



UNIVERSITAT DE
BARCELONA

El tratamiento de la delincuencia en Europa: un estudio meta-analítico

Santiago Redondo Illescas



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial – SenseObraDerivada 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial – SinObraDerivada 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial-NoDerivs 4.0. Spain License.**



UNIVERSIDAD DE BARCELONA

FACULTAD DE PSICOLOGÍA

EL TRATAMIENTO DE LA
DELINCUENCIA EN EUROPA: UN
ESTUDIO META-ANALÍTICO

TESIS DOCTORAL

AUTOR: Santiago Redondo Illescas

Vº Bº LOS DIRECTORES

Dra. María Teresa Anguera Argilaga

Dr. Vicente Garrido Genovés

BIBLIOTECA DE LA UNIVERSITAT DE BARCELONA



0700405074

A Angels, por su tolerancia.

ÍNDICE

FIGURAS 10

TABLAS 11

INTRODUCCIÓN 15

PRIMERA PARTE: REVISIONES PREVIAS DEL TRATAMIENTO DE LA DELINCUENCIA

1. EL TRATAMIENTO DE LA DELINCUENCIA 19

1.1. El debate social 19

1.2. El debate científico 21

1.3. Primeras revisiones de programas de tratamiento de la
delincuencia 22

1.4. El análisis Costo/efectividad 26

1.5. Un moderado optimismo 28

1.6. Nuestra investigación 29

RESUMEN 30

2. REGULACIÓN JURÍDICA DEL TRATAMIENTO DE LA DELINCUENCIA EN
EUROPA 31

2.1. Las normas penitenciarias europeas 32

2.2. La regulación del tratamiento penitenciario 34

2.3. Regulación legal del tratamiento y aplicación de programas de
tratamiento 36

RESUMEN 38

3. LAS TÉCNICAS DE REVISIÓN DE LA LITERATURA 41

- 3.1. Revisiones cualitativas y meta-análisis 42
- 3.2. La técnica meta-analítica 45
- 3.3. Principales fuentes de sesgo en el meta-análisis 49
- 3.4. La técnica costo-efectividad 51
- 3.5. El empleo de la técnica meta-analítica en España 54
- RESUMEN 56

4. LAS REVISIONES PRECEDENTES: TRES META-ANÁLISIS PARADIGMÁTICOS 59

- 4.1. La revisión de J.T. Whitehead y S.P. Lab (Lab y Whitehead, 1988; Whitehead y Lab, 1989). 59
- 4.2. La revisión de programas alemanes de F.Lösel y P.Köferl (Lösel, Köferl, y Weber, 1987; Lösel y Köferl, 1989). 62
- 4.3. La revisión de M.W.Lipsey (Lipsey, 1992a, 1992b). 69
- RESUMEN 77

5. META-EVALUACIÓN DE LOS META-ANÁLISIS PRECEDENTES 79

- 5.1. Criterios de selección de estudios 79
- 5.2. La codificación de los tipos de tratamientos 87
- 5.3. La codificación de las medidas resultado 93
- 5.4. Principales resultados obtenidos 99
- RESUMEN 100

6. ALGUNAS CONCLUSIONES SOBRE LAS REVISIONES PRECEDENTES 103

- 6.1. Problemas metodológicos que presenta los meta-análisis 103
- 6.2. Dificultades a la hora de comparar entre sí los diversos meta-análisis 106
- 6.3. Conclusiones sobre las revisiones y meta-análisis previos 109
- 6.4. ¿Qué es un programa de tratamiento efectivo? 110
- RESUMEN 115

SEGUNDA PARTE: LA INVESTIGACIÓN

7. OBJETIVOS DE INVESTIGACIÓN Y SELECCIÓN DE LOS ESTUDIOS 119

- 7.1. Objetivos de investigación y temporalización 119
- 7.2. Futuro desarrollo de esta investigación 120
- 7.3. Pautas para la localización de estudios en una revisión 121
- 7.4. Nuestro proceso de recolección de estudios 126
- 7.5. La heterogeneidad de los programas objeto de análisis 131
- 7.6. Criterios de recolección y selección de estudios 135
- 7.7. Estudios versus documentos o informes de investigación 139
- 7.8. Estudios localizados y seleccionados 140
- RESUMEN 143

8. METODOLOGÍA PROTOTÍPICA DE LOS ESTUDIOS SELECCIONADOS 145

- 8.1. ¿Cómo evaluar un programa correccional? 146
- 8.2. Ejemplificación de programas incluidos en el meta-análisis 150
 - *Social Skills Training with Young Offenders in a Borstal: An Evaluative Study (Hollin et al., 1986). 150
 - *The Teaching Family Model in the Netherlands: first results of a community based program for problem youths (Slot, 1984, septiembre). 153
 - *Containing Violent Prisoners. An Analysis of the Barlinnie Special Unit (Cooke, 1989). 155
- 8.3. Principales técnicas utilizadas en los programas analizados 158
- 8.4. Principales problemas metodológicos de la investigación evaluativa de programas de tratamiento de la delincuencia 161
- RESUMEN 163

9. LA PLANTILLA DE VACIADO Y CODIFICACIÓN DE LA INFORMACIÓN 165

- 9.1. El proceso codificador 165
- 9.2. El problema de la validez 166
- 9.3. Confección de la primera Plantilla de Codificación 170
- 9.4. Las sucesivas Plantillas y Manuales de Codificación 171
- 9.5. Determinación de las categorías tratamentales 177
- 9.6. Determinación de las categorías de las variables de medida o resultados 184
- RESUMEN 191

10. ESTUDIO PILOTO Y PROCESO CODIFICADOR 193

- 10.1. El estudio piloto 193
 - 10.1.1. Dificultades del proceso de codificación 193
 - 10.1.2. Primeros resultados 195
 - 10.1.3. Conclusiones 199
- 10.2. La codificación de los estudios 200
- RESUMEN 202

11. DISEÑO DE INVESTIGACIÓN Y ANÁLISIS DE DATOS 203

- 11.1. Diseño de investigación 203
- 11.2. Estudio de la fiabilidad de la codificación de las variables 205
- 11.3. Estudio de la fiabilidad de los cálculos de los TEs y de la potencia estadística 206
- 11.4. Estudio del sesgo de publicación 206
- 11.5. El cálculo de la efectividad media o de los TEs 207
- 11.6. Búsqueda de variables moderadoras (en función de variables independientes) 208
- 11.7. Formulación de un modelo explicativo 209
- RESUMEN 210

TERCERA PARTE: RESULTADOS

12. ESTUDIO DE LA FIABILIDAD	215
12.1. Fiabilidad de la codificación	215
12.1.1. Primer estudio de fiabilidad (por consenso de expertos)	215
12.1.2. Segundo estudio de fiabilidad (intra-codificador)	217
12.1.3. Tercer estudio de fiabilidad (inter-codificadores)	218
12.2. Fiabilidad de los cálculos	219
RESUMEN	220
13. ANÁLISIS DESCRIPTIVO DE LAS CARACTERÍSTICAS DE LOS ESTUDIOS	221
13.1. Estudios meta-analizados	221
13.2. Características del tratamiento	223
13.2.1. Características cualitativas	223
13.2.2. Características cuantitativas	226
13.3. Características de los sujetos	228
13.4. Características del ambiente/contexto de la intervención	232
13.5. Características metodológicas	238
13.5.1. Características cualitativas	238
13.5.2. Características cuantitativas	245
RESUMEN	249
14. ANÁLISIS DE LOS EFECTOS: META-ANÁLISIS GLOBAL	251
14.1. Análisis del sesgo de publicación	251
14.2. Exploración de posibles estudios meta-analíticos	252
14.3. Meta-análisis global	255
14.3.1. Resultados favorables, desfavorables y nulos	258
14.3.2. El tamaño del efecto	259

14.3.3. Exploración de la influencia del tratamiento	261
14.3.4. Exploración de la influencia de las características de los sujetos	267
14.3.5. Exploración de la influencia de las características del contexto de la intervención	270
14.3.6. Exploración de la influencia de las variables metodológicas	275
14.3.7. Exploración de la influencia de las variables cuantitativas	284
14.3.8. Algunas precisiones sobre las variables intensidad de la intervención (INTENSID) y seguimiento (SEG)	287
RESUMEN	289

15. ANÁLISIS DE LOS EFECTOS: SEIS META-ANÁLISIS

ESPECÍFICOS 291

15.1. Meta-análisis global en el postest	295
15.2. Meta-análisis global en el seguimiento	298
15.3. Meta-análisis sobre medidas de reincidencia	300
15.3.1. Otros análisis sobre la reincidencia	305
15.4. Meta-análisis sobre medidas institucionales (en el postest)	308
15.5. Meta-análisis sobre ajuste de las habilidades para la interacción social (en el postest)	311
15.6. Meta-análisis sobre ajuste de las habilidades para la interacción social (en el seguimiento)	316
15.7. Otras variables resultado no meta-analizables	317
15.8. Efectividad de los diversos tipos de tratamiento sobre las diferentes categorías resultado	320
RESUMEN	322

16. PROPUESTA DE UN MODELO EXPLICATIVO 325

16.1. El modelo propuesto	325
16.2. Regresión jerárquica con cada cluster	330
16.3. Análisis de regresión jerárquico	333
16.4. Análisis de regresión por pasos (stepwise)	336
RESUMEN	337

CUARTA PARTE: CONCLUSIONES

17. DISCUSIÓN Y CONCLUSIONES	341
17.1. El tratamiento de la delincuencia en Europa	341
17.2. ¿Fueron efectivos los programas analizados?	341
17.3. ¿En qué variables se lograron los mejores resultados?	344
17.4. ¿Qué programas fueron más efectivos?	346
17.5. ¿Qué es un programa efectivo a la luz de nuestros resultados?	347
17.6. Algunos problemas de nuestra investigación	349
17.7. La continuidad de los análisis	354
RESUMEN	355
18. EL PAPEL DE LA CIENCIA EN LA REHABILITACIÓN DE LOS DELINCUENTES	357
18.1. Delincuencia y opinión pública	358
18.2. Opinión pública y medios de comunicación	359
18.3. Ley del péndulo y delincuencia	361
18.4. Los conocimientos técnicos	362
RESUMEN	364
REFERENCIAS	365

FIGURAS:

3.1.	Fórmula del costo de la reincidencia	53
11.1.	Esquema del diseño de nuestra investigación	204
13.1.	Tipología de los tratamientos	225
13.2.	Tipología de los modelos teóricos de los tratamientos	225
13.3.	Duración de la intervención en meses	227
13.4.	Tipología de los sujetos en función de la edad	230
13.5.	Edad media de los sujetos	230
13.6.	Tipología delictiva de los sujetos	231
13.7.	Lugares en que se llevaron a cabo las intervenciones	235
13.8.	Régimen de vida en que se aplicaron las intervenciones	235
13.9.	País en que se realizaron las intervenciones	236
13.10.	Disciplina del primer autor de las intervenciones	236
13.11.	Fecha de publicación de los estudios	237
13.12.	Tipología de los diseños de investigación	242
13.13.	Calidad de los estudios	242
13.14.	Criterio de selección de los sujetos para los programas	244
13.15.	Fuente de publicación de los estudios	244
13.16.	Mortalidad experimental en los grupos tratados	247
13.17.	Mortalidad experimental en los grupos de control	247
13.18.	Duración del período de seguimiento	248
13.19.	Potencia estadística de los estudios	248
14.1.	Meta-análisis global	260
15.1.	Meta-análisis en el postest	297
15.2.	Meta-análisis en el seguimiento	301
15.3.	Meta-análisis sobre reincidencia	304
15.4.	Ajuste institucional en el postest	310
15.5.	Habilidades sociales (postest)	315
15.6.	Habilidades sociales (seguimiento)	318
17.1.	Factores susceptibles de influir en la reincidencia en el delito	353

TABLAS:

4.1. Modelos utilizados para el cálculo del TE	66
4.2. Características de los estudios relacionadas con los TEs	75
5.1. Criterios de selección de estudios	83
5.2. Tipos de tratamientos en diferentes meta-análisis: dimensionalidad tecnológica	94
5.3. Tipos de tratamientos en diferentes meta-análisis: dimensionalidad teórica	96
5.4. Variables resultados en diferentes meta-análisis	98
5.5. Tamaños del efecto promedio (TEs) en diversos Meta-Análisis sobre tratamiento de la delincuencia juvenil y adulta; coeficientes r_m o Φ	99
6.1. Componentes principales de los programas efectivos	112
7.1. Criterios utilizados en las búsquedas informatizadas realizadas en las bases documentales "Dialog", "Psyclit" y "Pascal"	128
7.2. Criterios de selección de estudios para la investigación meta-analítica	137
7.3. Número total de estudios recolectados y seleccionados para el meta-análisis	142
9.1. Esquema de valoraciones sobre la validez (Trascripción a partir del trabajo de Lösel y Köferl, 1989, pág. 340)	168
9.2. Informaciones integrantes de la versión final (Versión 4) de los instrumentos de recogida de información (Plantilla, Bases de datos y Manual de codificación)	175
9.3. Variables resultado obtenidas en el estudio piloto	187
10.1. Modelos teóricos que subyacen a los programas de la muestra, y técnicas o acciones interventivas más frecuentes, en el estudio piloto	197
10.2. Acciones y técnicas interventivas más utilizadas en los programas	198
12.1. Fiabilidad intra-codificación para las variables resultado	218
13.1. Relación de estudios meta-analizados	221
13.2. Características cualitativas (sobre N estudios=57)	224
13.3. Características cuantitativas	228
13.4. Características cualitativas (sobre N estudios = 57)	228

13.5. Características cuantitativas	229	
13.6. Características cualitativas (sobre N estudios = 57)		233
13.7. Características cualitativas (sobre N estudios = 57)		239
13.8. Características cuantitativas	246	
14.1. Resultados de los estudios publicados <u>versus</u> no publicados (análisis del <u>sesgo de publicación</u>)	252	
14.2. Tabla-resumen de la efectividad media y de la homogeneidad de los distintos meta-análisis posibles	254	
14.3. Meta-análisis global para cada uno de los 57 estudios analizados	257	
14.4. Frecuencia de resultados positivos, negativos y nulos		258
14.5. Primera agrupación de los tratamientos	262	
14.6. Segunda agrupación de los tratamientos	264	
14.7. Tercera agrupación de los tratamientos: modelo teórico		265
14.8. Resultados de los tratamientos según su homogeneidad		267
14.9. Resultados en función de la edad de los sujetos		268
14.10. Resultados en función de la presencia o no de alteración psicológica	268	
14.11. Resultados en función de la tipología delictiva de los sujetos		269
14.12. Resultados en función del lugar de la intervención		270
14.13. Resultados en función del régimen de vida de los sujetos		271
14.14. Resultados en función del país en que se realizó el programa		272
14.15. Resultados en función de la disciplina del primer autor		274
14.16. Modalidad del diseño	275	
14.17. Criterio de selección de los sujetos tratados		276
14.18. Modalidad de publicación o difusión de los trabajos		277
14.19. Calidad del diseño: (1) Tamaño muestral grande (n total > 30 en el postest)	278	
14.20. Calidad del diseño: (2) Aleatorización de sujetos a los grupos		279
14.21. Calidad del diseño: (3) Mortalidad muestral baja (< 20%)		279
14.22. Calidad del diseño: (4) Existencia de al menos una VD normalizada	280	
14.23. Calidad del diseño: (5) Existencia de grupo control		280
14.24. Calidad del diseño: (6) Equivalencia entre medidas <u>pre</u> y <u>post</u>		281
14.25. Calidad del diseño: (7) Existencia de pre-test		281
14.26. Calidad global del diseño	282	
14.27. Tipo de diseño: inter/intra	283	

14.28.	Variables cuantitativas y TE	285
14.29.	Análisis de regresión simple. Variable INTENSIDAD	288
14.30.	Análisis de regresión simple. Variable SEGuimiento	289
15.1.	Resumen de los siete meta-análisis realizados	300
15.2.	Meta-análisis global en el posttest	295
15.3.	Meta-análisis global en el seguimiento	298
15.4.	Meta-análisis sobre medidas de reincidencia	303
15.5.	Tabla de contingencia sobre la reincidencia de sujetos tratados y controles	305
15.6.	Meta-análisis de medidas institucionales en el posttest	309
15.7.	Meta-análisis sobre ajuste de las habilidades para la interacción social en el posttest	313
15.8.	Meta-análisis sobre ajuste de las habilidades para la interacción social en el seguimiento	316
15.9.	Promedio de los Tamaños del Efecto: Categorías de Modelos de Tratamiento por Categorías de Variables Resultado	321
16.1.	Cluster: Método	332
16.2.	Cluster: Tratamiento	333
16.3.	Análisis de regresión jerárquico	334
16.4.	Análisis de regresión por clusters	335
16.5.	Análisis de regresión por pasos	336
17.1.	Transformación de los diversos TEs a porcentajes de ganancia en los grupos tratados y controles	344

APÉNDICES

Apéndice 1: Relación de revistas revisadas para la localización de estudios	3
Apéndice 2: Cartas solicitando el envío de estudios correccionales	9
Apéndice 3: Relación de instituciones a las que se solicitaron estudios	15
Apéndice 4: Primera plantilla de codificación	27
Apéndice 5: Bases de datos para la codificación en Word Perfect	45
Apéndice 6: Manual de codificación (Versión 1)	51
Apéndice 7: Manual de codificación (Versión 2)	93
Apéndice 8: Manual de codificación (Versión 3)	135
Apéndice 9: Manual de codificación (Versión 4, última)	175
Apéndice 10: Informe sobre el pilotaje presentado en la "Third European Conference of Law and Psychology" (Oxford, Septiembre de 1992)	219
Apéndice 11: Relación completa de informes localizados	241
Apéndice 12: Bases de datos completa	277
Apéndice 13: Meta-análisis global en el posttest	319
Apéndice 14: Meta-análisis global en el seguimiento	335
Apéndice 15: Meta-análisis sobre reincidencia	353

Apéndice 16: Meta-análisis sobre medidas institucionales (en el postest)	369
Apéndice 17: Meta-análisis sobre habilidades sociales (en el postest)	381
Apéndice 18: Meta-análisis sobre habilidades sociales (en el seguimiento)	393
Apéndice 19: Análisis de regresión simple ponderado (para las variables cuantitativas de los diversos meta-análisis)	405
Apéndice 20: Cálculos del TE medio y la prueba de homogeneidad de todas aquellas variables con menos de 10 estudios	415

INTRODUCCIÓN

La delincuencia suscita en nuestro tiempo un gran debate público. Éste se encuentra presente en todos los ámbitos sociales: en la opinión de los ciudadanos, en los medios de comunicación, en el nivel político, y también en el plano científico. En muchos de estos sectores se plantea abiertamente qué debe hacerse para atajar la delincuencia. Con frecuencia este debate se mueve en parámetros dicotómicos, semejantes a los siguientes: ¿Debe castigarse con mayor dureza a los delincuentes? ¿O, por el contrario, deben suavizarse las medidas punitivas tradicionales, ofreciéndoles más oportunidades de reinserción social? Una pregunta implícita en la polémica podría ser formulada en los siguientes términos: ¿Qué utilidad tienen unos y otros planteamientos?

La presente investigación aborda esta controversia desde una perspectiva científica. Para ello, hemos acometido el análisis de programas de tratamiento de la delincuencia realizados en diversos países europeos a lo largo de la última década, intentando responder a la cuestión de su eventual efectividad o inefectividad. Ciertamente, la respuesta científica no es la única que debe darse a este problema, pero también es, sin duda, necesaria.

Hemos estructurado nuestra investigación en cuatro grandes apartados. En el primero -capítulos 1 al 6- repasaremos, sobre todo, las investigaciones anteriores que han analizado el funcionamiento de conjuntos diversos de programas de tratamiento de la delincuencia. En la segunda parte -capítulos 7 a 11- describiremos con detalle nuestra propia investigación de programas europeos: nuestros objetivos, cómo hemos seleccionado los estudios, a qué variables prestaremos atención prioritaria, cómo hemos extraído y codificado la información, y qué tipos de análisis realizaremos. La tercera parte -capítulos 12 a 16- se destinará a presentar los principales resultados obtenidos. En ella, mediante tablas y gráficas diversas, detallaremos los

efectos logrados por los programas de tratamiento que hemos analizado: qué técnicas interventivas se utilizaron y por cuánto tiempo, a qué tipos de delincuentes se dirigieron, dónde se llevaron a cabo, con qué metodologías se evaluaron los resultados, y qué resultados se obtuvieron. Finalmente, concluiremos nuestro trabajo en una cuarta parte -integrada por dos capítulos- en la que debatiremos las principales consecuencias y derivaciones de nuestro estudio. Paralelamente, en volumen aparte recogemos un conjunto de apéndices relativos al instrumental empleado para la localización y codificación de los estudios, así como el detalle de muchos de los resultados obtenidos, que en el informe de investigación han sido solamente resumidos.

En todo trabajo de investigación se hace uno deudor de muchas personas que lo han ayudado, de una u otra manera. Deseo mostrar mi agradecimiento a las siguientes: a Margarita Cortadella, Amanda Vidal y Yolanda Corredera, bibliotecarias del Centre d'Estudis Jurídics i Formació Especialitzada, cuya experta y siempre amabilísima ayuda me fue imprescindible durante la etapa de localización de materiales; a Eulàlia Luque, Mírian Thió y Dani, que con gran paciencia, me auxiliaron en el proceso de confección y revisión de las diversas plantillas/manuales de codificación; a los doctores Angel Blanco, María Dolores Peris, José Luis Losada, Cristina Rechea y Esther Giménez-Salinas, y a Eva Deumal, de quienes recibí orientaciones diversas y apoyo en cuestiones informáticas, teóricas y metodológicas; a mis directores de tesis, Dra. María Teresa Anguera y Dr. Vicente Garrido, cuyas sugerencias, supervisión, trabajo directo de codificación, amistad y permanente motivación han resultado absolutamente imprescindibles para esta investigación. Finalmente, al Dr. Julio Sánchez Meca, y al Prof. Fulgencio Marín. Su trabajo imaginativo y su ayuda paciente han permitido dotar a este estudio de la cualificación estadística y de la sutileza analítica sólo posible cuando un experto en meta-análisis es quien lo hace.

Agradezco la ayuda económica recibida del Instituto de Criminología de la Universidad de Barcelona que me permitió sufragar gastos de mecanografiado e informática.

**PRIMERA PARTE: REVISIONES
PREVIAS DEL TRATAMIENTO
DE LA DELINCUENCIA**

1. EL TRATAMIENTO DE LA DELINCUENCIA

1.1. El debate social

La delincuencia es un fenómeno estructural en nuestra sociedad. El conocimiento de las cifras "reales" de delincuencia presenta importantes dificultades metodológicas. Los indicadores más evidentes con que contamos para ello nos aportan información controvertida sobre la evolución del fenómeno delincuencial. Por un lado, algunas estadísticas policiales y ciertos estudios de victimización nos informan sobre un relativo estancamiento, e incluso una ligera reducción, del número de delitos que son cometidos y denunciados o detectados por los órganos de control. Por otra parte, si atendemos a las estadísticas judiciales y penitenciarias -siendo la cárcel el más extendido sistema de control penal- la conclusión a la que se llega es justamente la contraria: en casi todos los países occidentales se está operando un paulatino, y a veces rápido, crecimiento del número de delincuentes encarcelados.

Durante la pasada década en los Estados Unidos se produjo un decremento promedio de las tasas de delincuencia de un 10%, mientras que los encarcelados aumentaron en un 100% (Taylor, 1992). En España podemos observar también un ligero decremento a lo largo de los últimos años de las cifras de delincuencia, según datos policiales y victimológicos (El Periódico de Cataluña, 16 de octubre de 1992; Cuerpo Nacional de Policía, 1992; Alabart, Aragay y Sabaté, 1991). Por el contrario, los datos judiciales y penitenciarios nos refieren un evidente aumento del número de sentencias condenatorias y de las cifras de encarcelados (INE, 1989; Giménez-Salinas y Rifà, 1992), habiendo experimentado estas últimas un incremento del 70% en un período de tan sólo seis años (de 1985-1991). Muy probablemente, unas

cifras y otras -las que disminuyen y las que aumentan- nos estén informando sobre fenómenos distintos: las cifras de victimización y policiales hablarían de la dimensión "real" del fenómeno delictivo en una sociedad, mientras que las judiciales y penitenciarias nos estarían más bien mostrando en qué medida esa sociedad utiliza el sistema más duro de control penal -la cárcel- para reprimir su delincuencia (Redondo, Funes y Luque, 1993).

Hemos puesto de relieve en trabajos anteriores la magnitud social que tienen tanto el problema de la delincuencia en su conjunto como, más específicamente, el problema de las prisiones (Redondo, 1993). Significación social que puede ser identificada, en primer lugar, según ya hemos comentado, en parámetros del gran volumen de personas que son encarceladas en los distintos países para cumplir una pena de privación de libertad. En la actualidad existe en España una población carcelaria media superior a los 44.000 encarcelados, lo que representa una proporción de entorno a 112 encarcelados por cada cien mil habitantes. Si formulamos una mera traslación de estas cifras al conjunto de la población mundial, el número de individuos en prisión en todo el mundo podría estimarse en una cuantía no inferior a tres millones de personas. Cifra que, sin duda, será mucho más elevada, si consideramos que nuestra proporción de encarcelados no es, por el momento, de las más prominentes. Baste referir aquí que muchos países de nuestro entorno cercano -Francia, Italia, Gran Bretaña y Alemania- se sitúan por encima de nosotros en sus tasas de presos/ciudadanos (Giménez-Salinas y Rifà, 1992), y que el país más desarrollado del mundo -los Estados Unidos de América- roza los 500 encarcelados por cada cien mil habitantes, o sea multiplica por más de cuatro la proporción española.

No tiene una menor significación la magnitud de este fenómeno en términos económicos. En España los presupuestos estatales destinados por el Ministerio de Justicia a las prisiones fueron de más de 70.000 millones de pesetas durante 1990, lo que ha supuesto un incremento del gasto no inferior al 400% a lo largo de pasada década. En Cataluña, por su parte, esta cifra ha evolucionado desde 2.165 millones de presupuesto en 1985 hasta casi los 10.000 millones en 1991 -lo que implica un aumento del 360% en tan sólo seis

años⁻¹.

El tercer parámetro en que puede calibrarse la envergadura social del problema delictivo y carcelario se refiere a la amplia repercusión que tiene en la opinión pública. Baste mencionar la insistente atención prestada por los medios de comunicación de masas -televisión, prensa, radio y cine- a las cuestiones delictiva y carcelaria. La delincuencia, la reincidencia en el delito, la fuga de presos, los incidentes en prisión, los permisos de salida, la rehabilitación de los delincuentes, y temas similares, son cuestiones a las que se dedican amplios espacios televisivos o radiofónicos de debate, reportajes, encuestas ciudadanas, etc.

Ciertamente, una cuestión distinta, una vez que universalmente es reconocida la importancia de ambos fenómenos interactivos -delincuencia y privación de libertad- se refiere a qué es lo que debe hacerse para abordarlos de modo más efectivo. Éste es, tal vez, el punto donde se producen opiniones ciudadanas abiertamente divergentes entre sí: para la prevención de la delincuencia hay quienes proponen básicamente el mantenimiento y endurecimiento de las medidas punitivas -la cárcel, las multas, etc.- con el objetivo de favorecer la disuasión de los delincuentes; otros, por el contrario, postulan la necesidad del mejorar las condiciones sociales que afectan a los delincuentes, que podrían hallarse en el origen de sus inicios en la delincuencia; y existen también quienes plantean mayores medidas de educación y prevención comunitaria u otras, con tal de favorecer un desarrollo personal de los delincuentes que haga más probable su integración social.

1.2. El debate científico

En el plano de los expertos, la polémica más importante al respecto se ha producido en todo el mundo a lo largo de las últimas décadas entre los

¹El gasto público total en las prisiones es muy superior a estas cifras. Debe tenerse en cuenta que en los presupuestos enunciados, del Ministerio de Justicia, no se incluyen las partidas presupuestarias correspondientes a la custodia de las prisiones y al traslado de los presos, que competencialmente caen dentro de los presupuestos del Ministerio de Interior, sin que sea posible para nosotros su desglose.

partidarios del castigo justo de los culpables y los del objetivo de la rehabilitación de los delincuentes (Lipsey, 1992a, 1992b; Palmer, 1992; Lösel, en prensa). Esta polémica ha tenido entre sus principales mentores, en un extremo, a profesionales del ámbito del derecho y la justicia (como profesores de derecho penal, abogados, jueces y fiscales), en gran medida defensores de una justicia punitiva -el tratamiento, se ha dicho, haría perder el cariz punitivo que resulta intrínseco y necesario a la justicia penal (Lösel, en prensa)- y garantista -el tratamiento, se ha afirmado, modifica la personalidad y no tiene un suficiente control legal (en España, por ejemplo, Muñoz Conde, 1987; García Pablos, 1979; Mapelli, 1983; Bergali, 1987; Larrauri, 1991); el debate ha tenido, en el extremo opuesto, a profesionales e investigadores de las ciencias sociales, del comportamiento y de la educación (criminólogos, psicólogos, sociólogos, pedagogos, educadores, trabajadores sociales, etc.), que mayoritariamente se han decantado del lado de la idea de rehabilitación (en España, por ejemplo, Alarcón, 1989; Garrido, 1986; García y Sancha, 1985; Redondo, 1989, 1993b), como la única garantía de promoción de ambientes correccionales más humanos (Whitehead y Lab, 1989).

En la actualidad, tanto en el Estado Español como en los países europeos de nuestro entorno sigue debatiéndose la conveniencia de aplicar programas de tratamiento dirigidos a la rehabilitación de los delincuentes, ya sea en las prisiones, centros de reforma juvenil, o en la propia comunidad. Aunque este debate tiene múltiples componentes, entre los que se encuentran los ideológicos, presupuestarios, etc., uno de los factores más aducido en el mismo es el relativo a la efectividad o no de los programas que hasta ahora se han aplicado.

1.3. Primeras revisiones de programas de tratamiento de la delincuencia

En Norteamérica, donde se encuentran más desarrollados tanto este tipo de programas correccionales como divulgados los modelos teóricos (psicológicos, sociológicos, educativos, etc.) que subyacen a los mismos, se han llevado a cabo una veintena de revisiones con el propósito de corroborar la eficacia o

ineficacia de conjuntos amplios de programas aplicados². Las primeras de estas revisiones, realizadas en los años setenta, tuvieron por objetivo analizar la efectividad que habían tenido programas aplicados en años precedentes (décadas de los cincuenta y sesenta). Lipsey (1992a) y Palmer (1992) han recogido los principales resultados obtenidos por estas tempranas revisiones. Seguiremos a estos autores para su breve comentario. La primera y más divulgada revisión, es la llevada a cabo por Lipton, Martinson y Wilks (1975), en la que evaluaron 231 trabajos interventivos dirigidos tanto a delincuentes juveniles como adultos. La más afamada conclusión sobre la misma es debida a Martinson (1974), quien escribió que "Con aisladas excepciones, los esfuerzos rehabilitadores que han sido descritos hasta ahora no han tenido efectos apreciables sobre la reincidencia" (p.25). Romig (1978) dirigió su atención exclusivamente a programas de tratamiento de delincuentes juveniles, catalogando con gran detalle los trabajos realizados, con la pretensión de conocer las características más sobresalientes de las intervenciones exitosas. Sus conclusiones fueron igualmente negativas, detectando escasos resultados favorables. Wright y Dixon (1977) tampoco fueron más allá de encontrar reducidos efectos en las intervenciones que analizaron, aplicadas con delincuentes juveniles.

En síntesis, estas primeras meta-evaluaciones mostraron escasos efectos

²Los Estados Unidos y Canadá son países de una inmensa extensión territorial y población, si los comparamos con los países de la Europa Occidental. No mantendremos en este trabajo, ni mucho menos, la tesis de que en conjunto los países norteamericanos sean un modelo que deba ser imitado por nosotros por lo que se refiere al funcionamiento de sus prisiones y otras instituciones de reforma juvenil. Antes bien, en nuestra opinión, los sistemas correccionales de Norteamérica, tanto juveniles como de adultos, considerados globalmente, son excesivamente rígidos, violentos y deshumanizados, por lo que serían inaceptables para los cánones europeos, y, especialmente, españoles. Sin embargo, Canadá y Estados Unidos, por el hecho aludido de ser países muy grandes, y poseedores de un mayor desarrollo económico, social y científico, son, a la vez, lugares de grandes contrastes, en los que pueden coexistir, a un tiempo, centros carcelarios y de reforma de la máxima dureza y programas correccionales sumamente innovadores. Éstos son los que nos interesan aquí.

En un viaje de trabajo que realicé -dentro de una expedición catalana- en 1987 para visitar programas correccionales en los estados de Quebec y Ontario (Canadá) tuve ocasión de corroborar la afirmación anterior. Vimos prisiones y centros de reforma juvenil con regímenes extraordinariamente duros, que en España quebrarían principios jurídicos y constitucionales fundamentales. Paralelamente, en otros centros se estaban aplicando programas de gran interés desde la perspectiva de la rehabilitación.

positivos de los programas de tratamiento de la delincuencia sobre la reincidencia posterior de los sujetos, y concluyeron con un generalizado pesimismo al respecto. A raíz de estas primeras revisiones, el punto de vista más arraigado entre los investigadores, y en el público en general, fue que los programas de tratamiento no daban resultado y debían, por ello, ser abandonados (Whitehead y Lab, 1989; Lösel, en prensa). En la literatura especializada en la rehabilitación de delincuentes se hizo famosa la tajante frase de Martinson (1974): "nothing works", nada da resultado. Su laconismo se expandió de publicación en publicación, dentro de este ámbito, como un reguero de pólvora. Hasta tal punto que muchos investigadores de refrito citan aún el "nothing works", después de casi veinte años que lloviera, como si de la última conclusión científica en esta materia se tratara.

