theory, Practice, and Methods. New Directions for Evaluation. Progress and Future Directions in Evaluation: Perspectives on Theory, Practice, and Methods. New Directions for Evaluation San Francisco, Jossey-Bass. 69-78.

ANEXO A

DOCUMENTACIÓN CONSULTADA

DOCUMENTACIÓN GENERAL

- 4º Programa de Acción Comunitaria a Medio Plazo para la Igualdad de Oportunidades entre Hombres y Mujeres (1996-2000). *Documento de la Comisión Europea. Dirección Gnal. V (Empleo, Relaciones Laborales y Asuntos Sociales. Unidad para la Igualdad de Oportunidades)*

DEL INSTITUTO DE LA MUJER

1. I Plan de Igualdad de Oportunidades entre Mujeres y Hombres 1988-90
2. II Plan de Igualdad de Oportunidades de las Mujeres 1993-95
3. III Plan de Igualdad de Oportunidades entre Mujeres y Hombres 1997-2000
4. I Plan. Evaluación
5. Síntesis de resultados de la evaluación del II PIOM. 7 Marzo 1997. Documento
6. Síntesis de Resultados de la Evaluación del II Plan de Igualdad de Oportunidades de las MUJeres (1993-1995). Instituto de la Mujer. Madrid, 1997. Publicación.
7. “La evaluación de los planes de igualdad de oportunidades entre hombres y mujeres (PIOM) del Instituto de la Mujer: referencia obligada al segundo de ellos” Mariano Alvaro Page (Consejero Técnico de Estudios del Instituto de la Mujer), Miguel de la Fuente Sánchez (Director de Estudios de CUANTER, S.A.) y M^a Teresa Aranda Jaquotot (Técnica de Estudios de CAUNTER, S.A.). Ponencia presentada al seminario sobre *Evaluación de Políticas de Igualdad de Oportunidades entre Mujeres y Hombres*, celebrado en Santander los días 29, 30 de junio y 1 de Julio de 1998, en el marco de los cursos organizados por la Universidad Internacional Menéndez Pelayo.
8. Carta para la Igualdad de Oportunidades de las Mujeres en el ámbito rural. Las mujeres protagonistas en el desarrollo rural 1994-1995. 1995.
9. 10 años del Instituto de la Mujer. Publicación de 1994.
10. Entrevista a Concha Dancausa en “8 de Marzo”, revista de la CAM, nº 27, 1997
11. Instituto de la Mujer. Ley de Creación (Mº de Cultura)
12. Beijing 1995. IV Conferencia Mundial sobre las Mujeres. 1995.
13. ¿Qué ha supuesto la Conferencia de Beijing para las Mujeres? Aproximación a los acuerdos adoptados en la IV Conferencia Mundial de Naciones Unidas sobre las Mujeres. 1996.
14. Revista “Mujeres”: nº 21 (1er trimestre 96), nº 23 (3er trimestre 96).
15. La Mujer en cifras. 1987
16. La Mujer en cifras. 1992
17. La Mujer en cifras. 1997

ANDALUCÍA

1. Plan de Igualdad de Oportunidades de las Mujeres 1990-92. Publicación. 70 pp.
2. “Las Mujeres en Andalucía. Estrategias para avanzar”. II Plan Andaluz para la Igualdad de las Mujeres 1995-1997.
3. Plan de Igualdad de Oportunidades de Andalucía: Avances y Resultados del Plan. *Publicación. 14 pp.*
4. Informe de Evaluación del Plan para la Igualdad de Oportunidades 1990-93. Marzo de 1994.
5. Anexo al Informe
6. Proyecto de Evaluación del II Plan de Igualdad de Oportunidades de las Mujeres, 1995-1996. Instituto de Estudios Sociales Avanzados -Andalucía- del CSIC (Córdoba). *Fotocopias. 11 pp.*
7. II Plan Andaluz para la Igualdad de las Mujeres. Avance del Informe de Evaluación. Julio 1997 Instituto de Estudios Sociales Avanzados de Andalucía. Consejo Superior de Investigaciones Científicas. *Documento. 68 pp.*
8. Indicadores del Plan *Documento interno.*
9. Ley 10/98, de 29 de Diciembre, de presupuesto de la Comunidad Autónoma de Andalucía para 1989, en la que se crea el Instituto Andaluz de la Mujer. (BOJA nº 106, 30/12/88 y BOE nº 33, 8/2/89)
10. Decreto 1/1989, de 10 de Enero, por el que se aprueba el reglamento del Instituto Andaluz de la Mujer. (BOJA nº 4, 17/1/89)
11. Decreto 120/1997, de 22 de Abril, por el que se modifica el Reglamento del Instituto Andaluz de la Mujer, aprobado por el Decreto 1/89, de 10 de Enero (BOJA nº 49, 26/4/97)
12. Cuadro de presupuestos Junta de Andalucía e Instituto Andaluz de la Mujer para 1995-1998.
13. Cuadro presupuesto del II Plan de Igualdad correspondiente al IAM y a otros organismos
14. Presupuestos II Plan de Igualdad
15. Granados Vaquero, Elena (1997) “El Instituto Andaluz de la Mujer: Organización, presupuesto y actuaciones” Ponencia presentada en el III Congreso de la AECPA en Salamanca (2-4 de Octubre de 1997).