Estas primeras revisiones de los años 70 fueron pronto contestadas por algunos autores que como Palmer (1975, 1992) encontraron en ellas una doble negligencia: en primer lugar, precipitación en la generalización de sus negativas conclusiones, habiendo pasado por alto en sus análisis programas que habían sido más exitosos; por otro lado, no haber tomado en consideración la adecuación entre las características de los jóvenes tratados y el tipo de tratamiento utilizado.

Gendreau y Ross (1979) llevaron a cabo una nueva revisión de programas correccionales, en la que detectaron un conjunto amplio de intervenciones que sí habían resultado efectivas en el tratamiento de la delincuencia, revisión que presentaron en un trabajo con el sugerente título de Effective correctional treatment: Bibliotherapy for cynics.

Con posterioridad Sechrest, White, y Brown (1979), y Martin, Sechrest, y Redner (1981), de la National Academy of Sciences' (NAS) -Panel on Research on Rehabilitative Techniques- efectuaron nuevos análisis basándose para ello tanto directamente en programas evaluados como en revisiones previas de los mismos. Sus conclusiones fueron nuevamente pesimistas. Pese a ello, el NAS subrayó la posibilidad de que la no localización de resultados radicara más en la naturaleza de los programas evaluados que en la inadecuación de la idea rehabilitadora, señalando una serie de inconvenientes de la literatura correccional que había sido analizada, tales como la ausencia de diseños bien controlados, la carencia de

medidas suficientemente sensibles, y la falta de tratamientos potentes y adecuadamente aplicados.

Así pues con las últimas revisiones mencionadas se fue afianzando la esperanza en que intervenciones más recientes, con mayor fundamentación teórica e integridad en la aplicación que los antiguos y rudimentarios programas inspeccionados por R. Martinson, serían más efectivas (Whitehead y Lab, 1989).

Todas estas primeras revisiones de programas sociales en distintos ámbitos -también en el de la delincuencia- se formularon desde una perspectiva eminentemente cualitativa, que dejaba exclusivamente a la pericia y perspicacia del propio revisor o revisores la valoración final de la efectividad del conjunto de estudios evaluados. Sin embargo, desde el ámbito metodológico, a principios de los años ochenta fue recomendada, para llevar a cabo estas revisiones, la inclusión de técnicas cuantitativas, y más concretamente el uso del meta-análisis. En síntesis, el meta-análisis es una técnica estadística que permite formular en parámetros numéricos semejantes los resultados obtenidos por diferentes programas, de modo que sus efectos sean comparables entre sí (Gómez, 1987; Cooper, 1989).

A partir de entonces las revisiones de programas correccionales de la década de los ochenta han sido todas ellas meta-análisis de conjuntos numerosos de programas. Todas estas revisiones se han llevado a cabo en Norteamérica, con la sola excepción de una realizada en Alemania (Lösel, Köferl y Weber, 1987; Lösel y Köferl, 1989). La última de estas revisiones es la de Lipsey (1992a, 1992b), en Norteamérica, que analizó más de 400 programas de intervención con delincuentes, correspondientes a casi cuatro décadas.

El panorama Europeo es distinto. En primer lugar, es menor el número de programas correccionales llevados a cabo, si atendemos al pequeño número que son publicados. En segundo lugar, al tratarse de países y estados diferentes -los europeos- existe una mayor dispersión idiomática y conceptual. Todo ello hace que en la naciente "Europa unida" carezcamos de una integración y evaluación semejante a las realizadas en Norteamérica, que nos permita conocer qué se ha hecho y qué se está haciendo en materia de tratamiento de la delincuencia en el continente europeo, y cuáles son sus

resultados.

También en España se han llevado a cabo durante los últimos años algunos programas que han mostrado efectividad sobre la adquisición de nuevas habilidades sociales por los sujetos encarcelados, e incluso sobre la reducción de sus niveles de reincidencia (por ejemplo, Redondo, 1983, 1984; Garrido et al., 1989; Garrido y Sanchís, 1990; Legaz et al., 1990; Sánchez Benito y Polo Ruíz, 1990; Bayón y Compadre, 1991). Sin embargo, hasta ahora la efectividad de los mismos no ha podido ser comparada entre sí ni con programas de otros países. Por ello también éstos serán incluidos en esta investigación comparativa.

1.4. El análisis Costo/efectividad

Por otro lado, una cuestión clave del desarrollo de programas de rehabilitación con delincuentes, en las prisiones o en la comunidad, tiene que ver con la evaluación del costo/efectividad de los mismos. De este modo, no es lo mismo que estemos hablando de un programa susceptible de afectar a 8 sujetos, que de un programa efectivo capaz de influir positivamente a todos los habitantes de un centro penitenciario, o a todos los delincuentes participantes en una determinada intervención comunitaria.

Situándonos, pues, en la perspectiva costo/efectividad a que hemos aludido, es bien conocido que ciertas estrategias penales, como por ejemplo las penas de privación de libertad en régimen cerrado, tienen un coste económico muy elevado, derivado tanto de la carestía de las modernas construcciones carcelarias como de los gastos del personal y de las sofisticadas medidas de seguridad. Pese a su alto costo económico, la privación de libertad en regímenes cerrados produce elevadísimas cotas de reincidencia (Tournier y Barre, 1990; Redondo et al., 1993), lo que, a su vez, incrementa los gastos producidos por los nuevos encarcelamientos de los sujetos, el aumento de las medidas de seguridad, etc, en un círculo sine fine. En cambio, se ha sugerido que algunos programas de rehabilitación en contextos de control más blandos (régimen abierto, libertad vigilada, trabajo al servicio de la comunidad, etc.), que son mucho más económicos, han mostrado importantes efectos positivos tanto para la reducción de la reincidencia de los sujetos

como sobre la mejora del propio funcionamiento de las instituciones correccionales (entre otros: Taylor, 1992; Palmer, 1992; Lösel, en prensa).

Por todo ello, consideramos también de vital importancia, al analizar programas de tratamiento de la delincuencia, la cuantificación de aspectos relativos al coste/efectividad de su funcionamiento y resultados. Desgraciadamente, aunque éste era otro de los propósitos de partida de esta investigación, su análisis deberá ser aplazado por el momento. Como tendremos ocasión de exponer más adelante, son escasísimos los informes de investigación que hemos analizado que se refieren a los costos que fueron sufragados para la aplicación de los programas, por lo que nos ha resultado imposible la atención a este importante aspecto.

Finalmente, la cuestión sobre la efectividad de los programas de rehabilitación aplicados con delincuentes adquiere una relevancia especial ante las dos siguientes eventualidades: en primer lugar, las reformas de las legislaciones penales y penales-juveniles que se acometen en los distintos países europeos vendrían necesitadas, desde una perspectiva técnica, de una mayor información sobre la efectividad de las medidas e intervenciones utilizadas en la lucha contra la delincuencia; como, por el contrario, existe poca información al respecto de la efectividad diferencial de las medidas y programas aplicados con anterioridad (ya sean el mero internamiento, programas educativos, entrenamiento en habilidades sociales y de vida, desarrollo moral, "probation", atención social en la propia comunidad, etc.), estas reformas han de hacerse frecuentemente con mayores niveles de incertidumbre de los que serían aconsejables. Éste es el caso de España, en donde en este momento se está gestionando tanto la reforma de la legislación de menores como de la ley penal. En segundo término, también ciertos organismos internacionales -Naciones Unidas, Consejo de Europa- se plantean el dictado de normas y recomendaciones a los países miembros sobre el "Tratamiento" de los reclusos, de los jóvenes internados en centros de reforma, y de la delincuencia en su conjunto; dadas las diferencias existentes entre los distintos Estados europeos, tanto en el plano jurídico como en las prácticas correccionales, con toda probabilidad la profundización en el conocimiento sistemático de los puntos de confluencia y diferenciación existentes en este ámbito tiene gran interés para facilitar el trabajo en común al acometer la delincuencia.

1.5. Un moderado optimismo

Debido a todos estos factores, en la actualidad se aprecia a nivel internacional un renovado interés y mayores esperanzas en lo concerniente a las posibilidades que pueden tener los programas de tratamiento de los delincuentes. Lösel (en prensa) ha sugerido algunas de las razones que pueden explicar esta nueva situación más positiva:

1) La existencia de un mayor cúmulo de investigación psicológica sobre factores de interés criminológico (por ejemplo, en temáticas como la detección de los déficits de muchos delincuentes en habilidades sociales, autocontrol, etc.).

2) Frente a previas concepciones que estipulaban la especificidad del ámbito correccional, es una idea emergente que los principios básicos de la intervención psicológica y educativa son semejantes tanto para las prevenciones primaria -inespecífica- y secundaria -con sujetos de riesgo- como para la terciaria -esto es, para el tratamiento de delincuentes persistentes-, que es la que ahora nos ocupa. Esta perspectiva es "normalizadora" del ámbito correccional, y nos sugiere que cualesquiera técnicas y procedimientos de tratamiento útiles para la socialización de las personas -educación, formación laboral, habilidades sociales, terapias, etc.- pueden tener igual interés y efectividad en su aplicación con sujetos delincuentes.

3) La mayor experiencia acumulada durante los últimos años ha permitido diseñar y aplicar programas de modo más preciso y con mayor integridad e intensidad.

4) Se han operado sustanciales mejoras en las evaluaciones, que se han hecho más diferenciales, permitiendo discriminar de modo más eficiente la efectividad de los programas aplicados sobre grupos diversos de delincuentes, contextos sociales diferentes, y sobre criterios resultado distintos.

5) Los meta-análisis recientes han permitido una más sistemática integración

de información, proveniente de programas diversos, haciéndola comparable entre sí en términos cuantitativos (mediante el análisis de los tamaños del efecto, o TEs).

1.6. Nuestra investigación

La presente investigación se fundamenta en la consideración de la probable existencia de diferencias críticas en la efectividad y rentabilidad social de distintos programas de tratamiento con delincuentes, ya sean juveniles o adultos. Las revisiones efectuadas en Norteamérica han podido comprobar, por ejemplo, una mayor efectividad (reduciendo la reincidencia de los sujetos) de aquellos programas que han utilizado entrenamiento cognitivo entre sus estrategias (Gendreau y Ross, 1979; Ross, 1987; Andrews et al., 1990). Ello ha motivado el empleo de técnicas semejantes en algunos de los programas que se aplican en las prisiones de Cataluña (Garrido, Redondo y Pérez, 1989; Garrido, 1990). Por tanto, no es indiferente lo que se haga en esta materia: existen procedimientos más potentes y eficaces que otros, cuyo uso se debería promover.

Por lo que se refiere a Europa, ignoramos, tal y como ya hemos puesto de relieve, no sólo qué programas resultan más positivos en el tratamiento de la delincuencia, sino también, en buena medida, cuáles son estos programas, en qué modelos teóricos se fundamentan, qué técnicas han sido utilizadas, cuánto dinero se gasta en ellos, etc.

Por esta razón, creemos que tiene gran interés científico conocer esta información con la claridad suficiente para poder hacer un uso aplicado de la misma en el diseño de futuros programas con delincuentes. Y para acometer esta finalidad, consideramos que el doble empleo de una metodología cualitativa/cuantitativa, mediante el uso de procedimientos de análisis ya probados en otros países (Norteamérica, en este caso), puede ser considerado el más conveniente.

RESUMEN

Nuestro primer capítulo tiene un doble cariz conceptual y programático. En él hemos prestado atención, en primer lugar, al debate social que suscita la delincuencia y, más concretamente, el tratamiento y la rehabilitación de los delincuentes. El interés ciudadano sobre estas materias dimana tanto del gran volumen de delincuentes que son encarcelados como de los ingentes presupuestos públicos que se invierten en ello. Paralelamente existe también un debate científico, más focalizado en conocer y decidir cuáles son las medidas y técnicas más útiles para el tratamiento exitoso de los delincuentes. Dos son las posturas básicas defendidas por los expertos: unos se decantan por el mero castigo justo y ejemplar de quienes delinquen; otros postulan la necesidad de intervenir de un modo científico mediante diversos programas de tratamiento.

Para conocer la efectividad diferencial de las intervenciones con delincuentes se han efectuado diversas revisiones, especialmente en Norteamérica. Las primeras, referidas a programas más antiguos, obtuvieron resultados pesimistas y concluyeron que la rehabilitación de los delincuentes realmente "no daba resultado". En cambio las exploraciones de programas más recientes han hallado resultados más positivos y esperanzadores, concluyendo que algunas técnicas resultan efectivas con ciertos grupos de delincuentes. Este nuevo optimismo nos conduce a acometer esta investigación con el propósito de evaluar los programas de tratamiento de la delincuencia que se han realizado durante la última década en los países europeos.

2. REGULACIÓN JURÍDICA DEL TRATAMIENTO DE LA DELINCUENCIA EN EUROPA

El tema que nos ocupa tiene un referente normativo que no podemos pasar por alto. Por ello, en el presente capítulo atenderemos sucintamente a la regulación jurídica del tratamiento de la delincuencia en los países europeos.

El desarrollo de programas de rehabilitación de delincuentes en un determinado país depende, sin duda, de múltiples factores. Entre éstos se encontrarían características globales, inespecíficas, pero de gran importancia, tales como su sistema político, su nivel económico, social y cultural. Más directamente podría influir el progreso alcanzado en dicho país por la investigación criminológica, tanto académica como aplicada, así como la cualificación técnica del personal encargado de los servicios correccionales, y la dotación de recursos para los mismos.

No menos relevante es, con toda probabilidad, la concepción filosófica y científica existente sobre la delincuencia y su tratamiento. La preponderancia de opiniones favorables o contrarias a la rehabilitación de los delincuentes, la disponibilidad de modelos teóricos para su abordaje, y, finalmente, la concepción jurídica sobre la cuestión, plasmada en las leyes que estructuran la política criminal.

En términos absolutos, no creemos que una legislación criminal de corte estrictamente punitivo impida el desarrollo de programas de rehabilitación, ni, por el contrario, que una de cariz rehabilitador asegure la aplicación de aquéllos. El Derecho y la práctica social caminan, a veces, por sendas distintas. Sin embargo, presuponemos que una legislación que ampare ideas rehabilitadoras para con los delincuentes puede facilitar el avance en esta

materia, si concurren algunos de los elementos que antes hemos mencionado.

Aunque las políticas criminales suelen recogerse en diversas normas -leyes penales, procesales, penitenciarias o de ejecución de sentencias, de servicios sociales, etc.- aquí nos limitaremos al estudio de las leyes penitenciarias. Dado que la privación de libertad es la medida penal más ampliamente utilizada en todos los Estados, las normas penitenciarias acostumbra a ser el lugar idóneo para la definición filosófica y práctica del abordaje de la delincuencia. Por otra parte, la inmensa mayoría de estudios aquí recogidos corresponden a intervenciones llevadas a cabo dentro de los sistemas penitenciarios. Así, en las páginas que siguen, analizaremos específicamente cómo distintas legislaciones penitenciarias europeas afrontan la problemática del tratamiento de los delincuentes.

2.1. Las normas penitenciarias europeas

Ya desde el siglo pasado, distintos Congresos Penitenciarios Internacionales reunieron a criminólogos, penitenciaristas y otros expertos en delincuencia, que dictaron pautas de actuación para acometer este fenómeno y para el tratamiento de los delincuentes. Más recientemente, importantes organismos internacionales, como la ONU y el Consejo de Europa, han venido dictando principios, normas, y recomendaciones en estas mismas materias, cuyo objetivo ha sido el que las leyes y actuaciones de los estados se inspiren en ellas.

En un trabajo anterior hicimos un extracto de las más recientes normas europeas -las de la Recomendación N R (87) 3 del Comité de Ministros de los Estados miembros del Consejo de Europa-, que hacían referencia directa a la intervención o tratamiento de los delincuentes (Redondo, 1993a). En síntesis, los principios establecidos por esta Recomendación del Consejo de Europa, en lo relativo al tratamiento de los delincuentes, son los siguientes:

-Son objetivos prioritarios del tratamiento salvaguardar la salud y la dignidad de los encarcelados, así como "desarrollar su sentido de la responsabilidad y dotarles de competencias que les ayudarán a reintegrarse en la sociedad, vivir en la legalidad y subvenir a sus propias necesidades

(...)" (Art. 3).

-También, reducir los efectos perjudiciales que comporta la detención.

Para ello, la Recomendación europea propone la utilización de todos los medios que permitan un tratamiento personalizado (sean éstos curativos, educativos, morales, espirituales, etc.), una vez realizadas las oportunas consultas al personal competente. Específicamente, se consignan una serie de estrategias interventivas, que deberán ser empleadas:

-La orientación y formación profesionales de los sujetos, así como el ejercicio de una actividad laboral propiamente dicha.

-La aplicación de sistemas de comunicación y gestión que faciliten las relaciones positivas entre el personal y los internos, de modo que se propicien regímenes y tratamientos más eficaces.

-Favorecer el régimen abierto.

-Diseñar los programas de tratamiento, una vez que se haya efectuado un adecuado estudio de la personalidad de los sujetos.

-Establecer específicos sistemas que propicien la cooperación y participación de los sujetos en su tratamiento.

-Instaurar programas educativos normalizados, programas deportivos, y el uso de bibliotecas.

-Finalmente, se recomienda la aplicación de sistemas que permitan el retorno progresivo de los condenados a la vida en sociedad, mediante su liberación condicional, y la cooperación de los servicios sociales comunitarios.

No todos los países europeos, incluso los redactores y suscriptores de estas recomendaciones, han recogido por igual su contenido en sus respectivas legislaciones. Veamos a continuación cuál es la situación legal del tratamiento penitenciario en algunos de estos países.

2.2. La regulación del tratamiento penitenciario

Mercedes García y Ángel de Sola, penalistas de las universidades de Barcelona, han reunido en un trabajo reciente las legislaciones penitenciarias de una serie de países europeos, más cercanos a nuestra órbita cultural, y a nuestra praxis penal y penitenciaria (Alemania, Bélgica, Francia, Holanda, Inglaterra/Gales, Italia, y Suecia). Tomaremos este trabajo como fuente de nuestro análisis (García Arán y de Sola, 1992).

En primer lugar, un rápido comentario sobre las fechas de promulgación de las diversas leyes penitenciarias. La más antigua de las legislaciones penitenciarias es la holandesa, que data de 1951 (aunque guarda una estrecha relación con leyes penales reformadas en 1987). La más reciente es la española, promulgada por Ley Orgánica en 1979, y desarrollada en un reglamento penitenciario de 1981. En medio de estas dos fechas extremas se encuentran todas las demás: la legislación penitenciaria inglesa/galesa data de 1964 (con ciertas reformas posteriores); la belga es de 1965 (con modificaciones ulteriores); la ley sueca data de 1974; mientras que las leyes italiana y alemana corresponden a 1975 y 1976, respectivamente (con sendas reformas subsiguientes).

En general, se atribuye al tratamiento un carácter inseparable de la propia ejecución de la pena (legislaciones italiana, belga, alemana, sueca y española). Esta regla presenta importantes excepciones. La ley francesa no realiza una declaración formal sobre el tratamiento, y para la inglesa/galesa el objetivo principal de las instituciones carcelarias es la custodia de los presos. Por otro lado, en Holanda se realiza una distinción entre centros de mero cumplimiento de penas, y centros de tratamiento. Sobre el cumplimiento de las penas de privación de libertad se dice que, aunque han de mantener un carácter de "castigo a la medida, su aplicación también debe servir para la preparación (...) de la vida social" (art. 26). El caso favorable más extremo lo constituye la legislación penitenciaria sueca (Ley de tratamiento de los delincuentes en centros de rehabilitación), en la que la pena se equipara al tratamiento.

En la mayoría de estas legislaciones no existe una regulación concreta y

detallada del contenido del tratamiento. La única excepción es la ley española, que dedica todo el Título III (de los seis títulos de que consta) -o sea, 14 artículos de un total de 80 que integran su redactado- a la concreción del tratamiento penitenciario, que es definido como el "conjunto de actividades directamente dirigidas a la consecución de la reeducación y reinserción de los penados" (LOGP, art. 59.1).

Todas las legislaciones mencionadas optan por el principio de individualización del tratamiento. Más concretamente, consideran que el primer paso para un tratamiento adecuado debería ser una apropiada clasificación de los sujetos, que llevaría a la oportuna individualización de sus problemas y circunstancias, facilitándose con ello el empleo de la intervención más conveniente.

Un aspecto jurídico importante se refiere a los límites del tratamiento, esto es, hasta dónde debe y puede llegarse en el propósito de cambiar actitudes, comportamientos y hábitos en los sujetos, mediante el uso de técnicas psicológicas u otras. Tanto las Reglas penitenciarias europeas (Regla 65) como el Pacto de Nueva York (art. 10) establecen el principio de dignidad de las personas como límite crítico de la intervención con los privados de libertad, principio sólo presente en las legislaciones penitenciarias italiana, sueca, y española. Conectada con la cuestión de los límites se encuentra también la presencia o ausencia de un sistema de premios para incentivar la participación de los sujetos en el tratamiento, recogido tan sólo en las leyes francesa y española.

En lo tocante a las modalidades y grados del tratamiento, nuestro concepto de grado de tratamiento -correspondiente a un determinado régimen carcelario- resulta poco trasladable a otras legislaciones, distintas de la española. En Italia las modalidades son establecidas en el reglamento interno de cada centro. Mientras que en Holanda éstas dependen de si se trata de un centro con o sin tratamiento. Por lo que se refiere a los sistemas de progresión en el tratamiento, cabe establecer dos tipos: sistemas de progresión objetiva, determinada por el solo cumplimiento de plazos legales, y sistemas de progresión por individualización. Esta última tiene lugar en Alemania y Suecia, mientras que el sistema español es una mezcla de ambas modalidades. Esto es, la progresión en el tratamiento viene mediatizada por

el transcurso de unos determinados plazos temporales (la mayoría de ellos, proporcionales a la duración de la condena), pero básicamente depende del resultado de la evaluación periódica de cada sujeto.

Paralelamente, algunas legislaciones establecen criterios en lo relativo a la aplicación de situaciones de semilibertad (más o menos semejantes a nuestro régimen abierto). En Francia, se rigen por parámetros básicamente temporales. En Suecia y Alemania dependerían más de las características del recluso y de la evolución del mismo en el tratamiento. En Italia la semilibertad sería el producto de los dos componentes anteriores, temporalidad y evolución en el tratamiento.

Finalmente, en todos los países analizados se prevé la liberación condicional anticipada, en relación con la evolución del individuo en el proceso rehabilitador.

2.3. Regulación legal del tratamiento y aplicación de programas de tratamiento

Formulemos ahora abiertamente una pregunta que con anterioridad hemos sugerido: ¿Propicia la mayor aplicación de programas de tratamiento el que en un país exista una ley que regule esta materia de modo más claro, amplio y moderno? Y más concretamente, para los objetivos de esta investigación, ¿debemos esperar encontrar un mayor número de informes sobre programas de tratamiento en países con una más clara orientación jurídica hacia la aplicación del mismo?

De todas las legislaciones analizadas en relación con su regulación del tratamiento penitenciario, la italiana, la belga, la holandesa, la alemana, la sueca y la española tal vez sean las que de modo más explícito se refieran a una conexión entre privación de libertad y tratamiento. Por el contrario, las regulaciones francesa y, especialmente, la inglesa/galesa, pueden ser consideradas las más parcas, e incluso contrarias, a la filosofía del tratamiento. Si, contestando a nuestras anteriores preguntas, formuláramos una hipótesis que estableciera una relación proporcional entre regulación del tratamiento y aplicación del tratamiento, deberíamos esperar en Italia,

Bélgica, Holanda, Alemania, Suecia y España el mayor desarrollo de programas de rehabilitación en Europa, y, por el contrario, en Francia e Inglaterra/Gales un estancamiento.

De acuerdo con nuestros datos, de los estudios que hemos recogido el mayor número de informes corresponden a programas de tratamiento aplicados en Gran Bretaña, Alemania, España y Holanda. Mientras que es mucho menor o inexistente el número de programas Italianos, Franceses, Belgas, y Suecos. En el caso de Suecia, una explicación puede residir en el desconocimiento de la lengua sueca por parte del autor, unido a la inexistencia de revistas suecas en nuestras bibliotecas, lo que ha determinado, sin duda, un sesgo importante para la no selección de trabajos de aquel país, excepto cuando hubieran sido publicados en inglés. Una explicación semejante no es, sin embargo, válida para los posibles estudios italianos, franceses o belgas, que realmente no fueron localizados, pese a la revisión sistemática de bases de datos y revistas apropiadas.

Ciertamente, no se produce, como podemos ver, una relación tan simple como la que hemos hipotetizado. Antes que de la sola regulación formal del tratamiento, las iniciativas de aplicación de programas de tratamiento que, posteriormente, se vierten en estudios evaluativos, dependerían de factores diversos. El primero de ellos quizá sea la presencia en un país de una determinada filosofía criminológica que priorice, en los ámbitos académicos, políticos, y de opinión pública, la rehabilitación y la prevención de la delincuencia sobre los conceptos de mero castigo y represión de la misma. Esta filosofía rehabilitadora se plasma a veces en leyes y otras no. Piénsese, por ejemplo, que la mayoría de nuestras más recientes ideas rehabilitadoras nos vinieron desde el mundo anglosajón, Estados Unidos y Gran Bretaña, países en los que no existe, sin embargo, una regulación explícita del tratamiento semejante a la que después se plasmó en nuestra Ley penitenciaria.

En segundo lugar, es muy importante, sin duda, el nivel de desarrollo que tienen en una determinada comunidad las disciplinas sociales encargadas del diseño, aplicación, y evaluación de los programas de tratamiento de la delincuencia. Entre éstas se encuentran la psicología, la criminología, la educación, el trabajo social, etc. Y especialmente significativo a este

respecto puede ser la madurez e implantación que estas ciencias -y sus respectivos profesionales- hayan logrado en el campo correccional. La presencia en las prisiones, centros de reforma juvenil, y otros servicios para delincuentes, de psicólogos, educadores, trabajadores sociales, etc., con una sólida formación, no asegura pero hace eventualmente posible el desarrollo de programas de tratamiento de la delincuencia.

Por otra parte, tiene una importancia especial en lo concerniente a la evaluación de programas de tratamiento el tipo de formación o adscripción teórica y metodológica que tengan los profesionales que los diseñan y ejecutan (terapia conductual, psicoterapia, terapia transaccional, educación, etc.). De cual sea esta adscripción y formación depende, en buena medida, el mayor o menor interés que aquéllos tengan en los aspectos evaluativos y, por ende, en la publicación de informes sistemáticos sobre el desarrollo del programa.

Por todo lo expuesto, como en tantas otras cuestiones, tampoco aquí se observa una relación simple y directa entre la regulación jurídica del tratamiento de la delincuencia que tiene un país, y el grado de aplicación de dichos programas en el mismo.

RESUMEN

Hemos dirigido nuestra atención, en el presente capítulo, a la regulación jurídica del tratamiento de la delincuencia en Europa. Irremediamente, las normas son un referente necesario de la temática abordada. En primer lugar nos hemos referido a las recomendaciones del Consejo de Europa sobre el tratamiento de los delincuentes, que tienen, sin duda, un tono progresivo y renovador, procurando objetivos de dignidad, intervención técnica para dotar a los delincuentes de mejores oportunidades futuras, favorecimiento laboral, promoción del régimen abierto, etc.

Posteriormente hemos efectuado un repaso sucinto de las legislaciones penitenciarias de países de nuestro entorno. Hemos constatado como son Italia, Bélgica, Holanda, Alemania, Suecia y España, los países con un mayor desarrollo legal de ideas rehabilitadoras, mientras que Francia e Inglaterra

tendrían, por el contrario, una menor concreción normativa en esta temática. Pese a ello, hemos acabado constatando la no existencia de una relación directa entre desarrollo normativo y real aplicación de programas de tratamiento. Ésta parece depender más bien de factores diversos, como la filosofía social sobre la delincuencia, la madurez académica y profesional de las ciencias sociales conectadas con nuestra temática -psicología, criminología, educación...-, y sólo parcialmente de aspectos jurídicos.

3. LAS TÉCNICAS DE REVISIÓN DE LA LITERATURA

En la actualidad son muchas las revistas científicas existentes en cualquier campo de conocimiento. Así sucede también en criminología, psicología, sociología, y demás ciencias relacionadas con la delincuencia y su tratamiento. En muchas de las revistas especializadas, dentro de estos ámbitos, así como en libros e informes, aparecen múltiples estudios relativos a un determinado tema criminológico: causas de la delincuencia, prominencia de un determinado factor delictivo (edad, consumo de drogas, estructura familiar, aprendizaje, etc.), teorías criminológicas, y otros. Esto hace que cada sector de investigación se haga, paulatinamente, más voluminoso, complejo y disperso. Por ello, cuando el desarrollo de un campo es suficientemente importante, se requiere una revisión e integración sistemática de todos los conocimientos acumulados en el mismo, que permita "resumir, interpretar y reducir los resultados de los estudios primarios" (Sánchez-Meca, 1990b).

Para revisar la literatura ha habido tres caminos básicos (Gómez, 1987; Sánchez-Meca y Ato, 1989; Whitehead y Lab, 1989; Cooper, 1989):

1-El método de urna o recuento de votos: se trataría de recolectar el mayor número posibles de evaluaciones, en un determinado ámbito -de programas de tratamiento de la delincuencia, en nuestro caso-, que reúnan determinados criterios básicos (por ejemplo, comparación de grupos o medidas pre-post). O en definición de Gómez (1987) el método de urna pretendería "resumir los resultados de los estudios en función de su significación estadística ... positiva, ... negativa o no existencia de relación ...", y contabilizar "el número de estudios que caen dentro de cada

categoría, aceptándose como la mejor estimación aquella que obtiene más 'votos'" (p.21-22).

Los metodólogos, entre ellos los promotores del meta-análisis, han señalado las dificultades que comporta el método de votos, que han sido resumidas por Gómez (1987): la primera es que no existe un criterio objetivo de representatividad modal de la muestra, sino que la selección quedaría al arbitrio personal del revisor; por otro lado, prescinde de información descriptiva importante; además, no nos informa de en qué medida un tratamiento es superior a otro; también, sesga los resultados de conjunto en favor de los estudios con muestras más numerosas, que, por ello, hipervaloran la significatividad, frente a los de muestras más pequeñas; por último, el método de votos sesgaría favorablemente los estudios significativos que, como veremos más adelante, serían los más publicados.

2-El método analítico: implica ir más allá de las meras tablas de datos, para valorar críticamente la teoría, metodología y aplicación de diversas intervenciones, con objeto de extraer principios comunes o sus descubrimientos más sustanciales. Una variante consiste en seleccionar y analizar intervenciones ejemplares de diversos tipos (correspondientes, por ejemplo, a distintos modelos teóricos). Un problema importante, de no fácil solución, es ¿qué proyectos seleccionar como paradigmas que representen adecuadamente a los modelos o constructos teóricamente estudiados?

3-El más moderno método de revisión de la literatura científica es el meta-análisis. Se trata de una extensión perfeccionada del método de urna, que recoge e integra las evaluaciones de un sector científico, correspondientes a un determinado período temporal, y obtiene para todas ellas un estadístico común, que permite comparar la magnitud de sus efectos.

3.1. Revisiones cualitativas y meta-análisis

Hasta la década de los setenta y principios de los ochenta las revisiones de la literatura científica tenían prioritariamente un cariz narrativo, en donde el revisor resumía información seleccionada según su criterio. Entre los principales inconvenientes de estas revisiones narrativas se encuentran,

según Okun et al. (citados por Gómez, 1987), los siguientes: (1) implican una selección subjetiva de los estudios realizada por el revisor según su propio criterio de calidad de los mismos; (2) existe una ausencia de información sobre las magnitudes de los efectos producidos; (3) también comportan la carencia de información sobre posibles variables moduladoras; y (4) adolecen, con frecuencia, de falta de información metodológica sobre las muestras, medidas y estadísticos empleados.

Así pues, hasta muy recientemente, los investigadores que deseaban examinar un campo de conocimiento en psicología y otras ciencias debían recoger la literatura disponible en tal área de conocimiento y llevar a cabo un tedioso repaso de la misma, que, finalmente, no podía concretarse sino en un relato más o menos acertado (o desafortunado) de los principales resultados obtenidos por las investigaciones consultadas. Con demasiada frecuencia las conclusiones de un investigador entraban en abierta colisión con las obtenidas por sus colegas, incluso en la revisión de un conjunto semejante de trabajos. Nos estamos refiriendo a las revisiones cualitativas. Frente a ellas, el meta-análisis, que lleva a cabo una revisión cuantitativa, representa un avance, sin duda, ya que supone un acercamiento táctico mucho más riguroso a la integración de estudios evaluativos (Johnson, 1989). Para ello emplea los mismos procedimientos de análisis utilizados en sus propios desarrollos por las investigaciones primarias que son objeto de revisión, de lo cual pueden extraerse conclusiones más potentes, en términos de veracidad científica. Técnicamente, "lo único que cambia en la aplicación de métodos estadísticos del estudio primario a la revisión cuantitativa es el dato o unidad experimental: Mientras que en las investigaciones primarias la unidad de análisis es habitualmente la observación de cada sujeto, en las revisiones cuantitativas lo es un índice descriptivo de la efectividad del programas obtenida en cada estudio individual" (Sánchez-Meca, 1990b).

Según han puesto de relieve Fitz-Gibbon y Morris (1988) el meta-análisis se focaliza en la pregunta a la que realmente se desea contestar cuando se realiza una investigación evaluativa de programas, ¿qué grado de diferencia produjo una intervención? (en relación a otras intervenciones), en lugar de en aquella otra cuestión a la que, sin que sea preguntada, generalmente suele responderse, que a la postre resulta ser: si la hipótesis nula fuera cierta, ¿cual es la probabilidad de que estos resultados se hayan obtenido

por azar? Esta fijación en la significación estadística ha confundido tanto a la gente que Carver (citado por Fitz-Gibbon y Morris, 1988) se ha referido a ella como "una forma corrupta del método científico" (p.132).

Yendo aún más lejos, utilizamos el meta-análisis porque en investigaciones como la presente, en que existen intervenciones semejantes realizadas sobre distintos grupos y contextos, necesitamos contestar a tres preguntas básicas (Fitz-Gibbon y Morris, 1988): (1) para grupos y lugares distintos, ¿cuáles fueron los Tamaños del Efecto?; (2) ¿tiene la intervención un efecto consistente u homogéneo?; y (3) ¿existen características de los lugares o grupos en los que los efectos fueron distintos del efecto promedio?

Entre las razones de la utilidad del meta-análisis Fitz-Gibbon y Morris (1988) han encontrado las siguientes: en primer término, dado que muchas evaluaciones son replicaciones de una misma intervención en diferentes grupos y contextos, el meta-análisis constituye un camino claro y simple para resumir datos correspondientes a estos diversos grupos y contextos. Además, debido a que muchas intervenciones sociales (en educación, salud, terapia, prisiones, etc.) son aplicadas a grupos reducidos, suele obtenerse un nivel de significación bajo, lo que, a menudo, resulta descorazonador. La técnica meta-analítica, por el contrario, nos ayuda a evitar la confusión entre resultados de una intervención y tamaño muestral, ya que es éste, más que los resultados, el que determina el nivel de significación.

Pese a ello, el análisis cuantitativo de diversos estudios, realizado mediante procedimientos meta-analíticos, no debe excluir tanto el desarrollo de las necesarias sutilezas teóricas en el proceso de análisis, cuanto la detección y aprehensión de los matices y riqueza que los diversos trabajos de investigación puedan contener (Johnson, 1989). Así lo han puesto de relieve Masters, Cantor, y Hedges (citados por Johnson, 1989) que han llamado la atención sobre el peligro que pueden entrañar las revisiones meta-analíticas rutinarias, en estos términos: "Una revisión con un cuidadoso meta-análisis pero con una conceptualización débil o irrelevante tiene poco que ofrecer para el avance o la consolidación del conocimiento" (p.5). Por ello, el propio Johnson (1989), mentor de la sistematización y tecnificación de los

procedimientos meta-analíticos para los que ha diseñado el DSTAT³, ha sugerido la conveniencia de tomar tanto de las técnicas cuantitativas de revisión de la literatura -el meta-análisis- como de los procedimientos cualitativos, lo mejor que cada uno de ellos tiene, integrando, a su vez, ambos procesos en el decurso de una revisión científica.

En efecto, el meta-análisis no puede convertirse en algo mecánico. Su utilización presupone una serie de requisitos para el investigador (Gómez, 1987): éste debe plantearse preguntas claras, precisas y explícitas, y debe acometer la revisión de estudios o investigaciones que impliquen una hipótesis conceptual común. En nuestro caso, existe una pregunta inicial muy amplia, aunque necesaria, dado el desconocimiento existente sobre los programas correccionales en Europa. Ésta sería: ¿que programas de tratamiento de la delincuencia se han aplicado y se han evaluado en los países europeos? Cuestiones más precisas son: ¿en qué países?, ¿a qué sujetos se han dirigido?, ¿qué modelos teóricos se han utilizado para generar los programas?, ¿en qué han consistido las intervenciones?, ¿quiénes las han aplicado?, ¿cómo se han evaluado?, y, en suma, ¿qué tratamientos han resultado más efectivos?

Ciertamente, todos los estudios objeto de nuestro análisis tienen una hipótesis conceptual común, que nos sirvió como criterio selectivo de los mismos: todos consistieron en programas destinados al tratamiento de la delincuencia.

3.2. La técnica meta-analítica

El meta-análisis -también llamado investigación integrativa, evaluación cuantitativa de dominios de investigación y revisión cuantitativa (Gómez, 1987)- "es un conjunto de técnicas para resumir los descubrimientos de diversos estudios" (Glass, McGaw y Smith, 1981). En realidad, viene constituido por una serie de técnicas estadísticas que permiten la integración de investigaciones en un ámbito o campo determinado, con la finalidad de lograr una síntesis de la efectividad de un conjunto de programas de acción,

³DSTAT: Software for the Meta-Analytic Review of Research Literatures.

haciendo posible explicar los desacuerdos científicos acumulados a lo largo del tiempo (Johnson, 1989; Sánchez-Meca, 1990b). El meta-análisis permite, según Lösel (en prensa), la integración del "estado de conocimiento científico sobre la materia en una manera más comprensiva, sistemática, imparcial y diferencial, usando para ello procedimientos estadísticos" (p.3). Implica la transformación a un métrica común de los resultados de diversas investigaciones, de modo que sea posible su comparación y/o combinación, disminuyendo la subjetividad de las revisiones cualitativas (Gómez, 1987). O como han definido Whitehead y Lab (1989) "el meta-análisis es una técnica que reanaliza los datos encontrados en un informe de investigación original y obtiene una medida común para todos los estudios" (p.276), lo que permite comparar los distintos resultados. Es un "análisis de los análisis" (Gómez, 1987).

Para ello el meta-análisis examina estudios distintos con grupos o medidas de comparación diversas. Para homogeneizar los resultados de los diferentes estudios las dos medidas más habituales han sido el análisis de los niveles de significación -que comportan una gran dependencia de los tamaños muestrales- y la obtención de los tamaños del efecto, que permiten conocer la magnitud de la relación entre las variables estudiadas (Gómez, 1987). Puede ser calculado un TAMAÑO DEL EFECTO (a partir de ahora TE) para cada estudio con arreglo a la siguiente fórmula (el llamado índice de Cohen o diferencia media tipificada):

$$TE = [(X \text{ grupo E}) - (Y \text{ grupo C})] / Sd \text{ conjunta}$$

donde TE es el tamaño del efecto, X e Y las puntuaciones medias de los grupos experimental (E) y control (C), y Sd la desviación típica conjunta de ambos grupos. Se trata, pues, de la diferencia entre las puntuaciones medias del grupo de tratamiento y del grupo control, dividida por la desviación típica conjunta (Gómez, 1987). Lo que produce un valor positivo si el grupo de tratamiento obtuvo una puntuación superior al control y negativo si fue al revés. Este índice mide el número de unidades de desviación típica en que los grupos difieren entre sí en cada una de las variables resultado consideradas, según la formulación meta-analítica básica de G.V.Glass y

R.L.Hedges (Glass, McGaw, y Smith, 1981).

Más allá del cálculo de los TEs globales, tal vez la virtualidad más importante del meta-análisis es que permite efectuar un "análisis diferencial" de los efectos en función de diversas variables, ya sean sustantivas de los programas (tipo de tratamiento, sesiones aplicadas, contexto, etc.), ya sean metodológicas (Sánchez-Meca, 1990b; Lösel, en prensa). Además, debido a sus características, el meta-análisis resulta especialmente adecuado para estudiar el impacto de programas de intervención, ya que permite obtener índices globales de los TEs de conjuntos de tratamientos semejantes (Sánchez-Meca, 1990b).

Los pasos principales del meta-análisis pueden resumirse en los siguientes (Glass, 1976, 1977; Sánchez-Meca, 1986; Sánchez-Meca y Ato, 1989; Cooper, 1989; Lösel, en prensa):

- (1) Delimitación precisa de la cuestión que se estudiará.
- (2) Selección sistemática de los estudios primarios que serán analizados.
- (3) Codificación de los estudios en consonancia con diversas categorías (tipo de tratamiento, características de los sujetos, criterios resultado, etc.).
- (4) Evaluación de la calidad metodológica de los estudios.
- (5) Cálculo de medidas estandarizadas del TE.
- (6) Y análisis de las relaciones entre características sustantivas o metodológicas de los diversos estudios y los TEs obtenidos.

Según su inventor, G.V. Glass, el meta-análisis como técnica de revisión comporta tres características distintivas (Gómez, 1987):

1. Es un método cuantitativo, que permite organizar y reanalizar información procedente de grandes masas de datos.
2. No prejuzga la calidad del diseño de investigación, sino que convierte a

ésta en una cuestión empírica, cuya influencia sobre los resultados debe ser analizada también.

3. Persigue conclusiones generales. Esto necesariamente implica atender a las características esenciales comunes a los diferentes estudios e ignorar otras que les hacen distintos: "... el conocimiento sólo es posible a través de la exclusión ordenada de información" (Glass et al., citados por Gómez, 1987, p.27).

Gómez (1987) ha sintetizado las funciones y limitaciones del meta-análisis, a partir de los trabajos de Glass et al. (1981) y Green y Hall (1984). Entre las funciones destacan:

-Una función descriptiva, que permite sistematizar las características sustantivas y metodológicas de grupos de estudios dentro de una determinada área de investigación. El conocimiento de las primeras permitiría progresar en la aplicación de conocimientos teóricos a problemas concretos. El análisis de las segundas serviría para que el investigador primario pueda repasar, con facilidad, la situación de un determinado ámbito (diseños, estadísticos, etc.).

-Una función diagnóstica, que permitiría descubrir puntos oscuros en un sector de conocimiento, propiciando la investigación en el mismo.

-Una función predictiva, que favorecería la proposición de nuevas hipótesis de trabajo, que no habían sido anteriormente descubiertas o probadas por los estudios primarios que se analizan.

Entre sus principales limitaciones se encontrarían las siguientes (Gómez, 1987):

-El meta-análisis no permite corregir las carencias de información existentes en los estudios primarios, por lo que sus resultados y conclusiones vienen necesariamente determinados por éstos.

-Un meta-análisis se encuentra muy determinado por la selección de características que son codificadas por el autor, y éstas pueden ser muy

diferentes para distintos autores.

-Requiere un gran esfuerzo de codificación, cuando se cuenta con muchos estudios en un ámbito, y, además, pueden plantearse problemas de eficacia y consistencia de las codificaciones.

-Por último, existe un punto crítico de debilidad, constatado por D. Glass, según el cual las características de los trabajos sólo permiten predecir un porcentaje bajo, menor del 25%, de los resultados.

Con la finalidad de auxiliar y simplificar el proceso estadístico meta-analítico Johnson (1989) ha diseñado un paquete informático denominado Software for the Meta-Analytic Review of Research Literatures (DSTAT), que contiene las técnicas estadísticas más usuales en meta-análisis. Nos serviremos de él para el análisis de nuestros resultados.

3.3. Principales fuentes de sesgo en el meta-análisis

Gómez (1987) ha sistematizado las principales fuentes de sesgo que pueden darse en un meta-análisis. Son las siguientes:

1- Sesgo muestral. Sin duda uno de los peligros más importantes que nos acechan es que los estudios que hemos seleccionado no representen a la población de informes en la materia analizada. Especial importancia tiene aquí el eventual sesgo de publicación, que pudiera derivarse de incluir sólo trabajos publicados. Para mejorar la representatividad de los estudios incluidos en un meta-análisis se recomienda el empleo de distintos procedimientos de búsqueda y la inclusión del mayor número posible de estudios no publicados (Sánchez-Meca y Ato, 1989; Lipsey, 1992a; Lösel, en prensa), criterio que hemos procurado atender en nuestra investigación.

2- Pérdida de información. A este sesgo puede conducirnos el afán de síntesis implícito en el meta-análisis, al resumir gran cantidad de información en los TEs o en los niveles de significación. Para salvaguardar este sesgo se aconseja tomar medidas de homogeneidad/heterogeneidad de los estudios y, consiguientemente, identificar eventuales fuentes de variación asociadas a

las diversas características, tanto sustantivas como metodológicas, de los trabajos.

3- Uniformidad de los estudios. Hace referencia al viejo problema de las "manzanas y las naranjas", esto es a la posibilidad de que estemos mezclando cosas diferentes, no comparables. Aunque, según ha señalado D. Glass este problema no es distinto al que se enfrente el investigador primario cuando incluye en sus análisis a sujetos diferentes, sólo homogéneos en algunas de sus características. La codificación nos permitirá un tratamiento empírico de esta problemática.

4- Calidad de los estudios. Una cuestión importante en el meta-análisis es si deben ser eliminados o no aquellos estudios que presentan defectos metodológicos serios. La constatación de D. Glass es que no existe una relación importante entre tamaño del efecto y calidad metodológica de los estudios. No obstante, se recomienda realizar una ponderación de la calidad de los estudios, definiéndola en el manual de codificación, así como el que la codificación de los aspectos metodológicos sea efectuada por metodólogos expertos. Ambos criterios han sido también considerados en nuestro análisis, según comentaremos más adelante.

5- Independencia. Esta fuente de sesgo puede aparecer ya sea por la interrelación existente entre los diversos TEs calculados dentro de un mismo estudio, ya sea por la dependencia de las medidas de estudios similares entre sí porque, por ejemplo, tienen muestras interdependientes. Dentro de los meta-análisis correccionales efectuados hasta la fecha, éste fue el caso del trabajo de Lösel y Köferl (1989), algunos de cuyos estudios en prisiones alemanas eran producto de muestras integradas en parte por los mismos individuos. Para prevenir el sesgo de dependencia se ha propuesto hallar un estadístico promedio de los diversos TEs dentro de cada estudio o elegir aquél de ellos que se considere más importante.

3.4. La técnica costo-efectividad

Una cuestión de gran trascendencia en la justicia y la rehabilitación correccional se refiere, según hemos ya comentado, a los costos de las medidas que son aplicadas en relación con la efectividad que consiguen, de acuerdo con sus propósitos. Este problema es el que nos permiten resolver las recientes técnicas de análisis costo-efectividad.

Los análisis costo-beneficio y costo-efectividad tienen una consolidada tradición en sectores económicos, en donde las empresas, con ánimo de lucro, ponderan las ganancias que les producen sus inversiones; en suma, el aumento, reducción o reorientación de sus acciones productivas guardan una estrecha relación con los resultados que obtienen en términos de beneficios. Más recientemente, los análisis costo-efectividad están llegando a sectores sociales de ingente envergadura económica como, por ejemplo, la sanidad. Sin embargo, son casi inexistentes los estudios que hayan acometido esta cuestión dentro del marco de la justicia penal.

Una excepción a ello la constituye un reciente análisis efectuado por Taylor (1992) sobre la efectividad y eficiencia de la educación universitaria con delincuentes en Norteamérica. Taylor ha revisado un amplio elenco de estudios sobre efectividad de la educación universitaria con delincuentes, que analiza en relación con los bajos costos que ésta comporta, contrariamente a los onerosos presupuestos públicos del mero encarcelamiento. Sus conclusiones principales son que los programas de educación universitaria tienen una documentada mayor efectividad rehabilitadora, produciendo reducciones muy importantes en las tasas de reincidencia (en estudios entre 1974-1991); mayor efectividad para el empleo de muchos delincuentes, que una vez formados encontrarían un puesto de trabajo con más facilidad; efectividad institucional, reduciendo los problemas disciplinarios dentro de las cárceles, y aumentando la comunicación entre internos y personal; y eficiencia-costo, ya que la educación universitaria dentro de las cárceles tendría un costo muy inferior al que tiene en las universidades ordinarias, debido a que sólo se requieren pequeños gastos añadidos -libros, material, tutores-, ya que los grandes gastos de las instituciones correccionales tienen lugar igualmente, con o sin programas de educación dentro de ellas.

Los meta-análisis de programas de tratamiento de la delincuencia que nos han precedido no han abordado tampoco, en general, el problema del costo-efectividad de los programas. El único meta-análisis que hace una somera referencia a esta cuestión es el de Lösel y Köferl (1989), referido a programas de Socioterapia aplicados en prisiones alemanas. Lösel y Köferl (1989) mencionan que una crítica frecuente de los programas de rehabilitación en las prisiones es que su costo es superior al de las prisiones ordinarias. Para el caso alemán, estos autores cifraban el costo de una plaza en una prisión ordinaria en 101-109 marcos alemanes por día y preso, mientras que el de las prisiones socioterapéuticas se estimaba por encima de tres veces la anterior cantidad. Incluso en un supuesto de este tipo, Lösel y Köferl (1989) postulan que en este problema deben entrar valoraciones relativas a la inversión que los programas suponen en futuro -en eventuales reducciones de las tasas de reincidencia-; así pues, en su opinión, deberían entrar en el cálculo de costes de las prisiones ordinarias los períodos de prisión de quienes repiten condenas, los costes de la detección de los delitos, de los tribunales de justicia, los daños materiales y personales de los delitos, y los costes sociales que todo ello comporta para víctimas, familias y delincuentes. No son pequeños los problemas metodológicos que implica la estimación de todos estos parámetros.

Taylor (1992) ha ensayado el cálculo de los costes de la reincidencia en el delito en Estados Unidos. Para ello ha estimado el ahorro que podría operarse mediante la aplicación de programas de educación, a partir de las cifras conocidas del costo del encarcelamiento y de las tasas de reincidencia promedio para distintos grupos de sujetos. El cálculo lo ha efectuado con arreglo a la siguiente fórmula (véase Figura 3.1):

Figura 3.1. Fórmula del costo de la reincidencia

Costo anual del encarcelamiento/preso	\$25,800
Tasa promedio de reincidencia general por cada 100 liberados de prisión	x <u>50</u>
Costo total anual de la reincidencia para cada 100 delincuentes liberados	\$1,290,000
	\$25,000
Tasa promedio de reincidencia para quienes recibieron educación univer- sitaria	x <u>15</u>
Costo total anual de la reincidencia para cada 100 delincuentes que recibieron educación universitaria	\$387,000
	\$1,290,000
	<u>-387,000</u>
Ahorro anual a los contribuyentes generado por los programas de educación universitaria con delincuentes	\$930,000

Fuente: Tomado de Taylor (1992).

Muy probablemente, el análisis costo-efectividad no sea tan sencillo como de la propuesta de Taylor (1992) pudiera derivarse. Sin embargo, no cabe duda que su cálculo constituye una propuesta meritoria y necesaria en materia de delincuencia. En la mayoría de los países, los ciudadanos no tienen ni la más

remota idea de los miles de millones que cuesta su justicia. Lo mismo sucede, en buena medida, con las diversas suertes de "expertos" en esta materia - jueces, fiscales, penitenciarios, criminólogos, investigadores, etc.-. Desgraciadamente, nuestros poderes públicos informan con gran parquedad sobre estas cuestiones, lo que dificulta enormemente un análisis minimamente riguroso. Hoy por hoy, resulta imposible en España contestar a cuestiones de costo-efectividad como: ¿cuánto cuesta investigar un delito?; ¿cuánto cuesta detener a un delincuente?; ¿cuánto cuesta la instrucción de un sumario?; ¿cuánto cuesta una medida de encarcelamiento, y cuáles son las diferencias entre si éste tiene lugar en régimen cerrado o abierto?; ¿cuándo costaría o cuánto se ahorraría si se aplicaran medidas alternativas al encarcelamiento?, etc.

Una intención inicial de nuestro análisis era estudiar el costo-efectividad de los programas que recogieramos. Enseguida nos hicimos conscientes de la imposibilidad de esta pretensión por el momento. La práctica totalidad de los programas seleccionados carecen de información alguna sobre los costos de su aplicación. Su averiguación requeriría una metodología distinta, directamente encaminada a obtener información extra de parte de los autores de los estudios o de las administraciones correspondientes, lo que trasciende con creces nuestros actuales propósitos.

3.5. El empleo de la técnica meta-analítica en España

En España, muy pocos trabajos se han servido del meta-análisis para revisar alguna parcela de investigación. Aunque no guardan relación temática con nuestro estudio, sin embargo, por su metodología, tienen un valor paradigmático para nosotros las investigaciones realizadas por J. Sánchez-Meca, de la Universidad de Murcia, quien comparte con M.T. Anguera y J. Gómez el mérito de haber sido pionero e impulsor en España de los estudios meta-analíticos. Inicialmente, Sánchez-Meca (1985a) meta-analizó una treintena de estudios sobre déficit perceptivo del retraso específico en lectura, con el objetivo de ver en qué medida los déficits en la memoria icónica -que permite el procesamiento rápido de la información- aparecían implicados en este síndrome lector. Sus resultados confirmaron que, en los estudios revisados, los niños con dificultades lectoras eran más lentos en el

procesamiento de la información visual que lo niños "buenos lectores".

Con posterioridad, Sánchez-Meca (1985b, 1990a) analizó mediante la técnica meta-analítica la consistencia entre la concepción tradicional de la memoria icónica, que interpreta ésta como "persistencia de la información", y la evidencia empírica acumulada a lo largo de veinticinco años de investigación. Para la localización de la literatura utilizó cuatro procedimientos paralelos: la consulta a investigadores expertos en el tema, la inspección de la bibliografía, la consulta del Psychological Abstracts, y del Current Contents: Social & Behavioral Sciences, entre 1980-85, y una búsqueda computerizada en el Dialog System. Pudo localizar 74 estudios. Para la codificación de los trabajos el autor elaboró una lista de 13 características, que fueron definidas en un "Libro de Codificación". Tres categorías tenían un cariz metodológico (fecha de publicación, tamaño muestral y fuentes de publicación), mientras que las 10 restantes correspondían a características sustantivas de la temática de investigación. En sus resultados, Sánchez-Meca ofreció información relativa a la fiabilidad inter-codificador, así como sobre la relación existente entre los diferentes factores codificados y los tamaños del efecto obtenidos.

Otros meta-análisis realizados en España son los siguientes: el de Torres (1988) -vinculado a los anteriores-, en el que se evaluaron 43 estudios para investigar la relación existente entre memoria a corto plazo y proceso de recodificación en dicha memoria, concluyendo que existe relación, empíricamente documentada, entre la menor capacidad de la memoria a corto plazo de algunos niños y sus dificultades lectoras. Mateo (1985) efectuó un meta-análisis sobre estudios de rendimiento académico en España. También en el ámbito académico, Loscos (1985) efectuó un meta-análisis sobre la predicción del rendimiento escolar. Marmolejo (1990) ha comparado la efectividad diferencial de distintos métodos para el aprendizaje de la lectura. Velandrino y Sánchez-Meca (1990, noviembre) han explorado la relación existente entre inteligencia y tiempo de inspección, o velocidad de respuesta de los sujetos en tareas simples, meta-análisis en el que han concluido que la heterogeneidad observada se explica más por variables metodológicas y aparataje estadístico de las mediciones, que por componentes sustantivos. Finalmente, en lo tocante a los ámbitos psicológico-educativos, Vázquez et al. (1991, septiembre) han meta-analizado investigaciones sobre estado de

ánimo y memoria. Dentro del campo clínico, Gómez et al. (1989) han investigado la influencia del consumo de tabaco sobre la génesis de disfunciones estomacales. Por su parte, Gómez Olmedo et al. (1990) y Fernández Sierra et al. (1990) han efectuado un meta-análisis de las tendencias temporales y geográficas de la infección por el virus de inmunodeficiencia humana en la población española.

El sucinto repaso que hemos efectuado de los meta-análisis realizados en España nos da una idea de la limitada aplicación que, hasta el presente, se ha hecho de esta técnica, así como, dentro de la escasez, las temáticas a que se ha dirigido: ámbitos psicológicos, como procesamiento de la información, lectura, tiempos de reacción, y rendimiento académico, y ámbito médico, con un estudio sobre tabaco y úlceras pépticas y otro sobre infección por VIH. Ninguno de ellos ha acometido el análisis de problemas relacionados con la delincuencia, que es nuestro propósito.

RESUMEN

En este tercer capítulo nos hemos referido brevemente a las principales técnicas existentes para la revisión de la literatura científica. Hasta recientemente las revisiones tenían un carácter cualitativo. En ellas, los expertos realizaban un compendio, más o menos acertado, de muestras seleccionadas de estudios sobre una materia, contabilizando los resultados a favor o en contra de una determinada hipótesis -método de urna o recuento de votos- o analizando críticamente las teorías o metodologías utilizadas -método analítico-. Se han señalado distintas dificultades inherentes a la metodología cualitativa, entre las que destacan la selección subjetiva de estudios que realizan, y su carencia de información cuantitativa relacional, que permita extraer conclusiones sobre la interacción existente entre factores implicados -muestras, teorías, técnicas y evaluación- y los resultados obtenidos.

Frente a ello, la técnica meta-analítica permite la revisión cuantitativa de la literatura. Su dos méritos fundamentales son que sintetiza los resultados de todo un ámbito de investigación en términos numéricos -mediante la obtención

de los Tamaños del Efecto o TEs-, y permite referir los resultados obtenidos a las diversas fuentes de variación implicadas -sujetos, tratamientos, contextos de aplicación, y metodología-. De ahí su especial valor para la evaluación de programas, como es nuestro caso.

Hemos comentado también la necesidad existente en el campo correccional de efectuar análisis costo-efectividad, para averiguar qué programas con delincuentes obtienen mejores resultados a más bajo coste. Hemos ilustrado esta técnica con una evaluación realizada en los Estados Unidos sobre la efectividad-costo de programas universitarios que se desarrollan en las cárceles. Desgraciadamente, la perspectiva costo-efectividad nos queda vedada por el momento, a causa de la ausencia en los estudios de la información económica necesaria para su análisis.

Por último, hemos relacionado la decena de trabajos meta-analíticos llevados a cabo en España hasta la fecha. Han sido estudios dirigidos, sobre todo, al campo psicológico-educativo, en temáticas como el procesamiento de la información, la lectura y el rendimiento académico, y, en dos casos, a cuestiones clínicas: la relación tabaquismo-úlceras, y el Sida.

4. LAS REVISIONES PRECEDENTES: TRES META-ANÁLISIS PARADIGMÁTICOS

En tres capítulos sucesivos repasaremos críticamente las revisiones y meta-análisis de programas correccionales que nos han precedido. En el presente capítulo nos referiremos con detalle a tres de los cuatro últimos meta-análisis aparecidos. El primero en nuestra referencia, de Whitehead y Lab (1989), consistió en el análisis de 50 programas norte-americanos, aplicados con delincuentes juveniles entre 1975 y 1984. Obviaremos ahora el comentario de la revisión de Andrews et al. (1990), ya que, en buena medida, se trata de un nuevo análisis de los mismos programas estudiados por Whitehead y Lab (1989). En segundo lugar nos detendremos en el único meta-análisis de programas correccionales llevado a cabo en Europa, desarrollado por Lösel y Köferl (1989) (informe original en alemán de Lösel, Köferl y Weber, 1987) sobre un conjunto de 16 programas penitenciarios alemanes, que son todos los que fueron publicados en el intervalo temporal 1977-1985. Y finalmente, nos extenderemos ampliamente sobre el último y más ambicioso de los meta-análisis americanos de programas de tratamiento de la delincuencia, llevado a cabo por Lipsey (1992a, 1992b), que ha revisado hasta ahora nada menos que 443 estudios, publicados entre 1950-1990. Sus análisis y procedimientos serán en buena medida paradigmáticos para nuestra propia investigación.

4.1. La revisión de J.T. Whitehead y S.P. Lab (Lab y Whitehead, 1988; Whitehead y Lab, 1989).

Objetivos. Whitehead y Lab (1989) recogieron y evaluaron los programas de tratamiento de jóvenes delincuentes que habían sido publicados en revistas profesionales -Americanas- entre 1975-1984. Seleccionaron para su meta-

análisis aquellos programas que se referían a alguna forma de tratamiento correccional juvenil que, desde una perspectiva metodológica, incluían en su diseño al menos un grupo control, y sus datos aparecían en forma susceptible para el meta-análisis.

Método. Como variables resultado los autores tomaron en consideración exclusivamente la "reincidencia". Ésta no aparecía de una manera uniforme, sino que los distintos trabajos incluían distintas formas de medida de la reincidencia; todas ellas reflejaban, ciertamente, alguna evaluación de los contactos subsiguientes de los jóvenes con el sistema de justicia a causa de algún comportamiento delictivo.

Para el análisis comparativo de los distintos resultados, Whitehead y Lab (1989) utilizaron, inicialmente, dos estadísticos: chi-cuadrado y Phi (0-1); aunque, finalmente, el mayor énfasis interpretativo fue concedido a las puntuaciones Phi. Como criterio, se consideró que habían tenido alguna efectividad aquellos programas que, en la comparación grupo experimental/grupo control, habían logrado una puntuación $\Phi > .20$, mientras que se atribuyó gran efectividad si $\Phi > 0.40$. Si la puntuación Phi resultaba negativa se consideraba que el tratamiento había sido contraproducente.

Resultados. Los programas fueron clasificados en una serie de categorías, según los tipos de intervenciones administradas a los sujetos. A partir de esta clasificación, se obtuvieron los siguientes resultados:

(1) Programas de diversion o derivación de los sujetos fuera del sistema de justicia; se incluyeron aquí aquellos programas que interrumpían completamente cualquier contacto del joven con el sistema formal de justicia (policía, tribunales, centros de corrección, etc.), o le contactaban con alguna agencia o institución no perteneciente al sistema de justicia. En esta categoría se agruparon 15 programas. De ellos un 20% resultaron efectivos mientras que un 33% aparecieron como contraproducentes.

(2) Programas de diversion o derivación de los sujetos a instituciones que aunque formaran parte, como una extensión, del sistema de justicia, no eran, sin embargo, centros de reforma juvenil cerrados, etc., en el sentido

más tradicional. Aquí se catalogaron 15 programas. De ellos el 46% resultaron positivos, mientras que un 6% fueron contraproducentes.

(3) Cualquier intervención correccional comunitaria, tal como probation o parole. Se incluyeron 10 programas, de los cuales un 40% fueron efectivos mientras que un 30% se consideraron adversos.

(4) Programas de tratamiento institucional o residencial. En esta categoría aparecieron tan sólo 7 programas. De ellos el 14,3% resultaron favorables y el 42,8% contraproducentes.

(5) En una categoría denominada intervenciones novedosas o de especialidad se incluyeron 3 programas. Un 33% -1 sólo programa- se consideró como de alta efectividad, mientras que otro 33% -1 programa- se catalogó como contraproducente.

Whitehead y Lab (1989) efectuaron otras agrupaciones de los programas analizados. Las intervenciones conductuales obtuvieron una efectividad del 44%, resultando también contraproducentes en un recíproco 44%, mientras que las no conductuales se mostraron algo menos efectivas (el 35% de ellas) pero también algo menos contraproducentes (el 23,5%). Por el rigor en la asignación de los sujetos a los grupos, los estudios con una asignación aleatoria fueron efectivos en un 33%, mientras que los no-aleatorios resultaron eficaces en un 37%. En función del sexo de los sujetos, se encontraron diferencias mínimas entre los estudios. Finalmente, según la cronología del tratamiento, los estudios correspondientes a la primera parte de la década cubierta por el análisis (1975-1979) resultaron más efectivos (con un 63,6%) que los del segundo período (1980-1984), que lo fueron en un 22,7%.

La conclusión a la que llegan Whitehead y Lab (1989) es lacónica y desesperanzada: "Los resultados de este meta-análisis difieren poco de otras revisiones de la literatura. Los resultados apoyan con claridad las constataciones... [de algunos revisiones precedentes] en el sentido de que el tratamiento correccional tiene un escaso efecto sobre la reincidencia" (p.291).

El meta-análisis de Whitehead y Lab (1989) cuenta entre sus méritos principales su claridad analítica, y su especial atención a la variable reincidencia como factor decisivo del éxito de los programas. Sin embargo, son varias las críticas que pueden realizarse a la revisión efectuada por Whitehead y Lab (1989) y más explícitamente a sus conclusiones. Estos autores establecen arbitrariamente como criterio de efectividad -para considerar que un programa había tenido éxito- un valor de Phi $>.20$ (que equivaldría a un tamaño del efecto de alrededor de $.41$ unidades *sd*). Según Lipsey (1992a) y Palmer (1992) este criterio resulta exageradamente estricto, y condujo a Whitehead y Lab (1989) a una conclusión precipitada sobre la no efectividad de los programas. Frente a esta negativa conclusión, los propios autores reconocen que su meta-análisis obtuvo un TE promedio positivo de $.12$, que equivaldría a una diferencia de en torno a $.25$ unidades *sd* entre grupos de tratamiento y controles. Con estos resultados no puede deducirse la absoluta inefectividad de los tratamientos.

4.2. La revisión de programas alemanes de F. Lösel y P. Köferl (Lösel, Köferl, y Weber, 1987; Lösel, y Köferl, 1989).

Las prisiones socioterapéuticas. En comparación con los Estados Unidos y algunos países de la Europa Occidental, el ideal del tratamiento llegó bastante tarde a las prisiones de la Alemania Federal: en la reforma penal de 1969 se introdujeron las prisiones socioterapéuticas, cuyo objetivo eran los presos con severos trastornos de personalidad, los delincuentes sexuales peligrosos, los delincuentes jóvenes-adultos diagnosticados como especialmente propensos al delito, y los delincuentes inimputables o con la capacidad disminuida, cuyo internamiento en una prisión socioterapéutica pudiera favorecer sus oportunidades para la resocialización. Sin embargo, la aplicación de esta reforma fue pospuesta durante 20 años, que se dedicaron a la típica controversia doctrinal sobre el fundamento, funciones, justificación, etc., de las prisiones socioterapéuticas. Tras el inevitable debate, se acometieron algunos programas a partir de 1984⁴.