ARAGÓN

1. Plan de Acción Positiva para las mujeres en Aragón 94-96 (aprobado 21/11/94). *Documento fotocopiado. Diciembre de 1994*
2. Plan de Acción Positiva para las mujeres en Aragón. Volumen I. *Documento fotocopiado. Junio 1994.* Documento introductorio y general de justificación del Plan (concepto de acción positiva, planes, marco jurídico,, demografía y datos sobre mujeres y recursos).
3. “Acortando distancias”. Memoria 1996 (Instituto Aragonés de la Mujer)
4. La Mujer Aragonesa en cifras. 1997.
5. II Plan de Acción Positiva para las Mujeres en Aragón. Documento para el debate. Junio de 1997.

6. II Plan de Acción Positiva para las Mujeres en Aragón. Publicación. 64 pp.
7. Informe sobre el grado de cumplimiento del Plan de Acción Positiva para las mujeres de Aragón. 1994-1996. Enero de 1997.
8. 2º Plan de Acción Positiva para las Mujeres en Aragón 1997-2000. Proceso de elaboración. Octubre 1997. *Documento fotocopiado*.
9. Ficha de evaluación de las acciones del Plan de Acción Positiva para las mujeres en Aragón 1994-96

CANARIAS

1. I Plan de Igualdad de Oportunidades 1995-1996. Instituto canario de la Mujer. *Documento interno 90 pp.*
2. II Plan Canario de Igualdad de Oportunidades de las Mujeres 1997-2000. (Documento para el debate. 8 de Julio de 1997). *70 pp.*
3. II Plan Canario de Igualdad de Oportunidades de las Mujeres 1997-2000. Publicación. 89 pp.
4. Auditoría del I Plan de Igualdad de oportunidades para las mujeres canarias 1995-1996. Instituto canario de la Mujer. Realizado por Asesores en Imagen Corporativa S.A. Valencia, 20-10-96. *265pp.*
5. Memoria del Instituto Canario de la Mujer Año 1996. *Documento interno. 114 pp.*
6. Ley 1/1994 de 13 de Enero, sobre creación del Instituto Canario de la Mujer
7. Decreto 1/1995, de 13 de Enero, por el que se desarrolla la Ley...
8. Folleto: "¿Qué es el Instituto Canario de la Mujer?" Gobierno de Canarias. Consejería de Empleo y Asuntos Sociales.

CASTILLA-LEÓN

1. Plan Integral de Igualdad de Oportunidades para la Mujer en Castilla y León (1994-1996). *Publicación 236 pp. 1994.*
2. Orden de 28 de Febrero de 1996, de la Consejería de Bienestar Social, por la que se crea el Consejo Regional de la Mujer y orden de 5 de Julio de 1996 por la que se modifica la del 28 de Febrero. *Fotocopia*
3. Resumen de la Evaluación del Plan Integral de igualdad de Oportunidades para la Mujer en Castilla y León 1994-1996. *Documento interno . 34 pp. y 44 tablas.*
4. Anteproyecto del II Plan de Igualdad de Oportunidades para la Mujer en Castilla-León (1997-2000). *Fotocopia. Junio 1997.*
5. Folleto sobre el II Plan Integral de Igualdad de Oportunidades para la Mujer en Castilla y León
6. Anteproyecto para la Evaluación del Plan Integral de Igualdad para las Mujeres en Castilla y León (1994-1996). Eurodoxa, s.a. *Fotocopia 14 pp.*
7. Resumen de la Evaluación del Plan Integral de Igualdad de Oportunidades para la Mujer en Castilla y León 1994-1996. Junta de Castilla y León. Consejería de Sanidad y Bienestar Social. Secretaría Regional de la Mujer. *Documento entregado en el Seminario sobre Evaluación de Políticas de Igualdad de Oportunidades entre Mujeres y Hombres, celebrado en Santander los días 29, 30 de junio y 1 de Julio de*

1998, en el marco de los cursos organizados por la Universidad Internacional Menéndez Pelayo.