⁴Según los autores, la discusión en torno al tratamiento en las prisiones, su conveniencia, utilidad, etc., ha seguido en Alemania un desarrollo semejante al resto de los países occidentales. Por una parte, resultó desmedida la perspectiva según la cual el tratamiento vendría a ser una solución al problema carcelario; por otra, desde el mundo penal se

Estos programas se concretaron en los siguientes términos: un preso -mujer u hombre- podía ser trasladado a una prisión socioterapéutica, con su libre consentimiento, si los recursos terapéuticos existentes en ella parecían indicados para su resocialización; paralelamente, podía ser devuelto a una prisión normal si se evidenciaba que tales recursos no resultaban exitosos. Las plazas de las prisiones o departamentos socioterapéuticos llegaron a un número total de 650 -el 1.5% del total de plazas penitenciarias-, todas ellas destinadas a adultos. Existían diez departamentos para hombres y uno para mujeres. Generalmente tenían autonomía física y organizativa, con una capacidad que oscilaba entre 26 y 60 plazas, salvo la prisión de Berlín Oeste con 230 plazas. La "ratio" presos/personal era superior a las prisiones normales, siendo ésta de 2:1 para el total del personal, y entre el 6:1 y 16:1 por lo que se refiere al personal de rehabilitación. La duración del tratamiento oscilaba entre 1 y 3 años, y la selección de los presos no se limitó a los grupos de delincuentes señalados como objetivo en la reforma penal de 1969 -de marcado cariz clínico o psicopatológico- sino que también se incluyeron muchos jóvenes reincidentes con problemas de personalidad.

No ha existido una concepción unitaria del tratamiento. De acuerdo con un concepto amplio de terapia social, las medidas interventivas no fueron sólo aplicadas a individuos con síntomas específicos, sino que se incorporaron, dentro de los programas, cambios en las condiciones de vida, procesos grupales, y factores organizacionales e institucionales de carácter general. Ninguna de las prisiones se encuadraba en un sólo modelo o escuela

produjeron múltiples críticas a la aplicación del tratamiento en las prisiones: no permitiría una suficiente disuasión y castigo de la culpa, sería excesivamente costoso, disminuiría la seguridad, individualizaría y patologizaría los problemas delictivos, llevaría a involuntarias adaptaciones de la personalidad, implicaría decisiones de insuficiente transparencia jurídica, sería contrarrestado por los nocivos efectos de la prisionización, y, además, sería incapaz de preparar para la situación de libertad por la propia naturaleza del encarcelamiento. No son distintos todos los anteriores argumentos, contrarios al tratamiento en las prisiones, a los aducidos por diversos penalistas en nuestro país (Mapelli, 1983; Muñoz Conde, 1987; Bergali, 1987), lo que no resulta extraño si tenemos en cuenta que la mayoría de nuestros penalistas beben en las fuentes del derecho penal alemán, reproduciendo en buena medida sus argumentos y debates. Lösel y Körfel (1989) señalan cómo se ha producido una permanente confusión entre las cuestiones de valores -fundamento, justificación, inconvenientes, etc., del "tratamiento"- y las cuestiones de hecho -generalizada inexistencia o mínima aplicación real de programas de tratamiento-.

terapéutica. Aunque en términos generales puede afirmarse que los conceptos psicoterapéuticos tradicionales cedieron en favor de formas de modelos de entrenamiento social y de habilidades de competencia para la vida diaria. La mayoría de las prisiones socioterapéuticas utilizaron técnicas como el fortalecimiento de los contactos con el mundo exterior, la promoción de puestos de trabajo en la comunidad, el uso intensivo de las salidas y permisos, y la preparación de la liberación.

Método. La base del presente meta-análisis la constituyeron 16 análisis primarios, algunos de los cuales tenían muestras parcialmente coincidentes. Se toleró esta dependencia entre algunos de los resultados para poder contar con una base de datos suficiente. Los trabajos estudiados son todos los publicados (total o parcialmente) en Alemania Federal hasta 1985 (de hecho entre 1977-85). Se combinaron distintas formas de meta-evaluación, entre ellas la revisión cualitativa de literatura, las sistematizaciones tabulares, y el meta-análisis formal o estadístico.

Se diferenciaron los estudios según hubieran empleado como criterio de éxito medidas de reincidencia o rasgos de personalidad. Para operacionalizar la reincidencia se habían empleado diversos criterios, como por ejemplo la asignación a los sujetos de nuevas sentencias de prisión y también la comisión de pequeños delitos, oscilando los períodos de seguimiento entre 3 y 5 años. Los constructos de personalidad se habían operativizado mediante cuestionarios, escalas de actitudes, o auto-valoraciones.

Para analizar el efecto de los programas de tratamiento se emplearon los siguientes procedimientos:

1. Sistematización según las categorías de contenido. Se desarrolló un esquema de análisis para ordenar los estudios primarios, que incluía características de la prisión, sujetos participantes, tratamiento aplicado, diseño, teoría evaluativa y de criminalidad, selección de las muestras, recogida y análisis de los datos, y presentación de los resultados.

2. Incremento de la información sobre el tratamiento y el contexto de la prisión. Ya que existía en los estudios primarios bastante información incompleta, se buscó nueva información y datos diversos sobre el desarrollo

de los programas, consultando para ello otras publicaciones, datos internos de la prisiones, etc.

3. Entrevistas de feedback. Con el mismo objetivo de obtener más información útil para el meta-análisis se aplicaron a los evaluadores originales entrevistas semi-estandarizadas, que permitieron profundizar en la concreción de algunos datos e incrementar, de este modo, la validez.

4. Recogida de datos no publicados. Las entrevistas facilitaron en ciertos casos la obtención de nuevos datos estadísticos que no constaban en los informes originales. Ello permitió una mayor unificación del cálculo del tamaño del efecto.

5. Codificación sistemática de las amenazas contra la validez. Cada estudio fue evaluado respecto de las amenazas contra la validez por dos jueces, según el procedimiento de S. Cook y K. Campbell. Este procedimiento incluye características de validez estadística (7 ítems), validez interna (13 ítems), validez de constructo (10 ítems), y validez externa (3 ítems). En este caso se añadió un ítem más para la validez externa, referido a la estabilidad temporal de los efectos de los programas, y una nueva categoría denominada validez descriptiva (5 ítems).

6. Cálculo unificado de los índices de tamaño del efecto. Para los criterios cuantitativos de efectividad se utilizó el coeficiente d de Cohen, aplicable en este caso a las variables de personalidad. Mientras que para las variables cualitativas/dicotómicas, como la reincidencia, fue necesario el uso del coeficiente r_m de Friedman. Este coeficiente puede ser interpretado como una medida del grado en que el grupo de tratamiento supera al control en los resultados deseados (en este caso, la no reincidencia). El coeficiente r_m fue de alrededor de 0.5 d .

Para el cómputo de los tamaños de los efectos se utilizaron cinco modelos distintos, lo que permite estimar la dependencia que tienen los resultados de los métodos utilizados para su cálculo (Véase Tabla 4.1).

Tabla 4.1. Modelos utilizados para el cálculo del TE

Modelo	Resultados incluidos en el mismo
1	Todas las puntuaciones \underline{r}_m de cada estudio con $P < 0.05$ Todas las restantes puntuaciones \underline{r}_m son tratadas como efectos nulos
2	Todas las puntuaciones \underline{r}_m de cada estudio con $P < 0.10$ Todas las restantes puntuaciones \underline{r}_m son tratadas como efectos nulos
3	Todas las puntuaciones \underline{r}_m de cada estudio sin considerar su significación
4	Todas las puntuaciones \underline{r}_m de cada estudio con $P < 0.10$ Todas las restantes puntuaciones \underline{r}_m son ignoradas
5	Todas las puntuaciones \underline{r}_m de cada estudio con $P < 0.05$ Todas las restantes puntuaciones \underline{r}_m son ignoradas

Fuente: Tomado de Lösel y Köferl (1989, p.342)

Estos cinco modelos pueden considerarse como una estimación graduada del éxito del tratamiento en las prisiones socioterapéuticas, siendo el Modelo 1 el más "conservador" o moderado en su estimación de los efectos, y el Modelo 5 el más "optimista".

Resultados: Análisis del Tamaño del Efecto. Se calcularon las medias aritméticas de los tamaños del efecto observados de acuerdo con los cinco modelos descritos para cada estudio valorativo. En todos los análisis de reincidencia todos los efectos fueron positivos -reducción de la reincidencia-, excepto en uno. En cambio, los estudios que utilizaron medidas de personalidad produjeron el problema de la manipulación de los resultados que eran contrarios a las expectativas.

Según las puntuaciones obtenidas en los diferentes modelos, los efectos fueron generalmente moderados, pero consistentes (rango entre $r_m = 0.309$ y 0.000). El tamaño del efecto medio de los cinco modelos osciló entre 0.075 y 0.136, y el de los estudios aislados a lo largo de todos los modelos entre 0.017 y 0.255. El intervalo de confianza del 95% estuvo entre 0.077 y 0.181 para la estimación del efecto más optimista, y entre 0.038 y 0.112 para la estimación más moderada. La puntuación estimada más corriente para el efecto de las prisiones socioterapéuticas fue una puntuación de $r_m = 0.110$. Por lo que se refiere a la diferenciación entre estudios que utilizaron como criterio la reincidencia y los que utilizaron variables de personalidad sólo se produjeron bajas diferencias en favor de los segundos.

Resultados: Relaciones entre Características del Tratamiento y Tamaños del Efecto. En el meta-análisis tiene, según ya hemos argumentado, particular importancia el estudio de las relaciones entre las características del tratamiento (tipo, sujetos, contexto institucional, etc.) y los tamaños del efecto, ya que de ello se derivan indicaciones diferenciales para posteriores desarrollos de tratamientos. En el estudio de Lösel y Köferl (1989) no pudieron establecerse estas relaciones de modo consistente, debido, por un lado, a que el número de estudios y prisiones fue demasiado pequeño, y, por otro, a que el diseño del tratamiento no fue con frecuencia demasiado homogéneo o no fue descrito con suficiente claridad para propiciar una diferenciación válida de las clases de tratamientos. Por lo que el análisis en este terreno sólo puede tomarse como un pequeño estudio piloto.

Pudieron ser diferenciadas las siguientes características del tratamiento: tamaño de la prisión (número de presos), razón presos/personal total, razón presos/personal terapéutico, duración media del tratamiento, edad media de los presos, proporción de tipologías delictivas en las muestras -contra la propiedad, delincuentes con psicopatologías (según criterios de la Ley de 1969)-, proporción de bajas, períodos de conflicto en el desarrollo de la prisión, y duración del seguimiento.

De todas las anteriores características sólo la variable tamaño de la prisión tuvo una correlación significativa con el TE, consistente a través de todos los modelos. No obstante, si atendemos a las correlaciones negativas se constata que la varianza atribuible al tamaño de la prisión vino básicamente

determinada por la mayor dimensión de la prisión de Berlín, que fue frecuentemente evaluada por varios de los estudios. Por ello, la fácil conclusión de que las prisiones pequeñas propician un mayor éxito del tratamiento no logró en este estudio suficiente base empírica.

Conclusiones. Este estudio difiere de los meta-análisis típicos en que se basa sólo en un pequeño número de trabajos y en unos cuantos tipos de programas (prisiones socioterapéuticas). Sin embargo la metodología integrativa se muestra útil para una mayor precisión evaluativa, especialmente en términos de validez y efectividad.

Al respecto de la validez, los resultados obtenidos por Lösel y Köferl (1989) les conducen a argumentar la necesidad de no limitarse a la perspectiva de la validez interna, que suele ser lo más frecuente. La elevada amenaza encontrada sobre la validez de constructo lleva a los autores a concluir que, con demasiada frecuencia, se deja de lado más de lo debido la fundamentación teórica y generalizabilidad de la terapia social y su evaluación.

El elevado nivel de amenaza hallado en términos de validez descriptiva sugiere que en el futuro debería prestarse más atención a una mejor y más comprehensiva documentación de los conceptos de tratamiento, contenido y resultados.

La principal conclusión de los autores, respecto de la efectividad de las prisiones socioterapéuticas, es que los efectos deseados (no reincidencia) pueden ser anticipados en un 10% más de presos en prisiones socioterapéuticas que en los que se hallan en prisiones normales. Lo que contraviene, en su opinión, la desencantada valoración de "nada da resultado".

Aunque generalmente en los meta-análisis suelen hallarse tamaños del efecto superiores en 2 a 4 veces a los obtenidos, los autores remarcan el hecho de que su evaluación se dirigió a programas desarrollados en la vida real de las prisiones, con todas las dificultades que ello conlleva, mientras que muchas de las intervenciones meta-analizadas en otros campos se refieren a intervenciones sobre estudiantes de psicología en el laboratorio, etc., de mucha más fácil y controlada aplicación y evaluación.

Finalmente no existieron diferencias claras entre diversos tipos de terapia social.

Para el futuro F.Lösel y P.Köferl recomiendan la necesidad de que los estudios primarios efectúen una descripción más precisa de los contenidos y objetivos del tratamiento, una mayor explicitación teórica de sus fundamentos, una mayor diferenciación de las evaluaciones globales, transformándolas en pequeños estudios de valoración más precisos, análisis empíricos, diferenciación de criterios de reincidencia, sistematización de múltiples criterios de éxito en que se integren variables de reincidencia y criterios de conducta y personalidad, una mayor orientación evaluativa, y una mayor documentación unificada de los datos de los estudios evaluativos para facilitar su integración.

Entre los méritos del meta-análisis de Lösel y Köferl está el de haber sido el único específicamente dirigido a programas llevados a cabo en las prisiones. Constituyen ventajas de orden metodológico su esmerado análisis de resultados, estructurado mediante cinco modelos de progresivo rigor analítico, así como su especial atención a eventuales sesgos de validez, a los que nos referiremos más adelante. Tal vez su inconveniente más notable es el pequeño número de estudios analizados, algunos de los cuales, además, incluían muestras dependientes entre sí, correspondientes a las mismas prisiones, en fechas parecidas, y, por tanto, integradas en parte con los mismos sujetos.

4.3. La revisión de M.W. Lipsey (Lipsey, 1992a, 1992b).

Dedicaremos las páginas siguientes a describir y comentar en detalle el último trabajo de revisión de la literatura correccional al que hemos tenido acceso, que es un estudio típicamente meta-analítico. Se trata de un trabajo de M.W. Lipsey (Lipsey 1992a, 1992b), en el que evaluó un amplio conjunto de 443 estudios, pertenecientes a cuatro décadas de investigación (1950-90).

Objetivos. Lipsey plantea como objetivos de su meta-análisis mejorar las revisiones previas, tanto narrativas como meta-analíticas, en los siguientes aspectos: (1) ampliación del espectro de la literatura analizada, llevando a

cabo una búsqueda exhaustiva de trabajos, sean publicados o no; (2) codificación detallada de cada trabajo, que permita un análisis probatorio de los efectos de tratamiento medidos, tanto en lo relativo a los métodos usados como a la naturaleza y circunstancias de los tratamientos empleados; (3) aplicación de las técnicas estadísticas meta-analíticas dirigidas a evaluar tanto la magnitud de los efectos de los programas como la variabilidad de dichos efectos. Pese a ello, por el momento, M.W. Lipsey sólo nos ofrece en el trabajo que analizamos resultados concernientes a la magnitud global de los efectos del tratamiento.

Método: Criterios de inclusión de los estudios. Para la selección de los trabajos que habrían de ser objeto de este meta-análisis se establecieron los siguientes criterios:

1. En primer lugar, debía haber existido alguna intervención o tratamiento, extensamente definida, cuyo objetivo (explícito o implícito) fuera la reducción, prevención, tratamiento, etc., de la delincuencia juvenil o de problemas de conducta antisocial semejante a la delincuencia. Se definió la delincuencia juvenil como aquella conducta susceptible de ser perseguida legalmente pese a que no se hubiera producido ni la detención ni el procesamiento, mientras que se definieron como conducta antisocial aquellas acciones que amenazan, destruyen o dañan la propiedad ajena, a otras personas, o a uno mismo. Se excluyeron todos aquellos programas que se decían dirigidos exclusivamente al tratamiento del consumo abusivo de sustancias tóxicas (drogas, etc.), sin que incluyeran entre sus objetivos componentes de conducta antisocial.

2. Los sujetos del tratamiento debían ser, en su mayor parte, jóvenes menores de 21 años y mayores de 12. Lipsey parte de la idea, ampliamente aceptada en la literatura correccional, de que el momento clave para llevar a cabo una eficaz intervención sobre el comportamiento delictivo es el período de la adolescencia, en que el individuo está todavía en proceso de formación. Así pues, este presupuesto le llevó a fijar como criterio inicial de selección de sus programas que los sujetos a quienes fueron aplicados tuvieran menos de veintiún años.

3. Deberían haberse medido en forma cuantitativa variables resultado,

incluyendo al menos una medida de la delincuencia. Debiendo, además, existir alguna comparación que contrastara una o más condiciones de tratamiento con una o más condiciones de control en las variables resultado establecidas.

4. Las comparaciones entre grupos de tratamiento y grupos de control deberían basarse, con carácter general, en la previa asignación aleatoria de los sujetos a las condiciones experimentales; en el supuesto de que se tratara de diseños con medidas pre y post de la variable resultado, se requería que hubiera alguna evidencia de equivalencia entre los grupos previa al tratamiento, o una serie de medidas de comprobación de esta similaridad que atendieran a parámetros como historia delictiva previa, sexo, y edad de los sujetos. Se excluyeron tanto los estudios con diseños pretest-postest, carentes de grupo de control, como las comparaciones post/test de grupos no asignados al azar y faltos de información al respecto de su equivalencia.

5. Se seleccionaron solamente trabajos llevados a cabo en Estados Unidos o países de habla inglesa (como Canadá, Gran Bretaña y Australia) y escritos en inglés. El autor fundamenta la inclusión de este criterio sobre la conveniencia de mantener cierta homogeneidad sociocultural sobre el significado de la delincuencia juvenil⁵.

6. Para limitar el análisis a programas correspondientes a la "era moderna", se incluyeron sólo estudios que hubieran sido publicados a partir de 1950 (después de la Segunda Guerra Mundial).

Método: Localización de estudios. Para identificar y recuperar los informes que podían tener interés para el análisis se emplearon las siguientes fuentes de información: (1) la primera aproximación se realizó a través de las referencias bibliográficas de las revisiones y meta-análisis previos; (2) el recurso más importante lo constituyó una búsqueda mediante las bases de

⁵Por nuestra parte no descartamos que la precedente argumentación teórica se complete con una razón práctica inconfesable, aunque no menos determinante: los angloparlantes -el autor lo es- generalmente desconocen cualquier otro idioma distinto del inglés. Difícilmente podrían ser, pues, incluidos trabajos llevados a cabo en Méjico, Venezuela, Argentina, u otros países de hispanoamérica, y escritos en español, o realizados en países europeos, y escritos, por ello, en francés, alemán, español, sueco, etc.

datos de "Dialog system", utilizando como palabras clave una serie de expresiones alternativas de los conceptos "investigación" ("research"), "delincuencia juvenil" ("delinquency"), y "tratamiento" ("treatment"), que permitieron generar conjunciones de estos conceptos que podían aparecer bien en el título, el resumen o en el índice de términos de los trabajos buscados; finalmente, (3) la tercera fuente de información provino de las referencias bibliográficas existentes en los informes identificados. Todo ello permitió constituir un catálogo inicial de más de 8000 citas bibliográficas de aparente interés meta-analítico, entre las que se incluían referencias a trabajos aparecidos en libros, informes técnicos, conferencias, tesis, memorias, y artículos publicados en revistas científicas. Aquellos informes y trabajos que no pudieron ser localizados mediante los procedimientos anteriormente descritos fueron solicitados directamente a los autores a través del envío de una carta.

Método: Codificación. La codificación de los informes fue realizada por estudiantes de doctorado en psicología, que habían sido entrenados y supervisados para ello, mediante el manual de codificación preparado al efecto.

Se confeccionó un plan de codificación que comprendía 154 ítemes o informaciones, en los siguientes apartados:

Tamaño del efecto. Para cada trabajo se identificaron todas las variables resultado, cuantitativas, que habían sido evaluadas en el grupos de tratamiento y control, dividiéndolas entre las referidas al comportamiento delictivo y las relacionadas con otras conductas o características. Se consignó la dirección del efecto para cada variable, constatando si ésta era favorable al grupo de tratamiento, al grupo control, o a ninguno de ellos.

Se calculó un índice de tamaño de efecto para cada medida, empleando para ello el índice d de Cohen. Si para el caso de algunas variables resultado no se informaba sobre las puntuaciones medias y las desviaciones típicas, el tamaño del efecto se estimaba a partir de cualquier otra información recogida, tal como los valores p , t , o F , tablas de contingencia, y otras.

Un modo frecuente de presentar información sobre variables delictivas fue

la tasa de reincidencia, o la proporción de sujetos que en cada grupo experimental habían sido nuevamente detenidos, etc., después del tratamiento. Mediante la transformación de arcoseno de Cohen estos porcentajes fueron convertidos en estimaciones del tamaño del efecto.

Muchos de los trabajos presentaban también información relativa a subgrupos de sujetos, como delincuentes varones o mujeres, etc. En estos casos, si era posible se consignó asimismo información sobre estos subgrupos, aunque no es ofrecida, por el momento, en el informe objeto de esta revisión.

Variables metodológicas. Para cada trabajo se codificó abundante información sobre el diseño de evaluación, las medidas empleadas, las muestras, etc., y, especialmente, sobre la equivalencia previa de los grupos experimental y control. Ello habría de servir de base para poder indagar la relación existente entre los efectos encontrados y la metodología usada en cada caso.

Contexto del programa. En cada estudio se recogieron también datos sobre el año y forma de su publicación, país, y características relativas a los investigadores, tales como la disciplina e institución a la que pertenecían.

Naturaleza del tratamiento. Como elemento crítico, se recogió y codificó información sobre el tipo de tratamiento, contexto de aplicación, patrocinador del mismo, duración, intensidad, y otras características de la intervención llevada a efecto.

Naturaleza de los sujetos. También se incluyeron informaciones sobre las características demográficas de los jóvenes delincuentes (por ejemplo, raza, sexo, edad), su historia delictiva previa y otras semejantes.

Resultados: Distribución del tamaño del efecto para variables delictivas. Se llevó a cabo un análisis previo para detectar los tipos de variables delictivas más frecuentemente utilizadas en el conjunto de los estudios, procediéndose a continuación a seleccionar una sola medida por estudio (pese a que muchos programas empleaban diversas medidas) en función de su similitud con la categoría más frecuentemente usada en los otros trabajos. La medida más utilizada se refería a alguna variación del concepto de reincidencia. Por ello, se favoreció la elección de este tipo de mediciones, que, además, habían de

ser medidas primarias, esto es expresamente recogidas por los investigadores a los efectos del estudio.

Se utilizó el coeficiente de Hedges $[1-(3/(4n_t+4n_c-9))]$ para ajustar los tamaños de los efectos, ya que aquéllos que se basan en muestras pequeñas tienden a sobreestimar los valores de las poblaciones a que se refieren, por lo que deben ser reducidos proporcionalmente.

Dirección y tamaño de los efectos. Según una primera aproximación, 285 programas, esto es el 64.3% de los evaluados, habían logrado efectos favorables al grupo de tratamiento, mientras que 131, o sea el 29.6% de los estudios, había obtenido efectos favorables al grupo de control. Finalmente, 27 trabajos (el 6.1%) no había producido diferencias entre los grupos. Ello permite rechazar la hipótesis nula -la no influencia del tratamiento- sobre la base de una $z = 7.32$, $p < .001$. Esto es, si el tratamiento no hubiera sido efectivo en reducir la delincuencia se habría esperado que las diferencias positivas y negativas se debieran a error de muestreo, habiéndose producido en proporciones semejantes.

Aunque no todos los estudios ofrecieron suficiente información para calcular el tamaño del efecto, 397 estudios disponían de esta información. La dirección de los efectos en este subgrupo de estudios fue idéntica a la del conjunto general de trabajos, mostrándose que este subconjunto de estudios encontraron una delincuencia inferior para los grupos de tratamiento que para los de control. Se obtuvo un TE medio no ponderado de .172, que es el valor más directamente comparable a los resultados de otros meta-análisis previos.

De acuerdo con este resultado, si -en palabras de Lipsey (1992a)- nos hacemos la pregunta general "¿Reduce el tratamiento la delincuencia?", la respuesta es "Sí, como promedio se observa un efecto positivo". Sin embargo, aunque este efecto es positivo y estadísticamente significativo, la media de los tamaños de los efectos encontrada aquí resulta relativamente modesta. Tomando como base la distribución de los efectos de las medias ponderadas por la varianza inversa, los jóvenes tratados mostraron en torno a .10 unidades de desviación típica menos delincuencia después del tratamiento que los jóvenes controles. En una primera impresión -comenta

Lipsey- esto suena bastante trivial.

La anterior medición tendría más significado si fuera transformada en algo más entendible que las unidades de desviación típica, por ejemplo expresándola en términos de tasa de reincidencia. Si asumimos que los grupos de control sin tratamiento reincidieron en una tasa media del 50%, el efecto logrado sería equivalente a un decremento promedio en los grupos de tratamiento de 10 puntos de porcentaje. Es decir, habrían reducido su reincidencia, como consecuencia del tratamiento, del 50% esperable al 40%.

Resultados: Análisis de la Variabilidad del Tamaño del Efecto. Con la pretensión de analizar la variabilidad del tamaño del efecto, las características de los estudios fueron agrupadas en 11 apartados. Estos grupos representan tres amplias categorías -contexto del estudio, método y tratamiento-. Fueron secuenciados desde las cuestiones metodológicas más fundamentales y generales hasta las más específicas del tratamiento y contexto del estudio (véase Tabla 4.2)

Tabla 4.2. Características de los estudios relacionadas con los TEs

Método

Grupos experimentales, tamaño de la muestra, muestreo

Equivalencia inicial de los grupos experimentales

Mortalidad de sujetos tras su asignación a los grupos

Naturaleza de la condición de control

Características de las medidas resultado de la delincuencia

Información sobre el cálculo del tamaño del efecto

Tratamiento

Características de los sujetos de tratamiento

Magnitud o intensidad del tratamiento

Modalidad de tratamiento y administradores del mismo

Filosofía del tratamiento y marco o escenario

Contexto del Estudio

País, año de publicación, disciplina del autor, etc.

Fuente: Adaptado de Lipsey (1992a).

En primer lugar se examinó la relación de cada agrupamiento individual de variables y el tamaño del efecto. Para ello se construyó una regresión múltiple ponderada por cada agrupamiento de variables, para estudiar si alguno de tales agrupamientos mostraba correlaciones notables con el tamaño del efecto y potencialmente explicaba parte de su varianza. O sea, se pretendía conocer si la variabilidad en los tamaños del efecto primarios reflejaba diferencias en los tratamientos y circunstancias de tratamiento que se analizaban, o era más bien el resultado de la variabilidad de los factores metodológicos considerados.

El análisis presentado por Lipsey (1992a) en este trabajo se dirige en menor grado a los detalles de estas circunstancias, que a cuestiones de conjunto. Sus primeros resultados indican que tanto los factores metodológicos como los de tratamiento influyen sobre los resultados. Consideradas en bloque, las variables metodológicas parecen jugar un mayor papel. Sin embargo, consideradas individualmente, la categoría más influyente fue la propia naturaleza del tratamiento.

El autor aplaza para posteriores informes el análisis de las específicas influencias e interacciones que tendrían, en su influjo sobre los resultados obtenidos, tanto las variables metodológicas como las de tratamiento.

Conclusiones. Resultan enormemente sugerentes algunas de las conclusiones de Lipsey (1992a). En su propia opinión, uno de los méritos de su trabajo es que aporta alguna luz sobre la historia de las investigaciones de revisión del tratamiento de la delincuencia. Afirma que el principal tamaño del efecto medio se acerca peligrosamente a cero, aunque no tanto como para justiciar el retórico "nada da resultado" de la década de los setenta. En gran parte de la literatura correccional revisada previamente, en especial si se ha realizado un mero recuento de votos, efectos positivos convincentes podrían estar siendo difícilmente discernibles cuando, como es el caso, se trata de muestras con tamaño modesto: no producen suficiente potencia estadística y resulta muy difícil lograr significación estadística para tamaños del efecto en el rango .10-.20 unidades de desviación típica.

Por último, en opinión de M.W. Lipsey, la amplia variabilidad en los efectos encontrada en este tipo de literatura significa que diferentes revisiones que

analicen distintas porciones de la misma pueden conducir, con honestidad, a muy diferentes conclusiones. En el extremo alto de la distribución hay estudios de revisión que muestran impresionantes efectos, como es el caso, por ejemplo, de los de Gendreau y Ross (1979), Palmer (1975), y Andrews et al. (1990). Contrariamente, en el extremo bajo de la distribución, e incluso en el medio, se han realizado numerosas revisiones que evidencian insignificantes e incluso aparentemente negativos efectos, como por ejemplo los de Martinson (1974), y Whitehead y Lab (1989). Y concluye Lipsey (1992a) con un circense e ilustrativo símil: "Si la heterogeneidad de la distribución de efectos en la investigación sobre el tratamiento de la delincuencia es tan grande como un elefante, quizá no es un milagro que cada revisor de la literatura, tomando una cola de la distribución aquí y una joroba allí, describa a la fiera de modo tan distinto" (p.42).

RESUMEN

En el presente capítulo hemos descrito con detenimiento tres de los últimos meta-análisis de programas de tratamiento realizados. El primero de ellos, el de Whitehead y Lab (1989) analizó 50 estudios con delincuentes juveniles, tomando en consideración la variable reincidencia en el delito. Estos autores establecieron un estricto criterio de éxito que cifraron en puntuaciones Phi superiores a .20, que equivaldría de un TE de alrededor de .41. Sin embargo, obtuvieron un TE promedio de .12, lo que les condujo a una conclusión negativa sobre la efectividad de los tratamientos analizados. El segundo trabajo, correspondiente a Lösel y Köferl (1989) -el único europeo-, se realizó con 16 estudios que analizaban la efectividad de la socioterapia aplicada en prisiones alemanas. Tomaron en consideración tanto variables psicológicas como reincidencia. Obtuvieron un TE promedio de entre .075 y .136, lo que les permitió afirmar un resultado moderadamente positivo. Por último, Lipsey (1992a, 1992b) ha descrito el más ambicioso y minucioso meta-análisis de literatura correccional realizado hasta ahora. Ha trabajado sobre una muestra superior a los 400 programas de tratamiento de delincuentes juveniles. Con ellos obtuvo un TE promedio no ponderado de .172. Su conclusión más general es también favorable: aunque de forma moderada, el tratamiento funciona. Más allá de sus propias conclusiones, el meta-análisis de Lipsey tiene para nosotros un valor paradigmático debido a su minucioso procedimiento y metodología de análisis.

5. META-EVALUACIÓN DE LOS META-ANÁLISIS PRECEDENTES

En el capítulo precedente hemos resumido tres de los meta-análisis previos, como paradigmas que guiarán en parte nuestro propio trabajo. Ahora efectuaremos un análisis sistemático de los meta-análisis de programas de tratamiento de la delincuencia que nos han precedido. Fijaremos nuestra atención, de manera especial, en cuatro grandes apartados de interés común a todos ellos y a nuestra investigación: cuáles han sido los criterios de selección de estudios que han utilizado, qué tipologías de los tratamientos han formulado, cuáles han sido sus taxonomías de las medidas resultado, y finalmente, cuáles han sido sus principales resultados.

Al igual que sucede con los propios estudios primarios de tratamiento de la delincuencia, los meta-análisis han ido aumentando, a lo largo de su pequeña historia, el número de estudios evaluados y, sobre todo, perfeccionando su metodología y procedimientos de análisis. Por ello, priorizando lo más actual, comenzaremos presentando en cada caso los meta-análisis más recientes e iremos retrocediendo hacia los más antiguos.

5.1 Criterios de selección de estudios

Recogemos en este epígrafe las pautas y criterios que los autores establecieron para aceptar o rechazar la inclusión de un determinado trabajo dentro de su investigación.

Lipsey (1992a, 1992b)

Lipsey (1992a), que incluyó 443 trabajos en su meta-análisis, empleó, según ya hemos comentado, los siguientes criterios de aceptación:

1. Presencia de alguna intervención, ampliamente definida, que tuviera por objetivo la reducción, prevención, o tratamiento de la conducta delictiva o antisocial.
2. Que se tratara de sujetos jóvenes, menores de 21 años.
3. Que el estudio informara sobre resultados cuantitativos, con, al menos, una medida de delincuencia, realizada en uno o más grupos de tratamiento, y en uno o más grupos de control.
4. Control experimental mediante asignación aleatoria de los sujetos a los grupos, o, en caso contrario, existencia de medidas suficientes de control de equivalencia previa entre los grupos (edad, sexo, antecedentes, etc.).
5. Sólo incluyó estudios en lengua inglesa, con el argumento de mantener la homogeneidad cultural.
6. Para limitar los programas a la "era moderna", sólo se incluyeron estudios posteriores a 1950.

Izzo y Ross (1990)

Por su parte Izzo y Ross (1990), que trabajaron con 46 estudios de los que derivaron 68 medidas de tamaño del efecto, establecieron los siguientes criterios:

1. Que las investigaciones describieran la evaluación de un tratamiento aplicado a delincuentes juveniles.
2. Que incluyeran datos empíricos provenientes de un diseño experimental o cuasi-experimental.
3. Que hubieran tomado alguna medida de reincidencia.
4. Se recogieron estudios publicados en revistas científicas entre 1970 y 1985.