CASTILLA-LA MANCHA

1. 2º Plan de Igualdad de Oportunidades. Las Mujeres de Castilla-La Mancha hacia el S. XXI (1995-1999). *Publicación 41 pp. 1995.*
2. Decreto 35/1990, de 13 de Marzo, por el que se crea el Consejo Regional de la Mujer de Castilla-La Mancha
3. Decreto 79/1995, de 29 de Agosto por el que se establece la estructura de la Consejería de Bienestar Social de la Junta de Comunidades de Castilla-La Mancha
4. Datos globales de evolución del presupuesto. Ejercicio 1989-1998. Programa “Promoción de la Mujer”
5. Organigrama Dirección General
6. Informe previo de evaluación del II Plan de Igualdad de Oportunidades “Las Mujeres de Castilla-La Mancha hacia el siglo XXI”. Borrador. 84 pp.
7. Proyecto de evaluación del II Plan de Igualdad de Oportunidades. Las mujeres de Castilla-La Mancha hacia el siglo XXI”. Octubre de 1998. Fundación Mujeres
8. Primer Instrumento para recogida de datos. Evaluación del II Plan de Igualdad de Oportunidades: “Las Mujeres de Castilla-La Mancha hacia el siglo XXI”. WG, Consultoras para el Desarrollo. 3 de Diciembre de 1998.
9. III Plan para la Igualdad de Oportunidades entre Hombres y Mujeres de Castilla-La Mancha (1999-2003). Un compromiso con nuestro futuro. Dirección General de la Mujer. Consejería de Bienestar Social. Junta de Comunidades de Castilla-La Mancha.

CATALUÑA

1. Plan de Actuación del Gobierno de la Generalitat de Cataluña para la Igualdad de Oportunidades para las Mujeres **1989-1992**. *Edición trilingüe de la Comisión Interdepartamental de Promoción de la Mujer, Departamento de Presidencia. 1989.*
2. Plan de Actuación del Gobierno de la Generalitat de Cataluña para la Igualdad de Oportunidades para las Mujeres **1994-1996**. *Edición cuatrilingüe del ICD y de la Comisión Interdepartamental para la Igualdad de Oportunidades para las Mujeres. Departamento de Presidencia. 1994*
3. Decreto 25/87 del 29 de Enero, sobre la creación de la Comisión Interdepartamental de la Mujer
4. Ley 11/89 del 10 de Julio de creación del Instituto Catalán de la Mujer
5. Decreto 209/1993 del 31 de Agosto por el que se establece la estructura orgánica del organismo autónomo administrativo Instituto Catalán de la Mujer y se modifican los decretos que lo regulan y la organización y nomenclatura de la Comisión Interdepartamental de Promoción de la Mujer

6. Comparecencia de la presidenta del Instituto Catalán de la Mujer en la Comisión de Seguimiento del proceso de equiparación Hombre-Mujer (Parlamento de Cataluña, 6 de Mayo de 1997). *Comparecencia sobre la evaluación del II Plan*
7. Informe d'avaluació del II Pla d'actuació del Govern per a la igualtat d'oportunitats per a les dones. Versió revisada 25/4/1997. 242 pp.
8. Políticas de los Ayuntamientos de Cataluña a favor de las Mujeres. *Publicación del ICD (colección estudios)*

EXTREMADURA

1. Copia del Plan para la Igualdad de Oportunidades de las Mujeres de Extremadura.
2. Evaluación del Plan para la Igualdad de Oportunidades de las Mujeres de Extremadura. Informe interno
3. Folleto sobre el SERFIDEM (Servicio de Dinamización, Información y Formación para el Empleo de la Mujer). Dirección General de la Mujer.

GALICIA

1. (I) Plan para la Igualdad de Oportunidades de las Mujeres Gallegas 1992-1994. *Publicación bilingüe. 34 pp. por 2. 1992.*
2. II Plan para la Igualdad de Oportunidades de las Mujeres Gallegas 1995-1997. *Publicación bilingüe. 54 pp. por 2. 1995.*
3. Tríptico sobre el II Plan.
4. Folleto sobre el Servicio Galego de igualdade (Servicios que presta)
5. Tríptico fotocopiado sobre "Evaluación del II Plan de Igualdad de Oportunidades para las Mujeres Gallegas. Información"
6. Diagrama 1.1: Relación entre planificación/intervención y evaluación
7. Composición de la Comisión Técnica de seguimiento del plan (con nombres borrados)
8. Informe de evaluación del II Plan para la Igualdad de Oportunidades de las Mujeres Gallegas 1995-1997 (Amparo Almarcha). Marzo 1998.
9. Ficha de seguimiento de acciones
10. Evaluación. II Plan de Igualdad de Oportunidades de las Mujeres Gallegas. Seminario de Evaluación de Políticas de Igualdad entre mujeres y hombres. Intervención de M^aCarmen Girón Daviña (Servicio Galego de Igualdade. Xunta de Galicia).
11. Diseño de propuesta de indicadores para el II Plan de Igualdad de Oportunidades para las Mujeres Gallegas". Avance 1^a fase. Amparo Almarcha y Benjamín González. La Coruña, Noviembre de 1996.
12. Diversas fichas utilizadas en el proceso de la evaluación: "Principais fontes de información para as 5 areas", "Doc.1 Información de base sobre Asociaciones de Mujeres" "Doc. 2. Información de base sobre educación y cultura", "Doc. 3. Información de base sobre formación y empleo", "Doc. 4. Información de base sobre ordenamiento jurídico", "Doc. 5. Información de base sobre servicios sociales y salud", "Ficha-guía para a recollida de información", "Ficha de visita",

“Valoración global de los temas relativos al Plan de Igualdad, por Areas Temáticas”.

13. Memoria 1991-94. Servicio Galego de Promoción da Igualdade do Home e da Muller.
14. Memoria 1995. Promoción da Muller.
15. Memoria 1996. Servicio Galego de Promoción da Igualdade do Home e da Muller.

MURCIA

1. Plan de Igualdad de Oportunidades para las Mujeres 1993-1995. Región de Murcia. *Editado por la Dirección general de la Mujer. Consejería de Asuntos Sociales. 1993.*
2. II Plan de Igualdad de Oportunidades entre Hombres y Mujeres 1997-2000. Región de Murcia. *Editado por la Dirección General de política Social y Familia. Consejería de Sanidad y Política Social. 1997.*

LA RIOJA

1. • Primer Plan de Igualdad de Oportunidades de las Mujeres en la Rioja 1991-95. *Editado por el Centro Asesor de la Mujer. 1991.*
2. • Plan Integral de la Mujer 1996-1999. *Elaborado por el CAM y editado por la Direc. Gnal.*
3. • Listado de indicadores de evaluación para el Plan Integral de la Mujer (1996-99). *Documento interno de la Direccion Gnal.*
4. • Ficha de temporalización de objetivos
5. • Ficha de evaluación de medidas y acciones
6. • Instrucciones para la cumplimentación de los documentos de evaluación (1 folio): “Criterios a seguir para la cumplimentación de los documentos de seguimiento y evaluación de los planes integrales de personas con discapacidad y de mujer (1996-99)”.

PAÍS VASCO

1. Ley sobre creación del Instituto Vasco de la Mujer. *Publicación bilingüe (27 pp.)*
2. (I) Plan de Acción Positiva para las mujeres en la Comunidad Autónoma de Euskadi. 1991-1994. *Publicación bilingüe (122pp. por 2)*
3. Propuesta de Acuerdo por el que se aprueba el II Plan de Acción Positiva para las mujeres en la Comunidad Autónoma de Euskadi (29/12/95). *Fotocopia. 2pp.*
4. II Plan de Acción Positiva para las mujeres en la Comunidad Autónoma de Euskadi. *Publicación bilingüe (158 pp. por 2)*
5. II Plan de Acción Positiva para las mujeres en la Comunidad Autónoma de Euskadi: CATEGORIAS. *Documento interno. 113 pp. Emakunde y UPV (Dpto. de Psicología Social y metodología de las Ciencias del Comportamiento).*
6. II Plan de Acción Positiva para las mujeres en la Comunidad Autónoma de Euskadi: ACCIONES E INDICADORES DE EVALUACION. *Documento interno. 102 pp.*

Emakunde y UPV (Dpto. de Psicología Social y metodología de las Ciencias del Comportamiento).