Andrews et al. (1990)

El meta-análisis de Andrews et al. (1990) evaluó 80 trabajos, integrados por dos submuestras: la primera, 45 estudios de los 50 previamente analizados por Whitehead y Lab (1989), relativos a delincuencia juvenil; la segunda, 35 más seleccionados por ellos sobre delincuentes adultos. Sus parámetros de inclusión fueron:

1. Tratamientos de delincuentes juveniles y adultos.
2. Publicados en revistas profesionales entre 1950 y 1989.
3. Que incluyeran alguna medida binaria (menos/más) de la reincidencia.

Lösel y Köferl (1989)

Lösel y Köferl (1989) analizaron 16 estudios primarios realizados en prisiones "socioterapéuticas" alemanas; algunos de estos trabajos incluían muestras de sujetos parcialmente solapadas, y por ello dependientes.

Whitehead y Lab (1989)

La investigación meta-analítica de Whitehead y Lab (1989) operó con 50 trabajos, seleccionados con arreglo a los criterios siguientes:

1. Evaluaciones de tratamientos correccionales juveniles.
2. Estudios publicados en revistas profesionales entre 1975 y 1984.
3. Que tuvieran grupo de control.
4. Que incluyeran datos cuantitativos susceptibles de ser meta-analizados.
5. Que hubieran efectuado alguna media de reincidencia.

Gottschalk et al., (1987)

Gottschalk, Davidson II, Gensheimer y Mayer (1987) llevaron a cabo un meta-análisis de 90 programas, que seleccionaron con arreglo a los siguientes

parámetros:

1. Incluyeron trabajos realizados entre 1967 y 1983.
2. Tratamientos dirigidos a delincuentes juveniles.
3. Que incluyeran una medida de la delincuencia posterior, como mínimo.
4. Sólo seleccionaron programas comunitarios, en oposición a tratamientos institucionales.

Con posterioridad, estos mismos autores (Gottschalk, Davidson II, Mayer y Gensheimer, 1987) realizaron una meta-análisis específico de los 25 programas de orientación conductual y otro de los 35 programas de diversion incluidos entre los 90 estudios de su primer trabajo. Prescindiremos aquí de estas investigaciones, por considerarlas subsumidas dentro de la primera mencionada.

Kaufman (1985)

Sobre el meta-análisis de Kaufman (1985) sólo hemos podido acceder a alguna información indirecta, muy incompleta, a través de las revisiones precedentes, ya que se trata de una tesis doctoral no publicada.

Garrett (1985)

El primer meta-análisis importante, el de Garrett (1985), que trabajó con 111 estudios, había seleccionados sus materiales en consonancia con que:

1. Los estudios hubieran sido realizados entre 1960 y 1983.
2. Los programas de tratamiento se hubieran llevado a cabo en una institución o comunidad residencial (no se incluyeron, por tanto, programas de "diversion", "probation", etc.)
3. Se tratara de delincuentes juveniles, menores de 21 años.
4. Hubieran tenido algún control experimental en el diseño.

Seguidamente, en la Tabla 5.1, se ha formulado un resumen de los criterios de selección de estudios empleados por los distintos meta-análisis.

Tabla 5.1. Criterios de selección de estudios. En la cabecera de las columnas se consignan los distintos meta-análisis llevados a cabo hasta la fecha. En el encabezamiento de las filas aparecen el número de estudios incluidos en cada meta-análisis, y los criterios utilizados para la selección de los estudios evaluados. Para cada meta-análisis se indica -mediante un círculo- cuáles fueron los criterios utilizados.

CRITERIOS DE SELECCIÓN DE ESTUDIOS	Lipsey (1992)	Izzo y Ross (1990)	Andrews et al. (1990)	Lösel y Köferl (1989)	Whitehead and Lab (1989)	Gottschalk et al (1987)	Kaufman (1985)	Garrett (1985)
Número de estudios evaluados	443	46	80	16	50	90	20	111
Tratamiento de pre-delincuentes							0	
Tratamiento de delincuentes juveniles (<21 años)	0	0	0		0	0		0
Tratamiento de delincuentes adultos			0	0				
Al menos una medida cuantitativa de la delincuencia (reincidencia)	0	0	0		0	0		
Medida de delincuencia o de variables psicológicas relacionadas				0				
Diseño experimental: grupo de control equivalente	0	0	0		0		0	0

CRITERIOS DE SELECCIÓN DE ESTUDIOS	Lipsey (1992)	Izzo y Ross (1990)	Andrews et al. (1990)	Lösel y Köferl (1989)	Whitehead and Lab (1989)	Gottschalk et al (1987)	Kaufman (1985)	Garrett (1985)
Diseño cuasi-experimental: grupo de control no equivalente	0	0						0
Alguna forma de control experimental				0		0		0
Sólo programas residenciales (no "diversion", "proba tion", etc.)				0				0
Sólo programas comunitarios (no institucionales)						0		
Estudios posteriores a 1950	0							
Estudios publicados y no publicados	0						0	
Estudios realizados entre 1960-83								0
Estudios realizados entre 1967-83						0		

CRITERIOS DE SELECCIÓN DE ESTUDIOS	Lipsey (1992)	Izzo y Ross (1990)	Andrews et al. (1990)	Lösel y Köferl (1989)	Whitehead and Lab (1989)	Gottschalk et al (1987)	Kaufman (1985)	Garrett (1985)
Sólo estudios publicados en revistas entre 1970-85		O						
Sólo estudios publicados en revistas entre 1950-89			O ⁶					
Sólo estudios publicados en revistas entre 1975-84					O			
Estudios publicados en Alemania entre 1977-85				O				
Sólo estudios publicados en inglés	O	(*)	(*)	(**)	(*)	(*)	(*)	(*)

(*) Aunque todos estos autores no refieren ningún criterio lingüístico para la selección de los estudios, cabe suponer que en todos los casos "de facto" el criterio lingüístico es el haber sido divulgados en inglés, o, como máximo, tal vez ampliado el criterio al francés en alguno de los trabajos canadienses (Izzo & Ross, 1990; y Andrews et al., 1990). (***) Probablemente, todos los estudios analizados por Lösel y Köferl (1989) habían sido publicados en alemán. Me dirigi por carta al primer autor pidiéndole copia de los estudios, sin recibir respuesta.

⁶Sin embargo, la selección no pretende ser representativa, según los propios autores (Andrews et al., 1990), de la población de estudios publicados durante el período cubierto.

En lo relativo a la selección de los estudios, pueden apreciarse diferencias entre los diversos meta-análisis en algunos aspectos fundamentales. El primero se refiere a quiénes son los sujetos que son destinatarios del tratamiento. La mayoría de los análisis (Lipsey, 1992a; Izzo y Ross, 1990; Whitehead y Lab, 1989; Gottschalk et al., 1987; y Garrett, 1985) trabajaron exclusivamente con tratamientos con delincuentes juveniles. Un trabajo, el de Kaufman (1985), lo hizo sólo con pre-delincuentes, y otro, el de Lösel y Köferl (1989), se dirigió a sujetos adultos de modo exclusivo. Mientras que sólo un análisis, el de Andrews et al. (1990) incluyó tanto tratamientos de delincuentes juveniles como de adultos.

El segundo apartado de discrepancia tiene que ver con el tipo de variables resultado que son requeridas. La mayoría de los trabajos exigen al menos una medida directa de la delincuencia/reincidencia (Lipsey, 1992a; Izzo y Ross, 1990; Andrews et al., 1990; y Whitehead y Lab, 1989), siendo únicamente el de Lösel y Köferl (1989) el que permite la sola existencia de medidas psicológicas indirectas.

El tercer punto de diferenciación se refiere al grado de control experimental exigido a los estudios para su inclusión en el análisis. La mayor parte de los trabajos americanos requirieron un control experimental duro que supusiera la existencia de grupo de control equivalente -diseño experimental- o grupo de comparación con control de equivalencia -diseño cuasi-experimental- (Lipsey, 1992a; Izzo y Ross, 1990; Andrews et al., 1990; y Whitehead y Lab, 1989; Kaufman, 1985). Sólo un trabajo americano, el de Garrett (1985), y el europeo de Lösel y Köferl (1989), permitieron a sus estudios un control experimental más blando, con la inclusión de diseños pre-experimentales (los diseños pre-post, por ejemplo).

En la dimensión contexto de los programas, la mayoría de análisis acogieron todo tipo de contextos interventivos (prisiones, comunidad, etc.), mientras que los trabajos de Lösel y Köferl (1989) y de Garrett (1985) sólo incluyeron programas residenciales.

Por lo que se refiere al período temporal de estudios muestreados cada trabajo incluyó, como puede verse, un período diferente. Algunos cubren períodos muy largos de entre veintitrés y treinta y nueve años (Garrett, 1985; y Lipsey, 1992a), siendo los períodos cubiertos por el resto de trabajos más reducidos, de en torno a diez años.

Finalmente, todos acogen de modo exclusivo literatura correspondiente a sus propios países o a su propia lengua, aunque ésta sea de tan extenso uso científico como el inglés.

5.2. La codificación de los tipos de tratamientos

Con objeto de establecer en nuestra investigación una tipología adecuada de los tratamientos utilizados, prestamos especial atención a este punto de los meta-análisis precedentes. Al igual que en otros aspectos, tenía un especial interés para nosotros el meta-análisis de M.Lipsey, debido a ser el más reciente y por su exhaustividad. Sin embargo, desgraciadamente, en los informes divulgados hasta la fecha por Lipsey (1992a, 1992b) no ha sido incluido el análisis diferencial de tipos o conjuntos de tratamientos, por lo que hemos tenido que prescindir de su consideración por el momento.

Izzo y Ross (1990)

Izzo y Ross (1990) clasifican los estudios meta-analizados en función de la concepción teórica de la conducta social que en ellos se formula: aprendizaje social, modificación de conducta, modelado, teoría sistémica, terapia de realidad, teoría del nivel de madurez interpersonal, teoría sociológica general, y otras conceptualizaciones teóricas. Este criterio no es, como puede verse, estrictamente clasificatorio de los tipos de tratamiento, sino de los modelos teóricos que subyacen a los mismos. Por esta razón, la clasificación de Izzo y Ross (1990) adquiere una dimensionalidad más vinculada a las teorías subyacentes a la intervención que a las propias técnicas interventivas utilizadas. Por ello, la taxonomía de Izzo y Ross (1990) nos será útil a la hora de formular una tipología de las conceptualizaciones teóricas de las intervenciones, a cuya codificación

también nosotros hemos prestado atención.

Andrews et al. (1990)

Andrews, Zinger, Hoge, Bonta, Gendreau y Cullen (1990) han sido los autores más atrevidos a la hora de clasificar los tipos de tratamiento utilizados por diversos programas. Razonan que no puede acometerse una revisión de programas sin antes ponderar adecuadamente la calidad de las intervenciones llevadas a cabo. Para ello, establecen tres criterios, que obtienen de una amplia revisión de la literatura correccional previa. Estos criterios son los siguientes:

1. El principio de riesgo y selección del nivel de servicio. En él establecen que, en general, se destinan mejores servicios correccionales a delincuentes de alto riesgo que a los de bajo riesgo, que sólo recibirían servicios minimalistas. Por ello, su primer criterio clasificatorio de los programas obedece a esta mayor calidad implícita de los programas con delincuentes de alto riesgo.

2. El principio de necesidad y selección de objetivos intermedios apropiados. Las necesidades criminógenas de los sujetos, o factores de riesgo dinámicos, son del máximo interés para la intervención, ya que serían aquéllos estrechamente asociados a la conducta criminal. Cuando tales factores de riesgo cambian, la conducta delictiva también cambia, reduciéndose.

En esta dirección, los objetivos intermedios más prometedores serían, según Andrews et al. (1990), el logro de cambios en las actitudes y sentimientos antisociales de los individuos, en su asociación con otros delincuentes, la promoción del afecto familiar en combinación con la mejora de la supervisión paterna, la identificación con modelos no delictivos, la enseñanza de habilidades de autocontrol y automanejo, la sustitución de habilidades de robo y agresión por habilidades prosociales, la reducción de las dependencias de sustancias tóxicas; y, en general, todos aquellos objetivos encaminados a invertir la relación entre recompensas y costes generados por las conductas delictivas y no delictivas en los diversos contextos conductuales de los sujetos (familiares, escolares, laborales y otros).

Por el contrario, en opinión de Andrews et al. (1990), entre los objetivos intermedios poco prometedores se encontrarían: el aumento de la autoestima de los delincuentes y de la cohesión de grupos delictivos sin alterar su motivación antisocial, la mejora de las condiciones de vida de ciertos vecindarios pero sin llegar a las familias de alto riesgo, y la focalización en etéreos problemas personales o emocionales no conectados con la reincidencia.

3. El principio de respuesta adecuada ("responsivity") y de selección del tipo de servicio. El principio de respuesta adecuada haría referencia, como criterio positivo, a la selección de aquellos servicios (1) susceptibles de influir concretos objetivos intermedios, que se dirigen a necesidades de los delincuentes, y (2) que se relacionan íntimamente con estilos de aprendizaje de los mismos.

En general, pueden considerarse tipos de servicios apropiados -concluyen Andrews et al, (1990)- aquéllos que implican la utilización de principios de conducta y aprendizaje social para influir la conducta interpersonal y mejorar las habilidades y desarrollo cognitivo de los sujetos. Se encontrarían aquí el modelado, la práctica graduada (o moldeamiento), el ensayo, el "role playing" o juego de roles, el refuerzo, la mejora de recursos [personales], y la mejora de las habilidades verbales y de razonamiento (expresión de sugerencias, expresión de razones, reestructuración cognitiva).

Por el contrario, se considerarían servicios inefectivos los programas de grupo no estructurados, los grupos de consejo y grupos permisivos de iguales, y los enfoques de terapia de "milieu" (o de acomodación contextual a las necesidades terapéuticas de los sujetos). Según Andrews et al. (1990), se ha documentado la inefectividad de todos estos enfoques tanto en ámbitos comunitarios, como hospitalarios y en prisiones. En este punto los autores (Andrews et al., 1990) advierten, además, contra la utilización indiscriminada con delincuentes de grupos no estructurados, que pueden convertirse, en la práctica, en contextos de favorecimiento criminógeno.

Dentro de los enfoques manifiestamente inefectivos Andrews et al. (1990) incluyen las terapias psicodinámicas y no directivas, centradas en el cliente.

Todas ellas estarían dirigidas, en esencia, a poblaciones libres con problemas neuróticos o de sobrecontrol, que no son precisamente los problemas más habituales e importantes de la mayoría de los delincuentes.

En base a estos criterios, Andrews et al. (1990) realizan una clasificación de todos los estudios incluidos en su meta-análisis, catalogándolos como:

a) Sanciones penales (sin componentes concretos de rehabilitación): entre ellas se encontrarían la restitución, la amonestación policial, las medidas de menos versus más probation, y probation versus custodia.

b) Servicios correccionales inapropiados: entre éstos sitúan (1) los servicios para casos de bajo riesgo y/o los sistemas de respuesta penal, (2) los servicios de relación no directiva y/o de consejo psicodinámico no estructurado, (3) todos los acercamientos de grupo o "milieu" con el énfasis puesto en la comunicación de grupo y sin un plan claro para mejorar el control sobre el modelado y refuerzo prodelictivo, y (4) los acercamientos académicos y vocacionales no directivos o pobremente objetivados.

c) Servicios correccionales apropiados: contemplan como tales (1) los servicios para casos de alto riesgo, (2) todos los programas conductuales (excepto los dirigidos a casos de bajo riesgo), (3) las comparaciones de específicos tratamientos de respuesta adecuada, y (4) los programas no conductuales con un claro abordaje de necesidades criminógenas mediante una intervención estructurada.

Lösel y Köferl (1989)

Por su parte, Lösel y Köferl (1989), en su meta-análisis de 16 prisiones "socioterapéuticas" alemanas, utilizaron como único criterio conceptualizador del tratamiento la "socioterapia", término bastante genérico para recoger una variedad de acciones interventivas diversas como educación, formación laboral o atención individualizada. Esta conceptualización responde más a una concepción teórica (intervenir frente a no intervenir) que a una tipología de técnicas, que en este caso no se explicitan.

Whitehead y Lab (1989)

Mayor interés tiene para nosotros, en cambio, la tipología de Whitehead y Lab (1989) que, en su meta-análisis de 50 programas juveniles, clasifican los tratamientos en cinco categorías: (1) Programas de "diversion" extrajudiciales, a los que los sujetos serían asignados sin participación de las instancias legales; (2) Programas de "diversion" dentro del sistema de justicia, cuyo desarrollo sería una alternativa legal a las penas tradicionales; (3) Programas de orientación comunitaria (como la probation y la libertad bajo palabra); (4) Programas residenciales; y (5) Intervenciones especializadas. También esta clasificación tiene un cariz más global, de segundo orden, o teórico, que de especificación de las técnicas concretas empleadas en cada caso.

Kaufman (1985)

Kaufman (tomado de Lipsey, 1992a) analizó 20 estudios sobre pre-delinquentes. Clasificó los tratamientos en cinco categorías: (1) Consejo, (2) Servicios, (3) Contrato de contingencias, (4) Asistencia social o supervisión ("casework"), y (5) Formación laboral.

Gottschalk et al. (1987)

En el meta-análisis efectuado por Gottschalk, Davidson II, Gensheimer y Mayer (1987), sólo se incluyeron programas comunitarios con delinquentes juveniles. Su tipología de tratamientos viene constituida por un elenco amplio de técnicas, conductuales la mayoría aunque no todas. Así Gottschalk et al. (1987), al analizar los tamaños de los efectos de los programas en función del tipo de tratamiento empleado, obtuvieron resultados positivos de la intervención académica, el reforzamiento positivo, así como las intervenciones en el ámbito profesional/vocacional; no produjeron efectos significativos técnicas como el modelado, la economía de fichas, el contrato conductual, la "probation", la asistencia social o supervisión ("intensive casework"), el apoyo, las técnicas psicodinámicas, la terapia centrada en el

cliente, la terapia cognitiva, y la terapia de realidad; mientras que la terapia de grupo y la transaccional tuvieron efectos negativos, de acuerdo con su análisis. Paralelamente, Gottschalk et al. (1987) agrupan las diversas técnicas empleadas en los programas en una serie de seis tipos de intervención de segundo orden: (1) Conductual, (2) Educativa/vocacional, (3) Psicoterapia individual, (4) Psicoterapia de grupo, (5) Asistencia social-supervisión/ probation, y (6) Inespecífica.

Garrett (1985)

Garrett (1985), al igual que hicieron Gottschalk et al. (1987), categorizó los tratamientos de los 111 estudios analizados en una doble taxonomía, una de segundo orden, más global, y, anidada en ella, una de primer orden, relativa a las técnicas específicas utilizadas en los programas. De acuerdo con ello, los tratamientos quedarían catalogados en: (1) Psicodinámicos (incluyendo las técnicas individual, grupal y consejo familiar), (2) Conductuales (con los subgrupos de manejo de contingencias, cognitivo-conductual, grupos de encuentro guiados, cultura de iguales positivos, y de contexto o "milieu"), (3) Habilidades de vida (drogas/alcohol, académico, profesional, y actividades al aire libre), y (4) Otros (musicoterapia, habilidades inespecíficas, etc.).

El repaso de las taxonomías interventivas de los meta-análisis precedentes nos obliga a diferenciar dos dimensiones distintas, superpuestas en algunos de ellos: una dimensionalidad de primer orden, estrechamente vinculada a las técnicas concretas predominantes en cada programa (p.e. psicoterapia individual, psicoterapia de grupo, asistencia social, "probation", modelado, contrato conductual, economía de fichas, etc.). Esta dimensionalidad se refleja en las tipologías de Gottschalk et al. (1987), Kaufman (1985), y Garrett (1985).

En segundo término, en algunos meta-análisis las técnicas concretas son anidadas en una dimensionalidad de segundo orden, más relacionada con modelos teóricos que subyacen a las diversas técnicas (p.e. terapias psicodinámicas, conductuales, sanciones penales, servicios correccionales apropiados/inapropiados, etc.). Esta segunda dimensionalidad teórica puede

ser hallada en las taxonomías interventivas de Izzo y Ross (1990), Andrews et al. (1990), Lösel y Köferl (1989), Whitehead y Lab (1989), Gottschalk et al. (1987) y Garrett (1985).

De acuerdo con esta doble diferenciación de dimensiones clasificatorias hemos confeccionado las Tablas 5.2 y 5.3. La Tabla 5.2 permite la visión de conjunto de las diferentes técnicas incluidas en las taxonomías de los meta-análisis precedentes. La Tabla 5.3 recoge las taxonomías de segundo orden, más vinculadas a los modelos teóricos utilizados.

5.3. La codificación de las medidas resultado

Dada su gran sencillez ahorraremos en este punto la narración de las medidas resultado utilizadas en los distintos meta-análisis, pasado directamente a su presentación esquemática en la Tabla 5.4, que incluimos a continuación.

Tabla 5.2. Tipos de tratamientos en diferentes meta-análisis: dimensionalidad tecnológica. En el encabezamiento de los columnas se referencian los meta-análisis de programas correccionales que han utilizados la dimensión "técnicas concretas". En el encabezamiento de las filas aparecen distintos tipos de técnicas treatmentales. Para cada meta-análisis se indica -mediante un círculo- qué técnicas fueron categorizadas en cada uno de ellos.

TIPOS DE TRATAMIENTO: DIMENSIONALIDAD TÉCNICA	Gottschalk et. al (1987)	Kaufman (1985)	Garrett (1985)
Número de estudios evaluados	90	20	111
Intervención académica	0		0
Terapia de grupo	0		0
Reforzamiento positivo	0		0
Intervención profesional/vocacional	0		0
Modelado/role playing	0		
Apoyo	0		
Contrato conductual	0		
Economía de fichas	0		
Asistencia social (casework)	0		
"Probation"	0		
Terapia centrada en el cliente	0		
Terapia psicodinámica	0		
Terapia cognitiva	0		0
Análisis transaccional	0		
Consejo [psicológico]		0	

TIPOS DE TRATAMIENTO: DIMENSIONALIDAD TÉCNICA	Gottschalk et. al (1987)	Kaufman (1985)	Garrett (1985)
Servicios		0	
Terapia individual			0
Terapia familiar			0
Grupos encuentro guiados/cultura iguales positivos			0
Terapia de contexto o "milieu"			0
Tratamiento de drogas/alcohol			0
Actividades al aire libre			0
Otras técnicas			0

Tabla 5.3. Tipos de tratamientos en diferentes meta-análisis: dimensionalidad teórica. En el encabezamiento de los columnas se referencian los meta-análisis de programas correccionales que han utilizados la dimensión "teórica". En el encabezamiento de las filas aparecen distintas agrupaciones teóricas. Para cada meta-análisis se indica -mediante un círculo- qué dimensiones/modelos teóricos fueron categorizados en cada uno de ellos.

TIPOS DE TRATAMIENTO: DIMENSIONALIDAD TEÓRICA	Izzo y Ross (1990)	Andrews et al. (1990)	Lösel y Körfel (1989)	Whitehead and Lab (1989)	Gottschalk et. al (1987)	Garrett (1985)
Número de estudios evaluados	46	80	16	50	90	111
Aprendizaje social	0					
Modificación de conducta/terapias conductuales	0				0	0
Modelado	0					
Teoría sistémica	0					
Terapia de realidad	0					
Teoría del nivel de madurez interpersonal	0					
Teoría sociológica general	0					
Sanciones penales		0				
Servicios correccionales inapropiados		0				
Servicios correccionales apropiados		0				
Socioterapia			0			
"Diversion" extra-judicial				0		
"Diversion" dentro del sistema de justicia				0		
Orientación comunitaria				0		

TIPOS DE TRATAMIENTO: DIMENSIONALIDAD TEÓRICA	Izzo y Ross (1990)	Andrews et al. (1990)	Lösel y Körfel (1989)	Whitehead and Lab (1989)	Gottschalk et. al (1987)	Garrett (1985)
Programas residenciales				0		
Intervenciones especializadas e inespecíficas				0	0	
Intervenciones educativas y vocacionales					0	
Psicoterapia de grupo					0	
Asistencia social y "probation"					0	
Psicodinámicas						0
Habilidades de vida						0
Otras concepciones teóricas	0					0

Tabla 5.4. Variables resultados en diferentes meta-análisis. En el encabezamiento de los columnas se referencian los diversos meta-análisis de programas correccionales llevados a cabo hasta la fecha. En el encabezamiento de las filas aparecen distintos tipos de medidas-resultado. Para cada meta-análisis se indica -mediante un círculo- qué tipos de medidas resultado utilizó.

META-ANÁLISIS MEDIDAS RESULTADO	Lipsey (1992)	Izzo y Ross (1990)	Lösel y Körfel (1989)	Whitehead and Lab (1989)	Gottschalk et. al (1987)	Kaufman (1985)	Garrett (1985)
Delincuencia / Reincidencia	0	0	0	0	0	0	0
Medidas psicológicas	0		0		0		0
Ajuste interpersonal	0				0		
Participación escolar	0				0		0
Resultados académicos	0				0		0
Logros profesionales/laborales	0				0		0
Conducta relativa a la participación en el propio programa					0		
Contactos con la policía						0	
Peticiones al tribunal						0	
Ajuste institucional							0
Ajuste comunitario							0
Otras							0
Puntuación de ajuste global	0				0		

5.4. Principales resultados obtenidos

Lösel (en prensa) ha revisado y resumido los resultados obtenidos por los diversos meta-análisis sobre tratamiento de la delincuencia efectuados hasta la fecha. De su trabajo, reproducimos parcialmente la siguiente tabla, en la que el autor ofrece los TEs medios de cada meta-análisis, manteniendo la puntuación original cuando como TE se había calculado un coeficiente de correlación (ϕ o r_m) y transformando el coeficiente d en $d/2$ (véase Tabla 5.5).

Tabla 5.5. Tamaños del efecto promedio (TEs) en diversos Meta-Análisis sobre tratamiento de la delincuencia juvenil y adulta; coeficientes r_m o ϕ .

Meta-análisis	N (estudios)	TE (r)
Garrett (1985) Tratamiento de delincuentes juveniles en contextos residenciales	111	.18
Gensheimer, Mayer, Gottschalk, y Davidson II (1986) Programas de <u>diversion</u> o derivación para delincuentes juveniles	35	.10
Gottschalk, Davidson II, Gensheimer, y Mayer (1987) Programas comunitarios para delincuentes juveniles	90	.10
Whitehead y Lab (1989) Tratamiento de delincuentes juveniles	50	.12
Lösel y Köferl (1989) Prisiones Socioterapéuticas para delincuentes adultos	18	.11
Andrews, Zinger, Hoge, Bonta, Gendreau, y Cullen (1990) Tratamiento de delincuentes juveniles y adultos	154 (80)	.10
Izzo y Ross (1990)	43	-
Lipsey (1992a) Tratamiento de delincuentes juveniles	397	.05-.08

Fuente: parcialmente tomado de Lösel (en prensa).

Como puede verse en la Tabla 5.5, el TE promedio se sitúa en torno a .10. El propio Lösel (en prensa) ha sugerido algunas claves interpretativas de este resultado. En otros ámbitos, en psicoterapia por ejemplo, los TEs suelen ser del doble o del triple de los aquí obtenidos. Sin embargo, debe tomarse en consideración la cuestión a la que nos estamos refiriendo: es, no cabe duda, muy difícil intervenir en conducta antisocial, a menudo con delincuentes con fortalecidas carreras criminales, y en contextos y condiciones frecuentemente adversos. Pese a ello, un TE de .10 favorable a los sujetos tratados supone, en términos relativos, una mejora importante, que, si asumimos una reincidencia promedio del 50% en grupos no tratados, implicaría una reducción al 40% de la tasa de reincidencia para los grupos tratados.

RESUMEN

Hemos efectuado un repaso sistemático de los meta-análisis sobre tratamiento de la delincuencia que nos han precedido, realizados entre 1985 y 1992. Para ello hemos fijado nuestra atención en cuatro componentes críticos de todas estas investigaciones. En primero lugar hemos descrito los criterios de selección de estudios que utilizaron, en cada caso, para la colección de los materiales que analizaron. Destaca el hecho de que casi todos estudiaron programas de tratamiento con delincuentes juveniles. Sólo dos trabajos incluyeron también estudios con adultos. Asimismo, en general requirieron la existencia de diseños de evaluación experimentales, con grupos equivalentes.

En relación con la codificación de los tipos de tratamiento, algunos estudios optaron por clasificar éstos en las concretas técnicas aplicadas (psicoterapia individual, de grupo, "probation", modelado, contrato conductual, ...); otros, en cambio, los agruparon según sus modelos teóricos de base (terapias psicodinámicas, conductuales, sanciones penales, ...).

Como medidas resultado se catalogaron muy variados indicadores relativos a la reincidencia, medidas psicológicas, de ajuste interpersonal, ajuste académico y laboral, etc.

El resultado promedio de los distintos meta-análisis revisados por nosotros se situó en torno a un TE de .10, lo que podría equipararse a una reducción aproximada de 10 puntos en la tasa de reincidencia de los grupos tratados frente a los no tratados.

6. ALGUNAS CONCLUSIONES SOBRE LAS REVISIONES PRECEDENTES

Trabajos recientes han abordado la revisión de los meta-análisis de programas correccionales habidos hasta la fecha. En 1992 se publicó en los Estados Unidos un estimulante libro titulado "The Re-Emergence of Correctional Intervention", de Ted Palmer (Palmer, 1992), que, en conjunto, aporta una visión positiva y esperanzada de la rehabilitación correccional. El autor dedica un capítulo específico al análisis de las diversas revisiones y meta-análisis llevados a cabo en Norteamérica en este sector temático. Otros autores han extraído también conclusiones al respecto de las revisiones y meta-análisis realizados, ya sea dentro de trabajos meta-analíticos propios (Whitehead y Lab, 1989; Lipsey, 1992a, 1992b) ya sea en estudios de revisión propiamente dichos (McGuire, 1992; Antonovicz y Ross, 1992; Lösel, en prensa). Destinaremos este apartado al comentario de algunas de estas conclusiones.

6.1. Problemas metodológicos que presentan los meta-análisis

En un recientísimo trabajo, Lösel (en prensa) ha incidido en los problemas metodológicos que debe afrontar el actual trabajo meta-analítico de programas de tratamiento de la delincuencia. En síntesis, estas dificultades serían las siguientes:

1- En relación con las características de los diseños, pocos estudios presentan el requisito metodológico de asignación aleatoria de los sujetos; mayoritariamente los programas evaluados utilizan diseños cuasi-experimentales, que comportan evidentes dificultades para una

interpretación causal de los resultados. Incluso, cuando se consigue plantear un diseño experimental, suelen producirse elevadas tasas de mortalidad -por abandonos, traslados de sujetos encarcelados, etc.- lo que termina rompiendo la aleatorización. En algunos casos, los sujetos que "abandonan" el programa suelen ser "aprovechados", en términos evaluativos, para formar grupos de comparación, con el riesgo de sesgo que ello supone. Pese a estos inconvenientes, Andrews et al. (1990) y Lipsey (1992a) han encontrado en sus respectivos meta-análisis que las diferencias en los resultados guardan una baja relación con las características del diseño de investigación. Las dificultades mencionadas y conclusiones como las de Andrews et al. (1990) y Lipsey (1992), han llevado a Lösel (en prensa) y a Hodgins et al. (1990, septiembre) a proponer la necesidad de que en contextos correccionales se toleren más flexibles estrategias evaluativas, que las requeridas por los diseños experimentales; más concretamente, se ha propuesto la combinación y síntesis entre estrategias longitudinales y experimentales.

2. En relación con el lugar o contexto del tratamiento las meta-evaluaciones realizadas han permitido concluir que apropiadas intervenciones psicológicas pueden resultar también efectivas en contextos residenciales y en prisiones. Sin embargo, también es verdad que algunos efectos de tratamiento tienen un alcance más limitado en las prisiones que en la comunidad. Una dificultad consiste en interpretar adecuadamente esta observación: discernir si ello es debido a que los encarcelados son delincuentes más consolidados, tienen más problemas de personalidad o psicopatologías, o si lo que ocurre es que, pese a la aplicación de un nuevo programa, en las cárceles continúa, por encima de todo, predominando el empleo del castigo, con los consiguientes efectos del fenómeno prisionización.

3. Se constata la existencia de grandes lagunas en la fundamentación teórica de los programas aplicados. Frente a ello, Lösel (en prensa) afirma que las "teorías empíricamente fundamentadas son necesarias no sólo para el diseño del tratamiento sino también para su aplicación y la medida de los resultados" (p.14); sin embargo, "se ha mostrado repetidamente que los conceptos de tratamiento han sido aplicados tan pobremente que no debe sorprender su inefectividad [cuando ocurre]" (p.15).

4. Con demasiada frecuencia existe una gran parquedad en la medida de los criterios efecto o resultado. Muchas veces una medida de reincidencia, durante un corto intervalo de tiempo, constituye el criterio exclusivo de efectividad. Frente a ello, deberían utilizarse más adecuados intervalos de seguimiento, así como una graduación de la gravedad de los nuevos delitos cometidos. Paralelamente, algunos estudios emplean exclusivamente criterios evaluativos de conducta institucional -por ejemplo, reducción del número de informes disciplinarios- o medidas de personalidad -percepción del clima institucional, extroversión, etc.-, que pueden resultar no válidos en términos criminológicos. La tercera dificultad estriba en la presunción, no fundamentada, de muchos estudios, al esperar que los efectos del tratamiento se mostrarán de manera consistente en diferentes criterios evaluados. Es, por ejemplo, muy corriente pensar que ciertas variables de personalidad se relacionarán con la conducta futura de los sujetos, lo que no siempre sucede.

5. Tal vez una de las mayores dificultades metodológicas señaladas por Lösel (en prensa) resida en los factores contextuales no controlables. "La mayor parte de los estudios de evaluación [correccional] todavía analizan programas de tratamiento relativamente aislados y específicos. Sin embargo, la investigación y la práctica sugieren que el contexto organizacional específico, las competencias y rivalidades, el ambiente social de la institución, el apoyo del director de la cárcel, y otros factores semejantes, pueden tener un gran impacto en las actividades cotidianas y en su efectividad" (Lösel, en prensa, p.14). La evaluación de los factores de contexto apuntados por Lösel (en prensa) comporta problemas metodológicos de difícilísima solución.

6. En síntesis, el meta-análisis ha supuesto un avance evidente frente a otras formas de meta-evaluación, como las revisiones narrativas o el recuento de votos de significación. Sin embargo, un adecuado meta-análisis viene requerido de un buen número de estudios primarios con apropiada y documentada metodología, no siempre presente en los trabajos. Por ello, en opinión de Lösel (en prensa), continúan siendo necesarias las revisiones cualitativas, para la integración de aquellos estudios con dificultades metodológicas insalvables.

7. Un último problema dimana de la imposibilidad actual de integrar dentro del meta-análisis los estudios de caso único, por más pulcritud metodológica que contengan. Sus resultados no resultan comparables con los obtenidos en estudios muestrales, y, de entrada, su integración debe ser descartada. Este problema constituye, sin duda, un reto importante que los metodólogos deberán abordar en los próximos años.

6.2. Dificultades a la hora de comparar entre sí los diversos meta-análisis

Hasta el presente, en Norteamérica (USA y Canadá) se han llevado a cabo 15 revisiones generales (cualitativas y síntesis estadísticas) de la literatura correccional, 8 específicas (de áreas temáticas, por ejemplo, programas de "probation", etc.) y 9 meta-análisis, mientras que en Europa la revisión más importante es, según hemos visto, el meta-análisis de programas alemanes realizado por Lösel y Köferl (1989). Palmer (1992), Lipsey (1992a) y Lösel (en prensa) han enfatizado algunas de las dificultades existentes en la comparabilidad de las revisiones y meta-análisis realizados hasta ahora.

La comparación entre sí de las distintas revisiones de programas comporta dificultades y discrepancias paralelas a las que supone la comparación de los propios programas, dentro de una revisión concreta. La primera razón de estas discrepancias puede hallarse en que los programas analizados en distintas revisiones no fueron enteramente los mismos, variando completamente o en parte de una revisión a otra. Incluso cuando algunas revisiones dicen ser totalmente idénticas a otras, de hecho sólo fueron similares en un 70-80% de los estudios analizados. Ello se debe bien a la no estricta coincidencia de los períodos temporales abordados por estas meta-evaluaciones, o bien a los diferenciales enfoques seguidos para la recolección de los estudios, ya que unas pusieron mayor énfasis, por ejemplo, en programas aplicados en la comunidad y otras en programas institucionales. Estas diferencias de muestreo pueden conducir a los autores a resultados dispares (según nos comentaba Lipsey, 1992a, a través del símil del elefante, que, dada su voluminosidad, sería descrito de modos diversos por observadores que sólo atendieran a la observación de un sólo sector de su cuerpo).

Pero cabe ir, incluso, más lejos. En ocasiones, meta-análisis semejantes, con resultados parecidos, llegan, sin embargo a conclusiones disímiles. Éste fue el caso de los meta-análisis de Whitehead y Lab (1989) y Andrews et al. (1990), que incluyeron una muestra semejante de estudios, y obtuvieron similares TEs (Whitehead y Lab, 1989: .12; Andrews et al., 1990: .10). Pese a ello, su teorización e interpretación estadística fueron diferentes: Para Andrews et al. (1990) un TE promedio de .10 resultó aceptable en términos de reducción de la delincuencia; en cambio, Whitehead y Lab (1989) situaron muy alto el listón de TE requerido, por lo que consideraron que un TE promedio de .12 fue irrelevante.

En segundo lugar, los acercamientos conductuales o cognitivo-conductuales -y también otros grupos de programas- fueron definidos con distinta amplitud en diferentes meta-análisis (unos incluían componentes profesionales/vocacionales, o habilidades sociales, etc., y otros no). Consideramos que esta segunda dificultad, resaltada por Palmer (1992), tiene una extraordinaria importancia, ya que con distintos agrupamientos de las intervenciones, finalmente será muy difícil comparar unas con otras. Del mismo modo que sería complicado comparar, para un ingeniero aeronáutico, las categorías de (1) los aeroplanos de hélice (que incluye sólo los denominados aviones), con la de (2) las aeronaves de hélice (que incluiría aviones, helicópteros y dirigibles). Muy probablemente, este punto tiene un valor crítico para nuestra investigación sobre programas de tratamiento en Europa. Nuestra categorización de los programas debería ser de tal suerte que, por un lado, permitiera agrupaciones exclusivas de programas, y, por otro, posibilitara futuras comparaciones con estudios semejantes.

Otro aspecto que debemos considerar es si las intervenciones han sido unimodales o multimodales, por lo que se refiere a los componentes interventivos que las integran. Es posible que programas incluidos en una categoría concreta -terapia individual, conductual, etc.- no fueran de hecho literalmente unimodales, y, en cambio, utilizaran más de un acercamiento diferente. En la práctica, lo más corriente es que los programas incluyan diversos componentes, y sean, por ello, más multimodales que unimodales. Palmer (1992) considera que, realmente, los programas sólo han funcionado cuando han tenido un cariz multimodal, integrando diversos componentes en la intervención. Las categorías genéricas suelen funcionar como

abstracciones, que facilitan la descripción de unidades individuales variadas, tanto en su concepción como en su gradación. Por ello, antes de sacar conclusiones sobre la efectividad o no de una determinada categoría de programas, Palmer (1992) sugiere que debería ser indagada la relación sustantiva existente entre: (a) los componentes cuya combinación se produjo en el programa, y (b) la comparación de dos o más categorías cuyos componentes puedan, en parte, solaparse.

En un capítulo anterior diferenciamos en sendas tablas dos niveles u órdenes superpuestos, utilizados en los distintos meta-análisis previos, de clasificación de los tratamientos e intervenciones utilizados. La clasificación de primer orden la referimos a aquellas taxonomías que habían utilizado como criterio agrupador básico la propia intervención o técnica, mientras que en la de segundo orden incluimos las tipologías más relacionadas con las teorías o modelos teóricos de soporte de tales intervenciones. Una intervención o técnica haría referencia a las concretas acciones emprendidas como aplicación de un programa. Por teoría se entiende, en cambio, "un principio general o una colección de principios generales interrelacionados que es propuesta como explicación de una serie de hechos conocidos y descubrimientos empíricos" (Reber, 1985, p. 769). Mientras que un modelo [teórico] puede ser definido como "... una suerte de miniteoría, una caracterización de un proceso y, como tal, su valor y utilidad dimana de las predicciones que pueden construirse a partir de ella y de su papel para dirigir y desarrollar teoría e investigación" (Reber, 1985, p.447). Según ello, en nuestro caso, esperaríamos encontrar teorías o modelos psicológicos o educativos habituales, tales como la "teoría del aprendizaje cognitivo", el "modelo conductual", el "modelo dinámico", etc.

Esta discriminación entre clasificaciones tecnológicas y teóricas no se hace tan evidente, sin embargo, en la práctica. En muchas taxonomías, un nivel y otro se entremezclan y solapan, según tendremos ocasión de comprobar dentro de nuestra propia investigación.

Como resultado de las dificultades existentes en la comparación entre sí de las diversas revisiones, Palmer (1992) ha enfatizado una sugerencia importante para el futuro: en los futuros meta-análisis se debería promover una descripción mucho más detallada de los diversos programas en juego,

aunque tales programas, en un esfuerzo analítico, sean agrupados en categorías genéricas. Descubrir estas combinaciones de factores interventivos, detectando cuáles de ellos fueron más relevantes para el éxito de los programas, es uno de los objetivos más importantes para el futuro tratamiento de la delincuencia.

Estimulados por esta razonable sugerencia, uno de nuestros objetivos en esta investigación fue, precisamente, la descripción detallada de los diversos componentes interventivos, que habrían sido utilizados en los diferentes programas evaluados. Para la codificación adecuada de los mismos se reservaron, en nuestro estudio, más de 40 ítemes, que indagaban la descripción de las diversas acciones emprendidas, su duración, frecuencia y periodicidad, y su relevancia en el conjunto del tratamiento.

6.3. Conclusiones sobre las revisiones y meta-análisis previos

El propio Palmer, resume cuáles son las principales conclusiones que, de lo realizado hasta ahora, pueden extraerse:

1. Cuando se toman en consideración conjuntos indiferenciados de programas, sin discriminar entre entidades y abordajes distintos (como por ejemplo, programas conductuales, intervenciones familiares, counseling, etc.), estos conjuntos se muestran, globalmente considerados, como inefectivos o poco efectivos en la reducción sustancial de la reincidencia de los sujetos. De hecho, esta falta de efectividad aparece, asimismo, para grupos específicos de aplicaciones, como los siguientes: confrontación, intervenciones preventivas de amplio espectro en barrios enteros o grandes áreas urbanas, diversion, terapia/counseling de grupo, terapia/counseling individual, probation, y parole.

2. Pese a lo anterior, muchos programas concretos consiguen importantes mejoras en los grupos experimentales frente a los controles. En concreto, con un criterio de $p < 0.5$, entre un 25 y un 35% de los programas logran significativos mejores resultados para los sujetos experimentales; mientras que lo contrario -la producción de mejores resultados en los grupos controles frente a los experimentales-, sólo tiene lugar en un porcentaje menor al 10%

de los estudios. Si atendemos, específicamente, a la variable reducción de la reincidencia ($p < 0.5$), los grupos experimentales superan a los controles en un 65% de los programas, mientras que lo opuesto sucede en un 30% de los casos.

3. En este mismo acercamiento global, las categorías de programas que suelen aparecer en las distintas revisiones como más efectivas -produciendo mejores resultados en los grupos experimentales que en los controles- son las siguientes: programas conductuales (Whitehead y Lab, 1989), cognitivo-conductuales o cognitivos, habilidades específicas o habilidades de vida, e intervenciones familiares.

Pese a lo establecido en los puntos 1. y 3. -la ineffectividad y efectividad de ciertas modalidades de programas- es también cierto que la efectividad o no de un determinado programa guarda una relación compleja con el grupo categorial en el que éste ha sido inscrito. En todas las categorías creadas por los revisores de la literatura, hay algunos programas que son efectivos y otros que no lo son. Sin que existan, en opinión de Palmer, categorías que, por su gran efectividad, puedan ser recomendadas con carácter general. Aunque, según hemos dicho, algunas resultan más esperanzadoras.

6.4. ¿Qué es un programa de tratamiento efectivo?

Hasta ahora nuestra pregunta principal ha sido: ¿Son efectivos los programas de tratamiento de la delincuencia? A la luz de las revisiones efectuadas hemos podido responder a esta cuestión en modo afirmativo, aunque parcial: algunos programas resultan moderadamente efectivos. Esta respuesta parcial nos conduce inexorablemente a una nueva pregunta: ¿Cuáles son los programas efectivos?

Su contestación ha sido uno de los principales objetivos que se han planteado los revisores de la literatura correccional durante los últimos años, y, de modo especial, los autores de las investigaciones meta-analíticas. El camino básico para ello ha sido tratar de identificar características o componentes comunes a los programas que mostraron mayor efectividad; y, paralelamente, detectar aquellos otros componentes o condiciones más comúnmente asociados

con la ineficacia. Esta perspectiva analítica habría de conducirnos a un resultado de altísima validez social en materia de tratamiento de la delincuencia: la confección del "retrato robot" de qué es un programa efectivo. De este modo podrían concentrarse esfuerzos en la aplicación de programas efectivos, evitando la dispersión en iniciativas de nula o baja eficacia.

La tarea no resulta sencilla aunque debe acometerse de modo inaplazable. Para ello seguiremos dos caminos complementarios. El primero será considerar, a efectos de ordenación de factores, los grandes ámbitos de características que pueden hallarse teóricamente relacionadas con la efectividad/inefectividad de un programa de tratamiento de la delincuencia, a saber: (1) las características del tratamiento aplicado, (2) las condiciones de los sujetos tratados, (3) las circunstancias del contexto o ambiente de la intervención, (4) los factores metodológicos concernientes a la evaluación del tratamiento, y, por último, (5) el tipo de variable resultado medidas. Todos estos sectores constituyen eventuales fuentes de variabilidad asociada a los resultados de una intervención, y de este modo han sido considerados en la investigación meta-analítica (Sánchez-Meca, 1990b; Lipsey, 1992a). Siguiendo este esquema, pasaremos revista a las principales conclusiones extraídas por los revisores correccionales al respecto de factores que han mostrado mayor asociación con la efectividad de las intervenciones. Algunos autores han realizado esfuerzos directamente encaminados a la detección de estos factores en los programas efectivos (Gendreau y Ross, 1979; Ross y Fabiano, 1985; Bartollas, 1985; Palmer, 1992; McGuire, 1992; Lösel, en prensa; Antonowicz y Ross, 1992); otros, simplemente han concluido la asociación de factores-efectividad tomando como base sus propias revisiones de literatura correccional (Garrett, 1985; Gottschalk et al., 1987; Whitehead y Lab, 1989; Andrews et al., 1990; Izzo y Ross, 1990). A partir de todos ellos en la Tabla 6.1 hemos ensayado una propuesta de "retrato robot" de aquellos programas que han mostrado una mayor efectividad en el tratamiento de los delincuentes. En la Tabla 6.1 se presentan, en primer lugar -en los apartados positivos- diversos factores que han mostrado una asociación importante con la efectividad; posteriormente se consignan, dentro de cada grupo de factores, las características de algunos programas que han resultado negativas o contradictorias en términos de efectividad.

Tabla 6.1. Componentes principales de los programas efectivos.

(1) Factores del tratamiento

Positivos

- Modelo conceptual sólido (Antonowicz y Ross, 1992).
- Programas cognitivo-conductuales (Garrett, 1985; McGuire, 1992; Antonowicz y Ross, 1992).
- Programas conductuales (McGuire, 1992).
- Programas familiares (Garret, 1985).
- Programas estructurados, claros y directivos (McGuire, 1992).
- Integridad del tratamiento: aplicación según lo planeado (McGuire, 1992).
- Mayor duración del tratamiento (Gottschalk et al., 1987; Izzo y Ross, 1990).
- Mayor intensidad del tratamiento (Lipsey, 1992a; Antonowicz y Ross, 1992).
- Abordan estilos de aprendizaje y habilidades de los delincuentes (Antonowicz y Ross, 1992).
- Incluyen técnicas susceptibles de influir los modos de pensamiento de los delincuentes: role playing y modelado, y entrenamiento en habilidades sociales (Garrett, 1985; Ross y Fabiano, 1985; Izzo y Ross, 1990; Andrews et al., 1990; Antonowicz y Ross, 1992).
- Programas multifacéticos, que incorporan diversas técnicas (Antonowicz y Ross, 1992).

Negativos

- Grupos guiados de interacción positiva de iguales (Garrett, 1985; Andrews et al., 1990).
- Programas "milieu", o de acomodación ambiental a las necesidades terapéuticas de los sujetos (Garrett, 1985; Andrews et al., 1990).
- Terapias individuales y grupales (excepto para terapia familiar) (Garrett, 1985).
- Programas de muy corta duración (Gottschalk et al., 1987).
- Programas basados en modelos de disuasión, psicodinámicos y sociológicos (Andrews et al., 1990; Lipsey, 1992a; Antonowicz y Ross, 1992).

Contradictorios

- Programas de economía de fichas (Antonowicz y Ross, 1992).
- Programas de contrato de contingencias (Antonowicz y Ross, 1992).
- Programas capaces de transmitir a los sujetos una filosofía de vida, un nuevo sentido, una misión vital, etc. (Antonowicz y Ross, 1992).
- Programas que incluyen castigos y recompensas lógicos (Antonowicz y Ross, 1992).
- Inclusión de atención posterior o seguimiento de los sujetos (Gendreau y Ross, 1979; Bartollas, 1985).

(2) Factores de los sujetos

Positivos

- Se dirigen a delincuentes de "alto riesgo" (Andrews et al., 1990; McGuire, 1992).
- Se dirigen a las necesidades de los sujetos directamente conectadas con la delincuencia (Ross y Fabiano, 1985; Andrews et al., 1990; Antonowicz y Ross, 1992).

Contradictorios

- Programas específicos para delincuentes específicos (Antonowicz y Ross, 1992).
- Motivación de los sujetos (Antonowicz y Ross, 1992).
- Participación de los sujetos en la planificación del programa (Bartollas, 1985).

(3) Factores del contexto de la intervención

Positivos

- Programas comunitarios si son potentes y aplicados con integridad (Gottschalk et al, 1987).
- Las estructuras físicas y sociales del contexto de la intervención influyen -facilitan/dificultan- los resultados (Gendreau y Ross, 1979; Redondo, 1993a).

Contradictorios

- Basados en la comunidad versus programas institucionales o residenciales (Gottschalk et al., 1987; Lösel y Köferl, 1989; McGuire, 1992; Lösel, en prensa; Antonowicz y Ross, 1992).
- Los programas deben evitar entrar en conflicto con los responsables de las instituciones correccionales (Antonowicz y Ross, 1992).
- Implicación de los investigadores en el diseño, aplicación y control de las intervenciones (Antonowicz y Ross, 1992).
- Cualificación profesional y/o experiencia de los operadores del programa (Gendreau y Ross, 1979).
- Supervisión adecuada de los operadores del programa.
- Motivación de los operadores del programa (Gendreau y Ross, 1979; Bartollas, 1985).
- Equipos de tratamiento integrados (Bartollas, 1985).
- Buena relación interpersonal profesionales-sujetos (Andrews et al., 1990).

(4) Factores metodológicos

Positivos

- Mayor efectividad de los estudios no-aleatorios (Whitehead y Lab, 1989).

(5) Objetivos y tipos de variables resultado medidas

Positivos

- Se dirigen al comportamiento delictivo y otras conductas relacionadas con éste (Andrews et al., 1990; McGuire, 1992).
- Muy influidas las variables de ajuste institucional (Garrett, 1985).
- Muy influidas las variables de mejora académica (Garrett, 1985).

RESUMEN

En el presente capítulo hemos repasado algunas de las dificultades y problemas que presentan tanto la integración de estudios correccionales dentro de un meta-análisis, como la comparación de los meta-análisis realizados entre sí.

Muchos estudios evaluativos del tratamiento de la delincuencia tienen problemas metodológicos que afectan a los diseños de investigación, en donde existe una infrecuente aleatorización de las muestras y una gran mortalidad experimental. Asimismo, en otros muchos, se producen grandes lagunas en la conceptualización teórica de los programas aplicados. No son menos los que carecen de suficiente información sobre las medidas o criterios resultado evaluados. Además, en los programas con delincuentes intervienen, con frecuencia, gran cantidad de factores contextuales o logísticos -el ambiente de las instituciones de reforma, los apoyos o rechazos oficiales a los programas, etc.-, que se solapan con las técnicas o acciones concretas del programa, y cuyos efectos resultan de difícil control evaluativo.

Por su parte los diversos meta-análisis realizados también comportan dificultades de comparación debido a que han incluido muestras de estudios diferentes, correspondientes a contextos y períodos temporales distintos. O bien a causa de una dispar taxonomización en ellos de las variables implicadas: tipos de tratamiento, modelos teóricos subyacentes, variables resultado, etc.

Pese a todas estas dificultades, la mayoría de los revisores coinciden en una conclusión general (Lipsey, 1992a y 1992b; Palmer, 1992; Taylor, 1992; Lösel, en prensa): los meta-análisis más recientes contradicen el "nothing works" afirmado por la investigación más antigua; muchos programas de tratamiento de la delincuencia resultan efectivos con amplios grupos de delincuentes. Según ello, "La falsa noción nada da resultado se encuentra ya a punto para el cubo de basura de la historia correccional" (Dilulio, citado por Taylor, 1992, p.133).

Finalmente, tomando como base las revisiones precedentes, hemos extractado todos aquellos factores que aparecen conectados con la

efectividad/inefectividad de los programas de tratamiento. En la Tabla 6.1 hemos consignado todos estos factores, agrupándolos en cinco ámbitos: (1) las características del tratamiento aplicado, (2) las condiciones de los sujetos tratados, (3) las circunstancias del contexto o ambiente de la intervención, (4) los factores metodológicos concernientes a la evaluación del tratamiento, y, por último, (5) el tipo de variable resultado medida.

**SEGUNDA PARTE: LA
INVESTIGACIÓN**

7. OBJETIVOS DE INVESTIGACIÓN Y SELECCIÓN DE ESTUDIOS

7.1. Objetivos de investigación y temporalización

La información hasta ahora presentada y debatida nos lleva a plantear esta investigación con los siguientes objetivos: (1) Recoger e integrar diferentes programas de tratamiento de la delincuencia llevados a cabo en Europa (bien en prisiones, instituciones juveniles, sistemas de "probation", etc.), hayan sido o no publicados; (2) Analizar la eficacia relativa de estos programas, comparándolos entre sí en forma cualitativa, y, sobre todo, mediante su integración cuantitativa a través de la técnica del meta-análisis; y (3) Determinar qué componentes de las propias técnicas interventivas, de sus contextos de aplicación, de los sujetos tratados, y de las características metodológicas de los estudios, dan mayor cuenta, y en qué medida, de los resultados que obtengamos. En suma, de todos estos análisis esperamos poder extraer conclusiones relativas a cuáles son los principales programas aplicados durante la última década en Europa, y cuáles resultan más efectivos y por qué.

Inicialmente se diseñó un calendario de actuaciones, para desarrollar el proceso investigador, que se proponía su finalización antes de mayo de 1993. Sin embargo este calendario ha sufrido diversas demoras, debidas, especialmente, al retraso en la obtención de algunos materiales, y al laborioso proceso de vaciado y codificación de estudios. Finalmente, el desarrollo de esta investigación ha seguido las siguientes etapas:

Septiembre-91/Junio-92. Localización y recogida de informes sobre programas correccionales aplicados en países europeos, siguiendo los mecanismos de búsqueda bibliográfica a que nos referiremos más adelante.

Febrero-91/Abril-92. Preparación del primer borrador de la "Plantilla de Codificación" de datos y de los procedimientos informáticos, que servirán para el análisis de la información.

Abril-92/Junio-92. Estudio piloto: tanto la primera "Plantilla de vaciado y Codificación" como los procedimientos informáticos para el análisis de los datos serán sometidos a un pilotaje en el que se incluirán al azar un 10% de los trabajos recogidos. A partir de este proceso se dará la forma definitiva a la "Plantilla de codificación".

Junio-92/Junio-93. Codificación de la globalidad de trabajos recogidos por codificadores expertos: uno de ellos -codificador principal- será el autor de esta investigación, y como segundos codificadores se contará con la ayuda de reconocidos expertos en la materia, tanto desde el punto de vista sustantivo -tratamiento de la delincuencia- como desde la vertiente metodológica. Más concretamente han ejercido este papel de co-codificadores mis directores de Tesis, los doctores María Teresa Anguera (Universidad de Barcelona), y Vicente Garrido (Universidad de Valencia); y para la codificación de resultados y el cálculo de los Tamaños del Efecto me han ayudado el doctor Julio Sánchez Meca y el Prof. Fulgencio Marín, ambos de la Universidad de Murcia. Estas codificaciones repetidas, tanto de elementos sustantivos como de resultados, han permitido la realización de una serie de estudios de fiabilidad a los que nos referiremos a lo largo de este trabajo.

Junio-93/Diciembre-93. Análisis final de los resultados y terminación del informe de investigación.

7.2. Futuro desarrollo de esta investigación

Esta investigación viene integrada por dos procesos básicos: el primero consiste en la confección de la "Plantilla de vaciado y Codificación" de programas, y de los procedimientos de análisis de todos los campos de información recogida; el segundo se dirige propiamente a la codificación y análisis de todo el material sobre programas europeos de tratamiento de la delincuencia, hasta este momento localizado, correspondiente al período temporal 1980-1992.

La culminación del último de estos apartados finalizaría el actual proceso de investigación, mediante la emisión del informe correspondiente al conjunto de material evaluado hasta 1992. Sin embargo, el método de codificación y análisis de la información se está preparando con el objetivo más ambicioso de trascender este primer período, y permitir la continuidad del proceso investigador del siguiente modo: (1) Futura localización de programas a los que no hemos tenido acceso en esta primera fase (especialmente, iniciativas de los países europeos más alejados de nosotros, como los Países Nórdicos o los Países del Este), y de futuros programas que se lleven a cabo; (2) Codificación de estos programas, mediante los instrumentos ya diseñados, y consiguiente incremento de nuestra muestra de estudios; y (3) Crecientes análisis de información de los nuevos conjuntos de muestras de programas codificados.

7.3. Pautas para la localización de estudios en una revisión

El primer problema con el que se enfrenta un revisión de literatura científica consiste en la búsqueda y localización de los trabajos que serán revisados. Tal y como ha señalado Gómez (1987), en la tarea inicial de localización de los estudios debe tenerse "una mínima seguridad de que los estudios revisados representan a la mayoría de los existentes" (p.31). Así pues, la búsqueda de estudios es comparable al muestreo en una investigación primaria. Y el procedimiento seguido para ello deberá ser objeto de una minuciosa descripción (Gómez, 1987), que permita una adecuada replicabilidad futura.

Cooper (1989) ha sistematizado las distintas fuentes con que contamos para la localización de estudios en una revisión. Ha distinguido tres grandes grupos de canales:

A) Canales informales: en este apartado el investigador cuenta con:

1. Su propia investigación sobre el tema. En algunos casos es corriente que un investigador acometa la revisión de un tema sobre el que él mismo ha trabajado previamente y del que, en consecuencia, ya tiene alguna información.

2. El "colegio invisible" ("invisible college"), término acuñado

por Price (1966) para reflejar la idea de la existencia de grupos de científicos que trabajan en el mismo campo, tienen noticia recíproca los unos de los otros, y suelen intercambiar información y copias de sus respectivos estudios.

El peligro de este último canal de suministro de información, empleado de modo aislado, es que los miembros de un "colegio invisible" tendrían una cierta uniformidad tanto en los resultados más importantes obtenidos como en la metodología utilizada. Por ello es imprescindible, si se desea una buena representatividad de los estudios, acceder a información procedente de otros canales.

3. Estudiantes y profesores intercambian información, artículos, etc., de interés recíproco.

4. Participación en reuniones profesionales, congresos, etc. (sociedades científicas diversas...). Por esta vía pueden descubrirse trabajos interesantes en el ámbito de estudio del investigador.

Una ventaja de los canales informales, frente a los formales -que veremos a continuación-, es que permiten acceder a información con menores exigencias metodológicas. Por ello, abren el acceso del revisor a estudios inferiores, o de menor rango científico, muchos de los cuales difícilmente llegarán a ser publicados.

Sin embargo, el uso exclusivo de esta vía informal resultaría también insuficiente. Sería, en símil del propio Cooper (1989), como aquel investigador primario que utilizara como muestra sólo a sus amigos.

B) Canales primarios: son aquéllos que nos permiten acceder directamente a investigaciones primarias de modo sistemático. Los principales serían:

1. Bibliotecas personales o revistas cuya publicación es seguida regularmente por el investigador en sus bibliotecas institucionales de referencia (universidad, colegios

profesionales, bibliotecas especializadas, etc.).

Ésta es, sin duda, una fuente primigenia de información que permite ir acumulando documentación paulatina sobre los ámbitos de interés de un investigador concreto.

Tal vez, su inconveniente más importante sea su uniformidad. Esto es, la constatación de que los trabajos aparecidos en la misma revista, o en revistas de un área determinada, muestran una alta focalización endógena, de modo que tienden a referenciarse recíprocamente entre sí, siendo poco permeables a la información aparecida en revistas distintas. En esta línea, Xhignesse y Osgood (citados por Cooper, 1989) hallaron que un 30% de las referencias de los estudios publicados en una determinada revista se referían a trabajos aparecidos en ella misma, y un 37% más correspondían a revistas del mismo campo.

2. Procedimiento ascendente ("ancestry approach"): el revisor puede rastrear aquellas investigaciones que han sido citadas en los estudios previamente obtenidos por él. Éste puede ser, sin duda, un buen comienzo ya que generalmente un revisor, antes de acometer un examen de literatura, ya posee trabajos sobre el campo objeto de análisis.

Sin embargo, los canales primarios comentados (bibliotecas y revistas personales, y procedimiento ascendente) no carecen tampoco de sesgos. El más importante de éstos puede ser el "sesgo de publicación". La publicación de un estudio no parece tanto depender el mérito científico de un trabajo como de sus positivos resultados (en términos estadísticos). Es conocido que existe una "selección editorial a favor de estudios con resultados significativos" (Gómez, 1987, p.32) frente a los que obtuvieron resultados nulos o no significativos. Greenwald (citado por Cooper, 1989) encontró que los investigadores tratan de publicar sus estudios en un 69% de los casos, cuando la hipótesis nula ha sido rechazada, y consecuentemente, se establece alguna suerte de relación entre las variables estudiadas; mientras que sólo lo harían para el 6% de los trabajos que aceptan la hipótesis nula, negando por ello la relación entre las variables investigadas. En suma, Rosenthal (citado por Gómez, 1987) ha estimado que no serían publicadas el

95% de las investigaciones con resultados no significativos. Esto implica que no puede presuponerse una equivalencia apriorística en Tamaño del Efecto (TE) entre estudios publicados y no publicados. Por tanto, una recogida de materiales que sólo se dirigiera a trabajos publicados podría incurrir en un fuerte sesgo.

Por otro lado, existiría también un sesgo de publicación que podríamos denominar de consistencia de la investigación con investigaciones previas: esto es, se publicarían más fácilmente estudios cuyos resultados confirman las creencias existentes que aquellos otros que entran en conflicto con las mismas. Por todo ello, para la recolección de los estudios, tampoco deberían emplearse los canales primarios como fuentes únicas de información.

Algunos de los meta-análisis que hemos revisado podrían haber incurrido en sesgos de publicación, al haber recolectado sólo estudios publicados (Izzo y Ross, 1990; Andrews et al., 1990; Lösel y Köferl, 1989; y Whitehead y Lab, 1989). Otros, sin embargo, han salvado adecuadamente este inconveniente y han recogido tanto estudios publicados como no publicados (p.e. Kaufman, 1985, y Lipsey, 1992a). En el Meta-análisis de Lipsey (1992a; 1992b), el más voluminoso en número de estudios, el sesgo de publicación no podría, sin duda, explicar los resultados obtenidos (según ha puesto de relieve Lösel, en prensa): las diferencias entre resultados favorables y desfavorables son abrumadoras, según vimos con anterioridad: en 285 estudios (el 64,3% de los analizados) los grupos tratados superaron a los controles, en 131 (el 29.6%) grupos tratados y controles obtuvieron resultados semejantes, y sólo en 27 programas (un 6.1%) fueron los controles quienes superaron a los grupos de tratamiento.

C) Canales secundarios: permiten acceder, por distintos procedimientos ordenados de localización (autores, "thesaurus", etc.), a un determinado sector de estudios, más o menos acotado. En relación con éste nos facilitan información básica sobre identificación de los autores, revistas de publicación, países de realización, organismos productores, resúmenes de los trabajos, etc. Constituyen una vía decisiva a la hora de efectuar una revisión sistemática y amplia de la literatura, permitiéndonos acceder a la mayor parte de la información disponible en una materia. Los principales canales secundarios son los siguientes:

1. Bibliografías: listados de artículos y libros relevantes en un ámbito determinado. Realizados, a menudo, por científicos o grupos de científicos interesados en un tema. Es un vía muy útil, porque generalmente un investigador ha invertido mucho tiempo reuniendo la información bibliográfica presentada.

2. Publicaciones de organismos oficiales: pueden ser una fuente interesante para ciertos sectores temáticos. Por ejemplo, si buscamos literatura española sobre delincuencia, será conveniente revisar los listados de publicaciones del Ministerio de Justicia, de la Dirección General de Protección Jurídica del Menor, del Centre d'Estudis Jurídics de la Generalitat de Catalunya, del Consejo General del Poder Judicial, de la Escuela de Estudios Judiciales, de los Institutos de Criminología, etc. Y procederíamos de modo semejante para otros países.

3. Servicios de índices y resúmenes ("indexing and abstracting services"): es hoy el método más ágil y amplio para obtener información sobre una determinada materia. Especialmente, si se dispone de la información en soporte informático, como sucede con servicios de "abstracts" tales como el Psychological Abstracts (disponible en soporte informático denominado "psyclit"), el sociological abstracts (cuyo soporte informático es el "sociofile"), etc.

Casi todo son ventajas en este canal. Mediante los servicios de índices y resúmenes podemos acceder a muchísima información en un tiempo mínimo. En varios ensayos podemos acotar progresivamente nuestro ámbito de búsqueda. Y, finalmente, obtener cientos de referencias y resúmenes de otros tantos estudios. Una pequeña limitación de este sistema es el relativo retraso en el proceso de inclusión de trabajos en los índices. Ello implica que en muchos índices de resúmenes no pueden ser localizados los trabajos publicados durante el último o últimos dos años. Éstos deberán ser localizados, en su caso, mediante la revisión directa de las revistas apropiadas.