7. INSTRUMENTOS DE EVALUACION DEL II P.A.P.M.E. (II Plan de Acción Positiva para las mujeres en la Comunidad Autónoma de Euskadi). *Documento que recoge los instrumentos de evaluación de las 7 categorías. Emakunde y UPV (Dpto. de Psicología Social y metodología de las Ciencias del Comportamiento).*
8. Memoria de actuación de los Poderes Públicos en 1997 en ejecución del II Plan de Acción Positiva para las mujeres en la Comunidad Autónoma de Euskadi. *Documento interno. 121 pp. Emakunde y UPV (Dpto. de Psicología Social y metodología de las Ciencias del Comportamiento.)*
9. Memoria de actuación de los Poderes Públicos en 1996 en ejecución del II Plan de Acción Positiva para las mujeres en la Comunidad Autónoma de Euskadi. Borrador. *Documento interno. 159 pp. + Anexos (fichas) Emakunde .*
10. Evaluación del (I) Plan de Acción Positiva para las Mujeres en la Comunidad Autónoma de Euskadi (1991-94). *Documento interno. 138pp.*
11. *Emakunde.* Instituto Vasco de la Mujer, PAPMES y Evaluación. Comunicación presentada por Ana Alberdi al seminario sobre *Evaluación de Políticas de Igualdad de Oportunidades entre Mujeres y Hombres*, celebrado en Santander los días 29, 30 de junio y 1 de Julio de 1998, en el marco de los cursos organizados por la Universidad Internacional Menéndez Pelayo.

VALENCIA

1. Orden del 14 de febrero de 1995, de la Conselleria de Cultura, por la que se determina la composición y funcionamiento del Consejo Asesor del Institut Valencià de la Dona.
2. (I) Pla per a la Igualtat d'oportunitats de les Dones. Comunitat Valenciana 1989-1991. *Publicación 86 pp. 1989.*
3. (II) Plan de Actuación del Gobierno de la Generalitat Valenciana para la Igualdad de Oportunidades entre Hombres y Mujeres (1997-2000). *Documento interno 107 pp.*
4. Decreto 29/1988, de 7 de marzo, del Consell de la Generalitat Valenciana, por el que se crea el Instituto Valenciano de la Mujer.
5. Ley 5/1992, de 30 de Junio, de la Generalitat Valenciana, del Institut Valencià de la Dona y Ley 6/1993, de 31 de diciembre, de Presupuestos de la Generalitat Valenciana para el ejercicio 1994, en la que se suprime el Institut.

ANEXO A

ENTREVISTAS REALIZADAS

Directora Entrevistada	Partido gobernante	Organismo	Fecha de entrevista	Antigüedad en cargo
Concha Dancausa	(PP)	Instituto de la Mujer	13/5/1998	1996
Carmen Olmedo	(PSOE)	Instituto Andaluz de la Mujer	12/12/1997	1988
Cristina Sanromán	(PP)	Instituto Aragonés de la Mujer	25/11/1997	1996
Inmaculada González	(PP)	Secretaría Regional de la Mujer (Asturias)	7/8/1997	1995
Pilar Mir	(PP)	Comisión Interdepartamental de la Mujer (Balears)	12/5/1997	1995
Eusebia Nuez	(CC)	Instituto Canario de la Mujer	26/11/1997	1995
Angela Sanroma	(PSOE)	Dirección General de la Mujer (Castilla-La Mancha)	30/4/1998	1997
Ana Parra	(PP)	Secretaría Regional de la Mujer (Castilla y León)	9/7/1997	1995
Nuria Llorach	(CiU)	Instituto Catalán de la Mujer	22/12/1997	1989
Josefa Caraballo	(PSOE)	Dirección General de la Mujer (Extremadura)	29/9/1998	1995
Carmen Bianchi	(PP)	Servicio Gallego de Promoción de Igualdad del Hombre y la Mujer	24/7/1997	1995
Asunción Miura	(PP)	Dirección General de la Mujer (Comunidad de Madrid)	21/8/1997	1995
Belén Fernández	(PP)	Dirección General de Juventud, Mujer y Familia (Murcia)	29/1/1998	1995
Sagrario Loza	(PP)	Dirección General de Bienestar Social (La Rioja)	25/6/1997	1995
Txaro Arteaga	(PNV)	Instituto Vasco de la Mujer (<i>Emakunde</i>)	21/7/1997	1988
Esther Fonfría	(PP)	Dirección General de la Mujer (Comunidad Valenciana)	27/5/1998	1995