Como hemos tenido ocasión de comentar, todos los canales de localización de

información tienen sus ventajas y limitaciones, del mismo modo que sucede con los instrumentos que utilizamos en la investigación primaria. Por esta razón, el mismo consejo de triangulación que cabe para ésta puede ser formulado para la localización de literatura sobre un ámbito. Cuantas más fuentes usemos, mayor será la probabilidad de exhaustividad, menores serán los sesgos de muestreo, y mayor equilibrio lograremos entre estudios publicados y no publicados.

Rosenthal (citado por Gómez, 1987) diseñó una fórmula para estimar el índice de tolerancia de estudios con resultados nulos -muchos de ellos probablemente no publicados- que serían necesarios para alterar los resultados de un meta-análisis:

$$N_{.05} = \{ k [k (Z_k)^2 - 2.706] \} / 2.706$$

donde N sería el número de resultados nulos no recuperados para la revisión, 2.706 el cuadrado de la puntuación típica para una significación del 5%, k el número de trabajos revisados y Z_k su puntuación típica media.

Esta fórmula ha sido transformada por Orwin (citado por Gómez, 1987) para ser calculada mediante índices de tamaño del efecto:

$$N = k (d - d_c) / d_c$$

donde N es el número de resultados nulos no recuperados, k el número de trabajos revisados, d el promedio del Tamaño del Efecto de éstos y d_c el valor criterio elegido. Como valores criterio -continúa Gómez (1987)- se proponen los señalados por Cohen: se entienden efectos pequeños para $d = .20$, efectos medios para $d = .50$ y efectos grandes para $d = .80$.

7.4. Nuestro proceso de recolección de estudios

Siguiendo las orientaciones comentadas, para llevar a cabo la recopilación de estudios evaluativos sobre programas de rehabilitación aplicados en Europa, que habrían de dar lugar a este meta-análisis, se emplearon los siguientes mecanismos básicos:

1- **Consultas en soporte informático.** Se efectuó una primera búsqueda informática, mediante el sistema DIALOG Information Retrieval Service, que permitió acceder al CRIMINAL JUSTICE PERIODICAL INDEX (CJPI), producido por la "Universidad Internacional de Microfilms". El CJPI posibilita el acceso a revistas de justicia criminal (publicadas en Estados Unidos, Gran Bretaña y Canadá) en las áreas de la criminología, la ley penal, la familia y la ley, sistemas de seguridad, sistemas penitenciarios, y policía. Para llevar a cabo esta busca de información se utilizaron las combinaciones de las siguientes palabras claves en inglés: en primer lugar, se seleccionaron una serie de términos y sinónimos de frecuente utilización en los títulos de informes relativos a programas correccionales para referirse a los sujetos destinatarios de los programas, así como en los descriptores que los autores suelen dar para la localización de sus trabajos, tales como "Delinquen*" (que cubría delinquent/s, delinquency), "Offender/s", "Inmate/s", "Prison*" (que cubría los términos prison/s, prisoners), "Probation*" (que abarcaba probation/, probationers); todos éstos se combinaron con otra serie de palabras clave, frecuentemente utilizadas, tanto en títulos de informes científicos como en los descriptores de los mismos, para significar o referirse a las acciones o intervenciones emprendidas con sujetos delincuentes, tales como "Treatment/s", "Rehabilitation", "Intervention", "Parole" y "Community"; y, finalmente, las referencias obtenidas, producto de la combinación de los dos conjuntos de términos anteriores, se restringieron mediante su combinación con la palabra "Europe". Desgraciadamente esta inicial aproximación resultó poco fructífera en la localización de programas de tratamiento aplicados en Europa entre 1980-90. Se obtuvieron en ella un conjunto de 128 referencias, de las cuales tan sólo 6 resultaron de interés para los objetivos de esta investigación.

Teniendo en cuenta este primer resultado, se llevó a cabo una segunda búsqueda informática a través de una base de datos francesa denominada PASCAL, que recoge información sobre trabajos científicos, tanto en los ámbitos de tecnología, medicina o biología, etc., como en diversas áreas de ciencias sociales, y en criminología. Para esta segunda búsqueda se utilizaron las mismas palabras claves que la vez anterior, salvo que se añadió el término "Therap*" (para significar therapy, therapeutic), y se excluyeron los términos "Comunity", considerada su gran amplitud, y "Europe", ya que se comprobó que la utilización del nombre del país o

continente en que se ha llevado a cabo un programa no es muy frecuente ni en los títulos ni en los descriptores de los trabajos. Además, se añadieron dos nuevas condiciones a esta segunda búsqueda, que fueron las claves "Not psychiatr*" (que incluye psychiatry, psychiatric), para excluir los trabajos relativos a terapias y programas desarrollados en hospitales psiquiátricos, y por lo que se refería al campo país de publicación, se consignó "Not USA", para evitar la inclusión de muchos trabajos aplicados en los Estados Unidos. Este segundo acercamiento informático a la captación de trabajos tuvo un mayor éxito que el anterior, localizándose de este modo 30 programas más, que inicialmente parecían de interés para nuestros objetivos. En la Tabla 7.1 se han recogido los diversos criterios empleados para efectuar las búsquedas informatizadas a las que nos venimos refiriendo.

Tabla 7.1. Criterios utilizados en las búsquedas informatizadas realizadas en las bases documentales "Dialog", "Psyclit" y "Pascal".

Términos relativos a los destinatarios de los programas:

- Delinquen*
- Offender*
- Inmate*
- Probation*

Términos relativos a las intervenciones de tratamiento emprendidas:

- Treatment*
- Rehabilitation
- Intervention
- Parole
- Community
- Therap*
- No psychiatr*

Términos relativos al ámbito geográfico de los estudios:

- Europe
 - Not USA, Not Canada, Not Australia
-

Una tercera búsqueda informática se realizó a través del Psychological Abstracts (en soporte informático "psyclit"), en fecha 18/6/92. Para ella se emplearon los mismos términos utilizados en las anteriores, excluyendo explícitamente de la búsqueda los trabajos realizados en Estados Unidos, Canadá y Australia. De este modo se obtuvieron inicialmente 305 referencias correspondientes al período temporal 1983-1991, y 271 del período 1974-1982. La revisión de todas estas referencias permitió la localización de 32 nuevos estudios que describían programas de tratamiento, y, en principio, eran susceptibles de inclusión en nuestra investigación. La mayoría de éstos habían sido publicados en revistas de difícil localización en España (bien por ser revistas muy especializadas, o bien por corresponder a países del Este de Europa). Por ello, se decidió enviar carta de petición de los trabajos a los propios autores, de acuerdo con las instituciones de pertenencia consignadas en los "abstracts".

Finalmente, efectuamos una segunda búsqueda de comprobación dentro del Psychological Abstracts con tal de calibrar el volumen de publicaciones psicológicas meta-analíticas en el ámbito de la delincuencia. Para ello consultamos en soporte informático "psyclit" las publicaciones recogidas entre 1974-1991. Solicitamos información sobre cuántas de ellas incluían entre sus descriptores el término meta-análisis. Aparecieron 655 referencias, con una gran dispersión temática. Sólo 12 de estos trabajos incluían también los términos delin* (delincuencia, etc.) o crim* (delito, etc.), o sea escasamente un 2%. A modo de parámetros comparativos, 12 estudios se describían también como theor* (teoría, etc.), 2 como clin* (clínica, etc.) y 62 (el 9.46%) se referían a educa* (educación, etc.), siendo éste el ámbito temático con mayor número de referencias meta-analíticas.

2- Revisión directa de revistas especializadas, libros y monografías. En segundo término, se efectuó una revisión sistemática de los números aparecidos, entre 1980-1991, de un conjunto de cincuenta y cinco revistas especializadas en Criminología y otras ciencias relacionadas, como Ciencia Penitenciaria, Psicología, Pedagogía, Trabajo Social, Derecho, etc. La mayor parte de estas revistas eran europeas, correspondiendo la publicación de las mismas a los siguientes países: 20 españolas, 6 francesas, 6 británicas, 4 italianas, 3 suizas, 1 holandesa, 1 belga, y 1 portuguesa; sin embargo también se revisaron algunas publicaciones editadas en norteamérica, de las

cuales 11 correspondían a los Estados Unidos y 2 eran canadienses (véase relación completa de revistas revisadas en el Apéndice 1).

Asimismo se revisaron también un conjunto de libros y monografías, seleccionados tanto de mi biblioteca personal como de la biblioteca del Centre d'Estudis Jurídics, especializada en temas criminológicos.

3- Petición de trabajos a expertos. En tercer lugar, se dirigió una carta, en doble versión inglés/francés, a un conjunto de 82 instituciones docentes y de investigación en los ámbitos de la criminología y los servicios correccionales de todos los países europeos (excepto Albania, de donde no poseía dirección alguna a donde dirigirme; e incluyendo a Israel, pese a no ser un país europeo⁷); estas instituciones fueron seleccionadas de las que aparecen relacionadas en A WORLD DIRECTORY OF CRIMINOLOGICAL INSTITUTES (UNSDRI, 1986). Asimismo, me dirigí personalmente a un total de 118 profesores e investigadores de criminología, psicología criminal, derecho penal, etc., y de países europeos, seleccionados de entre los participantes en una serie de recientes congresos de criminología. En esta carta, cuyo literal se incluye como Apéndice 2, informaba a los destinatarios de la misma que estaba llevando a cabo una revisión de programas correccionales aplicados en Europa, y les pedía que, si ellos personalmente habían aplicado y evaluado programas de este tipo o conocían programas realizados en sus respectivos países, hubieran sido publicados o no, que me los remitieran.

Paralelamente, aprovechando la celebración del I Encuentro Hispano-Británico de Psicología Jurídica, celebrado los días 6, 7 y 8 de junio de 1991 en Pamplona, entregué personalmente la misma carta a que me he referido (Apéndice 2) a los más de treinta psicólogos británicos participantes en dicho Encuentro. En el Apéndice 3 se relacionan las instituciones, países y personas a los que me he dirigido solicitando que me remitieran información sobre programas correccionales aplicados y evaluados.

⁷Nuestra investigación se dirigía, en principio, al ámbito territorial y cultural europeo. Sin embargo, como quiera que pudimos localizar un trabajo llevado a cabo en Israel, consideramos que no existían razones geográficas o socio-culturales suficientes para su exclusión.

Como resultado del envío de todas estas cartas recibí diversos materiales, que incluían tanto trabajos publicados en revistas especializadas y libros, como estudios no publicados. También recibí algunas cartas que me ofrecían la posibilidad de enviarme materiales en lenguas distintas del español, inglés, francés, italiano y portugués -que eran los idiomas preferentes que yo había consignado en mi carta de petición de trabajos, según Apéndice 2-, en concreto en Sueco y Hebreo, por lo que tuve que contestar de nuevo solicitando el envío de tales trabajos. Se contestó a todos los envíos y respuestas recibidas con una carta de agradecimiento.

Debemos mencionar la buena disposición hallada en algunos de los investigadores e instituciones a los que nos dirigimos, que no ahorraron esfuerzos para ayudarme en la búsqueda de literatura correccional. Por ejemplo, del Ministerio de Justicia Portugués me fue enviado un conjunto importante de números de su revista Infancia e Juventude, y el Dr. H. J. Albrecht, del "Kriminologische Forschungsberichte aus dem Max-Planck-Institute für ausländisches und internationales Strafrecht", de Freiburg (Alemania), me remitió amablemente la colección de libros de criminología editada por este Instituto, que recoge investigaciones desarrolladas en el mismo. Finalmente, la carta enviada al "Criminologisch Instituut Leiden" (Holanda) tuvo como resultado la remisión de sendas revisiones de la literatura correccional, llevadas a cabo por la Documentalista del Centro de Documentación de La Haya del Ministerio de Justicia Holandés: una de estas revisiones se refería a "programas en instituciones correccionales, prisiones, probation, etc., en Holanda (1987-1991)", y la otra a "programas correccionales en Europa (1987-1991)". En esta revisiones pude localizar otros 20 estudios potencialmente interesantes. Algunos de ellos, no publicados o publicados en revistas holandesas, hube de pedirlos a ese Centro de Documentación y me fueron amablemente enviados por la Documentalista. También es verdad que en otros muchos casos las cartas enviadas no recibieron contestación.

7.5. La heterogeneidad de los programas objeto de análisis

Cuando nos referimos al tratamiento o la intervención sobre la delincuencia estamos dirigiendo nuestra atención hacia un universo de

conceptualizaciones, procedimientos interventivos y técnicas evaluativas de la supuesta efectividad de tales intervenciones, realmente variado y dispar. Repasemos algunas de las dificultades con las que nos enfrentamos y las soluciones por las que hemos optado.

La variedad de las intervenciones. Por lo que hace referencia a los mecanismos y estrategias interventivas, con objeto de abordar la delincuencia se han utilizado -y se utilizan universalmente- tanto procedimientos meramente punitivos y de control, sobre la base de conceptos y teorías penales, tales como la privación de libertad, como más sofisticadas estrategias terapéuticas y educativas, fundamentadas en teorías o modelos psicológicos, sociológicos, etc., tales como las técnicas grupales, la modificación de conducta, la educación, el entrenamiento en habilidades sociales, los programas laborales de transición, etc. Como hemos mencionado, todo tipo de intervenciones entrarán en nuestra consideración.

Medidas de reincidencia versus otras medidas. Por otra parte, por lo que respecta a los conceptos y técnicas evaluativas utilizadas para conocer el grado de eficacia de los procesos de intervención sobre la delincuencia, también han sido diversas las medidas empleadas. No obstante, podemos diferenciar dos grandes bloques de medidas: en primer lugar, aquéllas que han tenido por objeto conocer, de modo directo, si cierta estrategia interventiva, aplicada a un grupo o población de delincuentes, producía alguna reducción o variación de su comportamiento delictivo futuro, o sea, si era afectada su reincidencia; y en un segundo término, aquellas mediciones encaminadas a detectar cambios o mejoras no directamente en la conducta delictiva de los individuos, sino en ciertos factores psicológicos y de comportamiento de los mismos, a los que se presupone alguna conexión, más o menos cercana, con la conducta delictiva futura.

Desde una perspectiva ideal, serían de nuestro interés, ciertamente, aquellos trabajos y programas que hubieran incluido entre sus medidas evaluativas algunas mediciones relativas a la delincuencia futura de los sujetos tratados. En suma, éste es el interés final de cualquier intervención sobre la delincuencia. Sin embargo, para que un programa interventivo contenga información sobre la reincidencia posterior de los sujetos que pasaron por el mismo, los investigadores han de haber llevado a cabo un

largo y costoso seguimiento de los sujetos tratados de entre uno y cinco años, o más, tras la finalización del programa, lo que no siempre sucede. Por esta razón, en la presente revisión serán incluidos tanto aquellos programas que tengan alguna medición de la delincuencia futura de los sujetos tratados, como aquellos otros que posean sólo medidas de otros factores y comportamientos de los sujetos sometidos al tratamiento.

Los diseños de investigación. El diseño de evaluación funciona como una guía que facilita la ordenación de los datos, tanto en el proceso de recogida como su almacenamiento y posterior tratamiento (Anguera y Redondo, 1991). La evaluación de programas interventivos ha sido, con frecuencia, una actividad generalmente descuidada o pobremente realizada. Una razón para ello es que muchos profesionales no dominan o no se sienten cómodos con el uso de procedimientos cuantitativos, debido a que en el campo social está muy extendido el mito de la subjetividad -la idea de que los programas sociales no son evaluables, persiguen transformaciones que no pueden verse-, o bien por escasez de tiempo y recursos (Sanz Oro, 1990).

Así pues, otro de los puntos de decisión en esta investigación tiene que ver con los criterios metodológicos mínimos que habrían de ser exigibles a los trabajos revisados. La tendencia de las revisiones llevadas a cabo en Norteamérica (por ejemplo, Lipsey, 1992a) es mantener un criterio severo, que requiere que, para su aceptación, los estudios posean un diseño de grupo control, con asignación aleatoria de los sujetos. Ello nos parece bastante lógico, si tenemos en cuenta la mayor tradición empírica que tienen tanto la psicología como la criminología norteamericana y anglosajona, en general, en comparación con la mayor pragmática teórica y discursiva que, en términos generales, tiene la criminología europea.

Esto hace que en la revisión que nos ocupa adoptemos un enfoque metodológico minimalista, que nos lleva a aceptar para su revisión aquí cualesquiera estudios evaluativos de programas correccionales sobre la delincuencia, con tal de que cumplan el requisito mínimo de haber empleado alguna estrategia interventiva -cualquiera que ésta sea-, con uno o más valores de aplicación, y hayan tomado como mínimo dos medidas de las variables que dicen pretender influir, sean mediciones directas del comportamiento delictivo o de otros factores comportamentales o psicológicos

relacionables con la conducta delictiva de los sujetos. Este amplio enfoque no excluye, obviamente, que en nuestro análisis llevemos a cabo cuantas matizaciones se hagan necesarias sobre la calidad metodológica tanto de los procedimientos interventivos aplicados como de los diseños evaluativos empleados para conocer la efectividad de los mismos. Y esto nos permitirá, sin duda, discriminar los diferentes estudios que analicemos en función de la pulcritud de sus características metodológicas.

Metodología experimental versus conductual. Una cuestión pareja, en este punto, tiene que ver con la diferenciación entre metodologías experimentales clásicas y metodologías conductuales. En la metodología experimental clásica, o correlacional, se pretende generalmente demostrar la influencia o relación de una variable X sobre otra Y mediante el empleo de dos grupos, uno denominado experimental, al que se aplica el impacto de una variable interventiva X, y otro denominado grupo de control, que no es expuesto a la variable interventiva X. Con posterioridad se emplea alguna técnica estadística dirigida a negar la veracidad de la hipótesis nula (esto es, que no existen diferencias significativas entre los grupos de tratamiento y de control, y, por tanto, que el tratamiento X no ha producido efecto alguno), afirmando, en su lugar, la hipótesis alternativa (o sea, que, tras la aplicación de tratamiento X al grupo experimental, se han encontrado diferencias significativas en las puntuaciones que los grupos experimental y control obtienen en la variable Y, lo que permite concluir que éstas son debidas al influjo de la variable X).

Por su parte, la metodología conductual no ha utilizado generalmente grupos de control puros, sobre los que no se interviene, sino que todos los sujetos o grupos han sido expuestos al influjo de un tratamiento X. En lugar de ello, el influjo de una variable interventiva X sobre una variable efecto Y se ha pretendido mostrar a través de procedimientos de control experimental que han incluido las siguientes etapas: en primer lugar, se han tomado diversas medidas de la variable Y, con anterioridad a la intervención, proceso que se conoce como línea base; a continuación, se ha expuesto a los sujetos o a los grupos al impacto de una intervención X durante algún tiempo, tomándose medidas continuadas de la variable Y; y finalmente, se ha retirado el influjo de la variable X, prosiguiendo la medición periódica del nivel de la variable Y -seguimiento sin intervención-. El control experimental, que permite

establecer algún nivel de influencia o relación entre la variable interactiva X y la variable efecto Y, vendría dado por la comprobación de que sólo cuando la variable X se encuentra presente, y sólo entonces, la variable Y cambia sus parámetros.

En la presente investigación hemos preferido no excluir ninguna de estas metodologías. También aquí hemos adoptado una perspectiva amplia, aceptando para nuestro análisis tanto trabajos con un diseño evaluativo de corte experimental clásico más estricto, que utilicen grupo de control, como estudios de grupo único, con tal de que tengan medidas de las variables dependientes antes y después de la intervención. Acogiendo igualmente trabajos con diseños evaluativos conductuales, que hayan empleado mediciones de línea base.

7.6. Criterios de recolección y selección de los estudios

Primera selección. Un punto crítico de decisión, en esta investigación, tuvo que ver con la mayor o menor amplitud de criterios para la aceptación de los estudios que, finalmente, habrían de integrar la muestra meta-analizada. A este respecto, en un primer momento se adoptó un criterio de exigencia minimalista, y se recolectaron todo tipo de trabajos con tal de que, como mínimo, formularan un proyecto o propuesta de intervención sobre sujetos delincuentes. Para esta primera etapa de recogida de estudios no se tomó en consideración ningún otro criterio restrictivo, referente, por ejemplo, a que tal proyecto hubiera sido realmente puesto en práctica, la metodología utilizada en el mismo, y la existencia o no de evaluación.

Todas las dificultades metodológicas a que nos venimos refiriendo han inclinado a algunos autores a rebajar el listón de las exigencias metodológicas en el diseño de las evaluaciones de programas con delincuentes. Tal es el caso de Hodgins, Cyr, Paquet y Lamy (1990, septiembre) que han propuesto una aproximación "minimalista" en los criterios metodológicos exigibles a los programas correccionales.

Como ya hemos comentado, nuestro primer objetivo consistía, precisamente,

en conocer las ideas, propuestas y programas de intervención, sugeridos o efectivamente realizados, que hubieren tenido lugar en los diversos países del marco europeo. Este primer amplio abordaje se justifica, en nuestra opinión, debido a la inexistencia de una panorámica de esta naturaleza, que globalice lo reflexionado o actuado en Europa en materia de rehabilitación de delincuentes. Bajo esta perspectiva se acometió el estudio piloto que sirvió para poner a prueba los instrumentos y la metodología de la investigación. Durante el pilotaje fueron codificados tanto programas que habían sido evaluados, cualitativa o cuantitativamente, como otros que, aunque puestos en práctica, no ofrecían evaluación alguna, e incluso meras sugerencias y propuesta de tratamiento, aún no aplicadas.

Esta opción inicial, en principio adecuada, presentaba, sin embargo, muchos problemas. En el estudio piloto pudimos comprobar la gran disparidad existente entre la codificación de unos estudios -los de metodología más depurada- y otros -los que eran meras ideas y propuestas interventivas-. Nuestros instrumentos (plantilla/manual de codificación/bases de datos) habían sido diseñados, con inclusión de un amplio elenco de informaciones, pensando en el número máximo de informaciones posibles, que, como es lógico, se corresponden con aquellos trabajos metodológicamente más completos. Esto se traducía en una infrautilización del aparato codificador cuando se trataba de estudios de baja calidad metodológica. Pero, por encima de todo, esta heterogeneidad en el nivel de las informaciones aportadas por diversos estudios, y de las subsiguientes codificaciones, mermaba, casi siempre, las posibilidades de análisis de cualquier ítem a lo largo de los distintos trabajos. Muchos de los ítemes sólo aparecían cumplimentados para porcentajes muy bajos de estudios, lo que impedía extraer conclusiones generales.

Con frecuencia en prisiones, centros de reforma juvenil y otros ámbitos de tratamiento de la delincuencia es muy costoso adoptar demasiadas precauciones metodológicas, tales como la introducción de grupos de control y el uso de diferentes medidas de los efectos producidos por los programas. Se trata de contextos en los que, además, ni los administradores y operadores del sistema de justicia penal tienen, por lo general, interés en la literatura criminológica científica, ni destinan recursos a la investigación de los programas y procesos de rehabilitación.

Sin embargo, no podíamos pensar -como es obvio- en la realización de un meta-análisis para aquellos trabajos que no aportaran información cuantitativa sobre sus resultados.

Segunda selección. Como consecuencia de todo ello, se efectuó una segunda selección de aquellos estudios que habrían de formar parte del meta-análisis, bajo los criterios conceptuales y metodológicos siguientes: se requirió que los programas (1) se dirigieran a sujetos -jóvenes o adultos, hombres o mujeres- sometidos a alguna suerte de control por los sistemas de justicia penal; (2) que se tratara de la aplicación de técnicas o estrategias para el logro de objetivos relacionados directa o indirectamente con la rehabilitación de los mismos, tales como su formación, su entrenamiento en habilidades de vida, la reducción de sus comportamientos violentos o delictivos, la reducción de su reincidencia posterior, etc.; (3) que se empleara y describiera algún diseño de investigación que pusiera en relación la aplicación de las estrategias interventivas con la recogida de los datos; y (4) que se ofrecieran resultados cuantitativos de los efectos producidos por el programa (véase Tabla 7.2).

Tabla 7.2. Criterios de selección de estudios para la investigación meta-analítica.

-
- (1) Grupo de sujetos bajo control de la justicia
 - (2) Aplicación de alguna estrategia dirigida a la rehabilitación
 - (3) Diseño de investigación que permita la comparabilidad
 - (4) Algún resultado cuantitativo
-

Por tanto, nuestra muestra de trabajos vendría constituida por todos aquellos programas de tratamiento de delincuentes -a que tuviéramos acceso- que hubieran sido aplicados en centros de reforma juvenil, prisiones y servicios comunitarios para el tratamiento de la delincuencia, en países europeos, desde 1980 hasta la fecha (realmente, en la práctica, que la fecha

del informe sobre los mismos se encuentre entre 1980 y 1991⁸), y reúnan los requisitos establecidos en el apartado anterior.

Así pues, procedimos a realizar una lectura rápida de todos los informes recolectados, discriminando aquéllos que reunían los criterios de meta-análisis establecidos de los que no los contenían.

Todos los estudios incluidos finalmente en nuestro meta-análisis poseían los criterios antedichos, con la excepción del criterio número (3) -que se emplee algún diseño de investigación que relacione la aplicación de la intervención con la recogida de los datos-. Unos cuantos estudios (Brown, 1985; Members of the Demonstration Unit 1981-84, 1986; Weaver y Fox, 1984; Singer, 1991; Redondo et al., 1990) reunían todos los requisitos con la siguiente salvedad: incluían alguna medida de la variable dependiente reincidencia en el post-test, pero, sin embargo, no ofrecían información de reincidencia comparable (pre-test o grupo de control). Con objeto de no perder esta valiosa información sobre reincidencia se procedió de la siguiente manera: en cada estudio con este problema, se seleccionó una medida de reincidencia estándar -sin intervención alguna- que resultara comparable para cada caso, en función del país del estudio, la edad de los sujetos, su tipología delictiva, y el período de seguimiento de la reincidencia. Esta información se extrajo de los datos de reincidencia en diversos países europeos publicados por el Consejo de Europa (Tournier y Barre, 1990) y del estudio de reincidencia de Redondo, Funes y Luque (1993). Los porcentajes de reincidencia de comparación oscilaron entre el 50% y el 70%, dependiendo, según hemos comentado, de la edad de los sujetos, su tipología delictiva, el país y el período de seguimiento.

⁸Nuestros procesos de búsqueda de información han cubierto realmente el período 1980-1991 (ambos años inclusive). Pese a ello, entre los trabajos recogidos aparecen algunos estudios de los años 1992 y 1993. Su localización fue fortuita, como producto de las revisiones de información efectuadas por el autor a otros efectos o de otras circunstancias parecidas. Aunque tales trabajos han sido incluidos en nuestro análisis, debe considerarse que probablemente existan otros trabajos recientes (correspondientes a 1992 y 1993) que no han sido localizados por ahora.

7.7. Estudios versus documentos o informes de investigación

Al principio del proceso de vaciado de información, cada documento (capítulo de libro, artículo, informe de investigación no publicado, etc.) se tomó como un estudio. Por tanto, inicialmente, el número de trabajos recogidos y estudios era coincidente. Sin embargo, al avanzar en la codificación, pudimos ir constatando cómo no siempre se produce esta paridad entre documentos y estudios, dándose dos supuestos claramente distintos a ello y contrarios entre sí: el primero tiene lugar cuando dos o más informes son complementarios entre sí y constituyen, en realidad, un solo estudio. Esta posibilidad ya había sido contemplada desde el principio en la codificación, y, para ella, se habían preparado los campos DOCUCOMP1, DOCUCOMP2 y DOCUCOMP3, para la referenciación de posibles documentos complementarios utilizados para la codificación un determinado estudio. El segundo supuesto acontece cuando algunos informes de investigación refieren dos o más programas distintos, aplicados en contextos diferentes. En este caso, un mismo documento o informe de investigación dará lugar, de manera obvia, a más de un estudio.

Existe un tercer supuesto no tan claro como los anteriores, en el que debe decidirse cuál es el número de estudios distintos acogidos en un informe. Éste se produce cuando en un mismo informe de investigación, y dentro de un mismo programa, son aplicados dos o más tratamientos distintos a otros tantos grupos de sujetos. Imaginemos, por ejemplo, que en un programa realizado en un centro de reforma juvenil, se seleccionaron tres grupos, integrados respectivamente por 15 sujetos cada uno. Con el primer grupo se trabajó mediante un programa de habilidades sociales, aplicándose al segundo grupo este mismo programa más sesiones de terapia no directiva centrada en el cliente. Los resultados obtenidos en uno y otro caso fueron comparados con los del grupo de control, sobre el que no se intervino. En tal supuesto, ¿debe considerarse uno o dos estudios?

Para casos como el descrito, nos decantamos por la opción que nos parecía más clarificadora: computamos como un estudio distinto cada comparación relativa a un tratamiento diferente. Es decir, consideramos que existía un estudio cuando, ya fuera dentro de un mismo programa o en programas distintos, un grupo de sujetos era sometido a alguna suerte de tratamiento

diferente, cuyos efectos eran comparados bien con su medida pre-test o con un grupo de control. Una opción semejante habían adoptado Gottschalk et al. (1987) al considerar para su meta-análisis la existencia de 90 estudios. Entendieron como tales cualesquiera diseños de investigación simples: la comparación entre un tratamiento y un control.

Con esta finalidad, una vez que se nos produjo este problema, redefinimos el campo GRUTRATA, que nos servía para codificación del número de grupos de tratamiento, de la siguiente manera: "Debe entenderse grupos con tratamientos distintos. P.e. Si en un misma aula 20 sujetos reciben una enseñanza y 20 otra, tendremos 2 grupos de tratamiento. Por el contrario, si en dos aulas distintas, todos los sujetos (N=80) reciben el mismo tratamiento, tendremos 1 solo grupo de tratamiento".

Con este acercamiento adoptamos, al igual que hicieron Gottschalk et al. (1987), una perspectiva "interventiva", en la que la prioridad se pone, por encima de todo, sobre los distintos tipos de tratamiento, cuyos efectos diferenciales deseamos estudiar.

7.8. Estudios localizados y seleccionados

Tras el descrito proceso de localización y filtrado de estudios, fueron recolectados inicialmente -en la primera selección- 250 trabajos, correspondientes a los países que se relacionan en la Tabla 7.3. De ellos se seleccionaron al azar 20 informes para la realización del estudio piloto al que nos referiremos más adelante. Finalmente, la aplicación de los criterios establecidos para su inclusión en el meta-análisis -segunda selección- redujo a 57 el número de estudios seleccionados, correspondientes a seis países -Alemania, España, Gran Bretaña, Holanda, Israel y Suecia-, entre los que no se encuentran ni Italia ni Francia⁹.

⁹Inicialmente ya contábamos con pocos estudios franceses e italianos, como puede verse en la Tabla 7.3. Más sorprendente aún fue que, una vez filtrados los estudios que reunían los requisitos necesarios para su inclusión en nuestro meta-análisis, nos quedamos sin trabajo alguno de Francia e Italia. En ambos casos nos resultaba difícil de creer esta ausencia, tratándose de países de un aparente gran desarrollo de las ciencias criminológicas y sociales. En el caso italiano, tras repasar los materiales localizados así como las revistas de su procedencia, encontramos una

explicación plausible que aventuramos aquí: En el pasado, Italia dio a la penología y criminología prohombres de la talla de Cesare Beccaria o Lombroso; aún en la actualidad en Italia existe, sin duda, una creativa reflexión penal-criminológica, que atestiguan nombres como Pavarini, Gaetano de Leo y otros. Sin embargo, en el presente, esta reflexión adquiere un carácter eminentemente teórico y propositivo, y es raras veces concretable -y concretada- en la barahúnda de la Administración italiana. Paralelamente, tanto criminólogos como psicólogos y expertos en delincuencia parecen ser -en términos valorativos generales y, por tanto, asumiendo, sin duda, error- ajenos a intereses empíricos. En revistas y monografías italianas hemos podido encontrar "progetti", "riflessioni", "valorizzazioni", y "esperienze", realizadas en cárceles, en la comunidad, etc., pero, en ningún caso, evaluaciones cuantitativas de programas que hayan sido realmente llevados a la práctica. La impronta empirista con que Lombroso iniciara la criminología científica a finales del siglo pasado parece haberse borrado por completo en Italia.

Más sorprendente nos parecía el caso francés, en donde el aparente desarrollo de la sociología, psicología, trabajo social, etc., haría esperar la existencia de estudios de programas diversos con delincuentes. Bajo esta inquietud realizamos un esfuerzo especial dirigido a localizar posibles trabajos franceses. Para ello, lo primero que hicimos fue, según ya se ha comentado, efectuar una búsqueda informatizada en la base de datos francesa Pascal, que recoge información de múltiples revistas de ciencias sociales y criminología. Esta búsqueda nos aportó alguna nueva referencia, pero ninguna de ellas francesa. Posteriormente, en marzo de 1993 tuve que dar un curso para técnicos de prisiones en el ENAP de París (École Nationale d'Administration Pénitentiaire). Ello me brindó la ocasión de contactar con su director, el Sr. George Camps (por cierto, hijo de padres catalanes, exiliados en la Guerra Civil), con diversos técnicos penitenciarios, y pude consultar la biblioteca del ENAP. Planteé abiertamente el problema -o sea, la no localización de programas franceses- a mis anfitriones, quienes me terminaron de confirmar la cruda realidad existente al respecto. Me aseguraron que en Francia, con carácter general, no se realizaban ni se publicaban estudios sistemáticos sobre programas que hubieran sido llevados a cabo con delincuentes. Sencillamente no se hacía.

Encontraron tres razones para ello: una, que la gestión de la Administración Penitenciaria había estado y seguía estando en manos de hombres de leyes, poco inmersos en la mentalidad científica (paradójicamente, mis interlocutores en esta conversación eran, respectivamente, un físico -el director del ENAP- y un trabajador social, profesor del Centro). Ello hacía que se iniciaran experiencias, reformas, etc., sin que fueran mínimamente controladas y evaluadas de modo sistemático. La segunda razón residía a su entender en el cariz clínico que en Francia, en general, adoptan en su intervención las ciencias sociales. La cultura psicológica francesa, el trabajo social, etc., están muy imbuidos de enfoques dinámicos, fenomenológicos, etc., y son ajenos a los abordajes más empíricos. La intervención con delincuentes, y más específicamente la intervención en las prisiones, no son una excepción en este punto. Finalmente, medio en broma medio en serio, adujeron una razón de "chauvinisme" que puede que no sea menos verdadera: los expertos franceses en materia de justicia, al repasar su propia historia, se sienten muy orgullosos de, al parecer, haber inventado, en la Revolución, la prisión moderna. Asimismo, sienten extraordinaria

Tabla 7.3. Número total de estudios recolectados y seleccionados para el meta-análisis.

PAÍSES	N INFORMES LOCALIZADOS	N ESTUDIOS SELECCIONADOS
Alemania	22	4
Bélgica	7	
España	55	13
Ex-países socialistas	6	
Francia	6	
Gran Bretaña	83	26
Holanda	26	9
Irlanda	1	
Israel	1	1
Italia	11	
Portugal	18	
Suecia	7	4
Suiza	7	
TOTAL	250	57

devoción y respeto por la obra del eminente pensador francés Jacques Foucault sobre la prisión. Pues bien, la contemplación de estas grandezas les tendría tan ensimismados que, como cultura, no les dejaría seguir trabajando para mejorar los abordajes de la delincuencia y las prisiones.

Aunque no se mencionó en la conversación con nuestros colegas franceses, encontramos una razón más que puede explicar la inexistencia de estudios de evaluación sobre el tratamiento de la delincuencia. Francia, en materia cultural y lingüística, vive en buena medida volcada sobre sí misma. Se minusvalora lo ajeno, y más aún si no es hablado en francés, la lengua de la cultura. Es un hecho conocido que el mayor desarrollo de la criminología empírica y ciencias relacionadas (psicología, sociología, metodología, etc.) tienen lugar en lengua inglesa (más del 70% de nuestros trabajos están escritos en inglés). Si una cultura es ajena a esa información, será poco influida por sus contenidos y modos de hacer. Es, a nuestro entender, el caso de Francia y, probablemente, también de Italia.

Nuestros amigos franceses sólo nos dieron una esperanza: la posibilidad de localizar algún trabajo evaluativo, de carácter académico, no publicado. Y se comprometieron con nosotros a realizar algunas indagaciones al respecto.

RESUMEN

En este capítulo nos hemos referido, en primer lugar, a los objetivos de nuestra investigación y a la temporalización de los mismos. Nos proponemos integrar, mediante el meta-análisis, diversos programas europeos de tratamiento de la delincuencia, determinar su efectividad y conocer sus principales componentes. Además, la metodología utilizada nos permitirá la futura integración y análisis acumulativo de nuevos programas.

En segundo término hemos abordado la descripción de los canales básicos existentes para la localización de estudios en una revisión de literatura científica. Entre éstos se encuentran tanto los canales informales -producto del contacto habitual de un investigador con un determinado campo, mediante su propio trabajo, participación en congresos, etc.- como los formales, en que se realiza una búsqueda sistemática de información, ya sea de modo primario -bibliotecas y procedimientos ascendentes- o secundario -bibliografías y servicios de índices y resúmenes-.

Hemos descrito nuestro propio proceso de búsqueda de estudios que utilizó los siguientes procedimientos: (1) una consulta sistemática en tres bases documentales especializadas -Dialog, Pascal y Psychological Abstracts-; (2) una revisión directa de revistas, libros y monografías especializadas en el tratamiento de la delincuencia; (3) y la petición de trabajos a un amplio directorio de centros de investigación y expertos en la materia que nos ocupa.

Se han evidenciado y debatido algunos de los problemas metodológicos con los que nos enfrentamos al seleccionar los estudios que formarán parte del meta-análisis: entre ellos, la heterogeneidad de las intervenciones, de las medidas resultado (reincidencia, psicológicas, etc.), de los diseños y de las metodologías evaluativas.

Finalmente, hemos formulado los criterios seguidos para la recolección y selección de los trabajos que integrarán la muestra analizada que, en síntesis, son: (1) que se trate de un grupo de sujetos bajo control de la justicia; (2) que se haya aplicado alguna estrategia dirigida a la

rehabilitación; (3) que exista un diseño de evaluación que permita la comparabilidad; (4) y que posean algún resultado cuantitativo. En total fueron localizados 250 informes, a partir de los cuales 57 estudios entraron a formar parte del análisis.

8. METODOLOGÍA PROTOTÍPICA DE LOS ESTUDIOS SELECCIONADOS

Fernández-Ballesteros (1987) ha definido la evaluación de programas como la "aplicación de procedimientos de investigación social para el "enjuiciamiento ... sobre el mérito, valor o eficacia de acciones humanas concretas implantadas en la realidad social" (p.2). Un término cercano al de evaluación, que es frecuentemente usado en las ciencias sociales, es el de seguimiento. Se ha denominado seguimiento a la "observación o estudio del porvenir de un sujeto después de un pronóstico o de un tratamiento" (Landsheere, 1985, p.300). Además, deben distinguirse los conceptos de evaluación y medición. La evaluación supone un proceso más amplio y complejo que la medición. Mientras que la medición es una descripción cuantitativa del comportamiento, la evaluación "comprende a la vez la descripción cualitativa y la descripción cuantitativa de los comportamientos y supone, además, unos juicios de valor que afectan a su deseabilidad" (Gronlund, citado por Landsheere, 1985).

En el ámbito correccional que nos ocupa la evaluación puede acometer distintas finalidades; puede servir (Anguera y Redondo, 1991; Redondo, 1993a):

-Para analizar la validez de distintos modelos explicativos de la delincuencia, su relevancia interventiva y su evaluabilidad.

-Para predecir la prominencia delictiva de ciertos sujetos o grupos de sujetos y para predecir la reintegración y/o reincidencia de los ex-condenados.

-Para tomar decisiones acerca del cambio de las instituciones correccionales.

Según Thornton (1987) la evidencia empírica puede emplearse tanto para optimizar el diseño de los sistemas de organización correccional como la selección de los programas que deben aplicarse.

-Para rediseñar programas de intervención, constatando su eficacia o ineficacia.

-Y, finalmente, para llevar a cabo análisis sincrónicos/diacrónicos de los programas que aplicamos. Esto es, para efectuar evaluaciones que en momentos críticos distintos nos permitan conocer la adecuación de funcionamiento de nuestras intervenciones (análisis sincrónicos), o con objeto de evaluar de modo permanente nuestros programas (evaluaciones diacrónicas).

8.1. ¿Cómo evaluar un programa correccional?

En la evaluación de programas correccionales Jones (citado por Thornton, 1987) ha prescrito que debe estar claramente definido:

- 1) Cómo van a ser tratados los sujetos: qué acciones van a ser emprendidas bajo el epígrafe "tratamiento" o "intervención".
- 2) Qué efectos inmediatos suponemos que producirá el tratamiento aplicado.
- 3) Cómo pensamos que estos efectos inmediatos van a influir la futura delincuencia de los sujetos tratados.

Por ello, la evaluación de los programas de tratamiento de la delincuencia no deberá interesarse sólo en conocer el resultado final que éstos producen sobre la reincidencia, sino también en los procesos teóricos que pueden acompañar a la producción en tales efectos (Anguera y Redondo, 1991)

Thornton (1987) ha dividido en tres grupos las variables empleadas en el diseño y la evaluación de programas correccionales: variables input, que "son aquellas que... pueden ser dispuestas con antelación al funcionamiento del sistema" (p.478) o programa (por ejemplo, las características de los

participantes y del personal, las condiciones del edificio, su localización geográfica, la organización formal del sistema -normas y pautas de funcionamiento-, etc.); variables de proceso, que intervienen a lo largo del funcionamiento del programa (por ejemplo, las actitudes y conducta de los sujetos tratados y del personal, y sus interacciones); y variables output, referidas al modo social y psicológico de funcionamiento de los delincuentes, ya sea dentro de las propias instituciones correccionales ya sea después de su liberación.

Entre las tareas de los profesionales encargados de este ámbito evaluativo se encontrarían, según Milan y Evans (1987), las siguientes:

- 1) Diseñar y seleccionar las medidas evaluativas que serán empleadas.
- 2) Evaluar la efectividad de los programas. Para ello, sería necesario, en primer lugar, el uso de evaluaciones formativas, en las que se revisaría la información disponible sobre el programa y se realizarían sugerencias relevantes sobre el mismo, que permitieran dar al programa su forma final; pero también, sería necesario el empleo de la evaluación sumativa, con mediciones antes y después de la intervención, para valorar si un programa, considerado globalmente, ha resultado efectivo.
- 3) Otros cometidos importantes serían los de entrenar al personal y divulgar los conocimientos adquiridos mediante la evaluación de los programas a otros ámbitos de las instituciones correccionales, los sistemas de justicia, o la literatura especializada sobre el tema. "En gran medida, es responsabilidad del psicólogo correccional -escriben Milan y Evans (1987)- describir los resultados exitosos, publicar informes sobre los programas desarrollados (tanto si tuvieron éxito como si fracasaron), y participar en la confección de manuales de entrenamiento y/o procedimientos diseñados para divulgar el empleo de programas útiles entre otros miembros de la institución" (pp. 579-580). Esto es, se hace necesario que las características básicas de los procedimientos, programas y contextos específicos de la evaluación e intervención con delincuentes sean identificados de un modo suficientemente completo y objetivo (Morris y Braukmann, 1987).

En un reciente trabajo, Yonai (1992) ha sistematizado una guía ideal que

debería ser empleada para la correcta evaluación de programas de educación correccional. Esta autora divide el proceso de evaluación en los siguientes tres grandes apartados:

1. El encuadre de la evaluación. Esta primera etapa tiene un carácter eminentemente reflexivo. En ella el evaluador debería delimitar, en primer lugar, el propósito de la evaluación, contestando a preguntas críticas como qué va a ser evaluado, si se trata de todo un programa o sólo de una parte del mismo, y por qué va a ser evaluado. Además, debe concretarse si se trata de una evaluación formativa, esto es, interactiva con el propio diseño del programa, o sumativa, o de mero resultado. Finalmente, el evaluador debería conocer en la medida de lo posible cómo van a ser utilizados los resultados de la evaluación.

En este apartado los evaluadores deberían también reflexionar sobre el contexto en el que se enmarca la evaluación, describiendo la estructura organizacional en la que se inscribe, y las correspondientes consideraciones políticas, sociales y económicas de la misma.

Finalmente, el evaluador debe reconocer a los destinatarios de la evaluación, cuyo abanico puede ser muy amplio, yendo desde los propios afectados por el programa o sus familias, los responsables del mismo, los empleados para su realización, quienes lo financian, hasta destinatarios más lejanos como legisladores o contribuyentes. Conocer quiénes son los destinatarios de una evaluación tiene gran importancia para delimitar o matizar cuáles son, en definitiva, las preguntas que deben responderse mediante la evaluación, cuál será la metodología más adecuada, y cuáles los mecanismos de información que se utilizarán para la devolución de resultados.

2. El diseño de la evaluación. En este ámbito debe ser explorado un amplio rango de preguntas que podrían ser contestadas mediante la evaluación. Ello nos ayudará a clarificar los objetivos concretos y las variables que serán medidas.

Tal vez el apartado más crítico resulte aquí la delimitación del método, que dará cuenta de cómo, en suma, van a ser respondidas las preguntas de evaluación. ¿Van a utilizarse, para ello, encuestas, escalas de conducta,

tests, observación de grupos, entrevistas, análisis de documentos, observación participante, etc.? ¿Va a llevarse a cabo un proceso de triangulación, para obtener evidencia sobre la misma pregunta desde distintas fuentes de información? Y, posteriormente, ¿cómo van a ser aplicados estos instrumentos, en qué momento y por quién?

A la vez, será necesaria también una reflexión sobre las fuentes de datos posibles. Esto es, quiénes poseen información sobre nuestro objeto de evaluación: ¿Los sujetos participantes, el "staff", los educadores, las familias, los trabajadores sociales, los agentes de probation, los jueces...? Diferentes fuentes de información pueden tener distintas perspectivas de la misma actividad, y un suministro múltiple puede ayudarnos a confeccionar una imagen más completa del programa.

Conectadas con los métodos se encuentran las técnicas de análisis de los datos, en las que deberemos decidir si se utilizarán técnicas cualitativas, cuantitativas o ambas. Ambas tecnologías -cualitativas y cuantitativas- permiten incorporar análisis de costes y de tiempo.

Finalmente, deben adoptarse decisiones, en el diseño de la evaluación, sobre el sistema de informe que se empleará para la comunicación de la evaluación. Su nivel de detalle, comprensividad y formato vendrán determinados por la audiencia o audiencias de destino.

3. La gestión de la evaluación. En el ámbito que nos ocupa, una de las primeras preocupaciones debería dirigirse a definir el personal que llevará a cabo la evaluación. Convendrá decidir si el personal propio, que pretenda implicarse en el proceso evaluativo, posee los conocimientos y habilidades necesarios para dirigir o efectuar la evaluación, o si, por el contrario, se requieren consultores externos.

No menos importante es la formalización del presupuesto de los costes de la evaluación, en los que deberán incluirse, para evitar sorpresas posteriores, todos los conceptos necesarios (viajes, visitas, transcripción de informes, teléfono, costes de los análisis de datos, etc.).

Finalmente, debe planificarse de modo preciso la temporalización, en la que

deben preverse todas las actividades necesarias, desde el proyecto de evaluación, el diseño de los instrumentos, el pilotaje, las cadencias de la intervención, los análisis, etc. Una vez diseñada, la temporalización debería ser contrastada y ajustada con los propios destinatarios de la evaluación.

Es evidente que las pautas formuladas por Yonai (1992) para la evaluación de programas de educación correccional constituyen un proceso ideal al que debería, tal vez, aspirarse. Sin embargo, también es cierto, según iremos viendo, que son pocos los informes de evaluación correccional en los que puedan encontrarse los pasos e informaciones descritos por Yonai (1992). Desgraciadamente, muchos de ellos se encuentran, con frecuencia, ausentes.

8.2. Ejemplificación de programas incluidos en el meta-análisis

Para ejemplificar el proceso evaluativo de los programas que nos hemos encontrado, hemos seleccionado algunos estudios incluidos en este meta-análisis de los que realizaremos una descripción narrativa más detallada. Más concretamente, hemos elegido tres estudios, cada uno de los cuales representa a una de las técnicas interventivas más frecuentes entre los trabajos recogidos: el de Hollin et al. (1986) es un programa típico de terapia cognitivo-conductual, tal vez el modelo de mayor utilización actual en el tratamiento de delincuentes; el de Slot (1984, septiembre) representaría a los programas de contingencias que poseen un carácter globalizador de la vida de los sujetos; finalmente, el informe de Cooke (1989) nos servirá para ejemplificar otra estrategia frecuente en el tratamiento de la delincuencia: las comunidades terapéuticas.

Social Skills Training with Young Offenders in a Borstal: An Evaluative Study (Hollin et al., 1986).

Este estudio se propuso comparar la efectividad de un programa tradicional de habilidades sociales con uno complementario de auto-instrucción, ambos sobre la base del modelo cognitivo-conductual. Se trabajó con 10 sujetos jóvenes (cuya edad media era 19 años), encarcelados por robos y delitos violentos, que fueron asignados aleatoriamente a dos grupos de tratamiento,

correspondientes con las dos técnicas aplicadas. Además se constituyeron tres grupos de control, integrados por un total de 25 individuos. Para seleccionar a los sujetos se efectuó una entrevista y se aplicó una escala de problemas de conducta, lo que permitió catalogarlos como apropiados/inapropiados para el programa. Los sujetos controles fueron en parte reclutados mediante apareamiento en función de las variables edad, inteligencia y tipo de delito, para controlar la equivalencia grupal.

La intervención se llevó a cabo en un borstal o prisión juvenil de Feltham, Gran Bretaña. Las principales actividades interventivas fueron las siguientes: para el primer grupo de tratamiento se trabajó en modificación de conducta cognitiva, combinando para ello (1) la enseñanza de habilidades sociales manifiestas y (2) autoinstrucción encubierta; el segundo grupo recibió sólo (1) entrenamiento en habilidades sociales manifiestas. En ambos casos la intervención se realizó en diez sesiones semanales de hora y media. Por tanto, la intervención tuvo un intensidad de conjunto de quince horas. Por su parte, los grupos controles adoptaron dos modalidades experimentales: dos de ellos no recibieron intervención alguna, y el otro recibió también diez sesiones semanales de hora y media con discusión grupal no estructurada sobre temas de interés juvenil. En todos los casos las intervenciones fueron llevadas a cabo por un psicólogo y un oficial de prisiones, entrenados al efecto. El programa duró en total diez semanas a las que se añadieron diez semanas más de seguimiento.

Para evaluar la efectividad de los tratamientos se utilizaron cinco medidas, que se aplicaron a todos los grupos, antes y después del tratamiento: (1) se midieron variables de personalidad -extroversión, neuroticismo, psicoticismo- mediante el Eysenck Personality Questionnaire; (2) se empleó la escala Jesness Behavior Checklist (Jesness, 1988), aplicada por funcionarios entrenados, para medir variables como consideración de los otros, efectiva comunicación, amistad, etc.; (3) se tomaron en cuenta las faltas disciplinarias de los sujetos, a través de los informes disciplinarios emitidos por el personal; (4) se evaluaron las respuestas verbales y soluciones de los sujetos -que fueron gravadas en casete y posteriormente categorizadas- a veinte situaciones problemáticas que les fueron presentadas; (5) finalmente, se grabó en vídeo su comportamiento en una situación de "role-play" con una chica desconocida, durante tres minutos.

Estas grabaciones fueron posteriormente analizadas y categorizadas en conductas como el contacto ocular, las manifestaciones de ansiedad, etc.

Como principales resultados se obtuvieron los siguientes: (1) el tratamiento no afectó a las medidas de personalidad, lo que en absoluto sorprendió a los autores, que argumentan la falta de relación, informada por la literatura científica, entre concretas habilidades sociales y medidas de personalidad; (2) por su lado, la escala conductual aplicada (Jesness Behavior Checklist) resultó poco útil para medir los resultados de los distintos grupos, no permitiendo la discriminación de efectos diferenciales en ellos; (3) todos los grupos entrenados -entrenamiento cognitivo, habilidades sociales, y discusión grupal- mejoraron su conducta en prisión, disminuyendo sus informes disciplinarios durante el período de un mes de seguimiento; (4) los datos de autoinforme, grabados en casete, que tuvieron una buena acogida entre los sujetos, presentaron, sin embargo, grandes dificultades para su categorización posterior -dificultades para decidir qué era, en muchos casos, una buena o mala solución a un problema-, por lo que fueron, a la postre, poco útiles; (5) las grabaciones en vídeo constituyeron el instrumento más válido, permitiendo comprobar la efectividad del tratamiento: las técnicas conductuales aplicadas produjeron sustanciales mejoras en los índices conductuales establecidos, en la situación de interacción de los sujetos con una desconocida.

Pese a estas mejoras conductuales en situación de "role-play", los autores concluyeron la dificultad existente para la generalización de tales habilidades desde el contexto de entrenamiento a la vida diaria de la propia prisión. Ello lo atribuyeron a la diferencia estimular existente entre una y otra situación. Por otro lado, no fue directamente investigada la posible mejora de la autorregulación encubierta, que había sido entrenada en uno de los grupos. Esto habría requerido un diseño no grupal, que no fue utilizado. Los autores acaban proponiendo una futura estrategia interventiva más radical, que en su opinión debería no sólo poner el énfasis en el entrenamiento individual a los encarcelados sino, más aún, en la "reestructuración del ambiente carcelario para facilitar el cambio y desarrollo personal" (Hollin et al., 1986, p. 297). Sin embargo, ven difícil que este cambio pueda llevarse a cabo en las prisiones inglesas, por lo que sugieren que, al menos, deberían minimizarse los constreñimientos ambientales existentes, programando el entrenamiento

en habilidades sociales más cerca del cumplimiento de la condena, de modo que pueda tener una mayor utilidad para la integración social de los sujetos.

The Teaching Family Model in the Netherlands: first results of a community based program for problem youths (Slot, 1984, septiembre).

Este programa consistió en una aplicación del modelo de padres enseñantes, de la Universidad de Kansas (Wolf et al., 1972; Phillips et al., 1980; Braukmann y Wolf, 1987), en dos centros holandeses para jóvenes delincuentes. El programa de padres enseñantes se basa en el concepto de conducta deficitaria: la falta de habilidades sociales es considerada un importante factor en el desarrollo de una conducta problemática como la delincuencia.

Se trabajó con un grupo experimental de 24 sujetos jóvenes (promedio de edad de 16.8 años), que fueron comparados con un grupo control equivalente de 72 individuos. Los datos ofrecidos en la evaluación no correspondieron siempre a todos los sujetos tratados, sino que en algunas variables fueron evaluados los 24 sujetos, mientras que en otras sólo lo fueron 12 ó 6 de ellos.

El programa se dirige a un pequeños grupos de 6 jóvenes problemáticos que viven en un casa con dos padres enseñantes -hombre y mujer-, que suelen ser técnicos sociales, estudiantes universitarios, etc., parejas reales o no, que llevan a cabo esta actividad durante un tiempo como trabajo remunerado. El centro se ubica en un área normal de la ciudad, y los jóvenes asisten a las escuelas y a otras actividades formativas y asociativas del propio barrio. El tratamiento se estructuró en una serie de niveles secuenciados:

(1) Fase primera, al ingreso de los sujetos en el programa: se pone el énfasis en la enseñanza de habilidades útiles para el funcionamiento social tanto dentro del propio centro como fuera del mismo: asistencia escolar, trabajo, etc.; para ello los educadores utilizan técnicas de instrucción de conducta, modelado, y extinción, que facilitan los comportamientos apropiados de los jóvenes y les dan "feedback" inmediato sobre sus ejecuciones. Asimismo, se emplea un programa de economía de fichas, con las mismas finalidades. Durante esta fase inicial, se entregan puntos por la conducta apropiada y se

retiran por la inapropiada. En todos los casos se ganan más puntos que se pierden, y los sujetos pueden canjear sus ganancias por privilegios una vez por día.

(2) Fase segunda, de semi-autogobierno: se presta una mayor atención a los déficits individuales y se promueven dos mecanismos de participación de los jóvenes en la planificación y funcionamiento del programa: por el sistema de administrador un joven se responsabiliza durante una semana del funcionamiento y control de las tareas domésticas; por el sistema de conferencia familiar se permite que los jóvenes tomen decisiones grupales para fijar normas, sanciones, etc.

(3) Fase final, de autogobierno: se prioriza la enseñanza y práctica de habilidades para la independencia exterior.

El programa tuvo una duración media de 12,5 meses por sujeto, y se efectuó un medida de seguimiento a los 6 meses de que los sujetos hubieran abandonado el programa.

Para la evaluación se emplearon los siguientes indicadores e instrumentos: (1) se midieron los problemas psicosociales de los jóvenes, utilizando para ello una serie de escalas de relaciones familiares, aspiraciones escolares y laborales, frecuencia de problemas de conducta, frecuencia de delitos, y problemas en casa o en sus residencias posteriores; (2) los trabajadores sociales que supervisaban cada caso, evaluaban periódicamente -en cada visita- los progresos de sus pupilos, juntamente con los padres enseñantes y los propios jóvenes, en un cuestionario con once preguntas relativas al progreso de cada joven; (3) mediante entrevistas se evaluaban el clima social de la casa, el tiempo dedicado por los padres a discutir el futuro, la utilidad del programa, la claridad de las instrucciones dadas a los jóvenes, la imparcialidad de los padres enseñantes, el nivel de atención prestado a los chicos, etc.

Se obtuvieron resultados pre-post para 6 sujetos del grupo de tratamiento, de los cuales 5 mostraron claramente una mejora y 1 un deterioro. Los oficiales y trabajadores sociales de caso valoraron muy positivamente el desarrollo del programa, constatando mejoras importantes en la conducta de

los chicos y una importante reducción de su violencia, tanto durante el programa como posteriormente. Por su parte, las entrevistas a los jóvenes permitieron obtener conclusiones positivas sobre las valoraciones que éstos hacían del clima social y otras dimensiones consideradas.

Todo ello conduce a los autores a concluir que el programa de padres enseñantes se muestra una alternativa atractiva y eficaz para el tratamiento de jóvenes delincuentes. Además, este programa resulta sumamente interesante en términos de costo-beneficio: es sustancialmente menos costoso que los internamientos tradicionales de larga duración.

Containing Violent Prisoners. An Analysis of the Barlinnie Special Unit (Cooke, 1989).

A raíz de la abolición de la pena de muerte en Escocia se produjo preocupación entre el personal penitenciario en relación con aquellos presos condenados a cadena perpetua, algunos de los cuales habían llevado a cabo agresiones graves a funcionarios.

Por ello en 1973 se creó en la prisión de Barlinnie (Glasgow) la Special Unit, para contener a este tipo de presos. Se basaba en los principios de comunidad terapéutica: esto es, las relaciones personal-presos deberían cambiar para parecerse a las de enfermeros-pacientes; los presos habrían de tener un mayor peso en la toma de decisiones; deberían poder ser responsables de su propia conducta y de la de sus compañeros, y habría que enseñarles a verbalizar sus sentimientos agresivos. Este régimen no fue especificado desde el principio en detalle y su funcionamiento evolucionó a lo largo del tiempo, como resultado del entrenamiento del personal y de las presiones y demandas de la comunidad.

La estancia en la Unidad Especial comportaba una serie de privilegios: posibilidad de recibir frecuentes visitas, de prepararse sus propios alimentos, de elegir trabajar o no, de decorar su celda y equiparla con televisión; y una serie de desventajas: reducido espacio de la unidad (menor que el de una prisión), dificultad para la privacidad, y presión de las asambleas de comunidad (integradas por presos y personal). Éstas podían

decidir el establecimiento de ciertas consecuencias para el comportamiento de los internos: si no cooperaban con el régimen podían ser trasladados a otra prisión más dura; si quebrantaban las 'normas' comunitarias la comunidad podría decidir 'sanciones' como la pérdida temporal del privilegio de recibir a las visitas en la propia celda o de salir fuera de la unidad.

Los presos eran seleccionados para la Unidad por haber sido violentos, destructivos o subversivos en otras prisiones, una vez que se les había realizado un estudio psicológico individual y se habían descartado otras posibilidades de intervención con ellos, la existencia de problemas psiquiátricos profundos, o enemistades con otros presos de la unidad, y se había constatado su motivación para un cambio positivo.

Por la Unidad pasaron 25 individuos entre 1973-1986. La estancia media por preso fue de 41 meses (rango 4-90 meses). El 64% de ellos habían sido condenados por asesinato y un 76% por robos; el 60% tenían cadena perpetua, y sólo 3 de ellos debían cumplir menos de 10 años cuando ingresaron en la Unidad. El 68% habían recibido condenas adicionales por delitos dentro de la propia prisión. El 80% estaban clasificados -a su ingreso en la Unidad- en categoría A o de máxima seguridad.

Como metodología evaluativa se utilizó el estudio sistemático de casos mediante informes de la prisión, delictivos, psiquiátricos y psicológicos, y observaciones independientes efectuadas por el autor del estudio y un colega clínico.

El autor ofrece resultados tanto sobre las respuestas de los sujetos dentro del propio contexto carcelario como en la comunidad, para aquéllos que fueron liberados. Dentro del ambiente carcelario se evaluaron tres indicadores: (1) número de individuos agredidos físicamente, en comparación al cálculo que se hizo de la proporción esperada para cada preso según su propia historia carcelaria; sólo ocurrieron 2 agresiones, frente a un volumen esperado de 105; (2) número de incidentes graves ocurridos (cantidad de intentos de evasión, huelgas de hambre, "accidentes violentos", manifestaciones, barricadas, fuegos provocados, campañas de suciedad, autolesiones y fugas), frente a los esperados; sucedieron 8 incidentes graves, mientras que, según la estimación realizada, hubieran sido

esperados 49; (3) número total de incidentes (suma de (1) y (2))

Aunque este programa no pretendía, por definición, la reforma, sino sólo la contención de los encarcelados, 12 presos obtuvieron la libertad. Se calculó la proporción de delitos esperada (según la ecuación de predicción de Nuttall, que permite estimar una predicción de reincidencia a partir de 16 factores asociados a ella como tipos de delitos previamente cometido, edad del primer delito, duración de la condena, etc.). Según ello, cabía esperar que 8,31 sujetos reincidieran en el plazo de dos años; sin embargo, sólo 4 lo hicieron.

Finalmente, para evaluar el clima generado en la Unidad Especial se aplicó la Escala de Ambientes Correccionales de Moos (1975). Los internos consideraron el régimen habitual de las otras prisiones como más alto en "control del personal" y "orden y organización", y el de la Unidad Especial alto en su "orientación hacia los problemas personales", "orientación práctica", "autonomía", "apoyo", "expresividad", y "compromiso".

Según el autor, los factores situacionales de la Unidad Especial que parecieron tener una mayor influencia en el cambio de comportamiento de los sujetos fueron los siguientes: el mayor nivel de autonomía; se proveyó una saludable válvula para la expresión de sentimientos agresivos; las "reuniones de comunidad (presos/personal)" mostraron ser un medio poderoso de imponer el control social mediante la presión a conformarse con las normas del grupo; la oportunidad para romper las barreras entre el personal y los internos; el acceso a regulares y frecuentes visitas de la familia y amigos; y, finalmente, otros privilegios existentes como la elección por parte de los encarcelados de actividades diarias, la posibilidad de confeccionar su propia comida, y el acceso a una televisión propia.

Interpreta el autor que los anteriores factores ambientales pudieron generar algunos cambios "intrínsecos" en los sujetos, por la mediación de dos procesos: por un lado, algunos internos pasaron varios años en la Unidad y, por este influjo contextual, pudieron aprender a relacionarse con la gente de un modo no agresivo; además, algunos pudieron desarrollar nuevas redes sociales no delictivas, lo que habría facilitado su inserción a la salida.

8.3. Principales técnicas utilizadas en los programas analizados

En el epígrafe anterior hemos presentado brevemente tres programas típicos que ejemplifican las técnicas interventivas con delincuentes más frecuentes en Europa. Sin embargo, estos tres programas, y las técnicas por ellos representadas no dan cuenta de todos los programas que hemos seleccionado y analizado. Veamos a continuación un detalle de las características más sobresalientes de las diversas técnicas de tratamiento que hemos podido categorizar.

1-Terapia de conducta: en las terapias de conducta el modelo básico utilizado es el condicionamiento operante, cuya ley empírica del efecto establece que las consecuencias de una respuesta son un determinante de la probabilidad futura de esa respuesta. Una aplicación muy frecuente de la terapia de conducta con encarcelados viene constituida por las economías de fichas, cuyos elementos básicos son (Redondo, 1993a): (1) se establecen una serie de objetivos conductuales (higiene, cursos, programas laborales, reducción del consumo de drogas, de las agresiones y violencia, etc.); (2) se determinan una serie de reforzadores o situaciones gratificantes (visitas familiares, dinero, reducciones de condena...); (3) se pondera la relación de valor entre conductas y refuerzos; (4) y, finalmente, se estructura un sistema de fichas o puntos, que son entregados a los sujetos por sus logros conductuales, y que pueden intercambiar por los refuerzos.

Como operadores de estos programas suelen actuar tanto terapeutas expertos como para-profesionales previamente entrenados en técnicas conductuales. En la categoría de las terapias de conducta se enmarcaron los programas de Redondo (1983; 1984), Daniel (1987), Daniel y Dodd (1989), Martín Aixelà (1989), Kravetz et al. (1990), y Sastriques (1993).

2-Programa ambiental de contingencias: se asienta sobre los modelos teóricos de condicionamiento operante y del aprendizaje social o cognitivo. Un ejemplo típico de estos programas lo constituye el sistema de fases progresivas, cuyos componentes básicos son los siguientes (Redondo et al., 1990): (1) se establecen una serie de objetivos conductuales, que suelen abarcar toda la vida diaria de los sujetos; (2) se estructuran una serie de unidades de vida o fases, distintas en exigencia de conducta y disponibilidad de reforzadores;

(3) los sujetos son periódicamente asignados a unas fases u otras en función de sus logros conductuales. Para la aplicación de este tipo de programas se requiere la práctica implicación de todo el personal de una institución, liderados por un reducido grupo de expertos, que se encargarían del diseño, la supervisión y la evaluación de los mismos. Se encuadraron como programas ambientales de contingencias los de Slot (1983; 1984, mayo), Brown (1985), Slot y Heiner (1986, mayo), Day (1988), Tomás Bardisa (1990), y Redondo et al. (1990).

3-Terapia cognitivo-conductual: se fundamenta en el modelo cognitivo conductual o de aprendizaje cognitivo. Tal vez el programa cognitivo-conductual más completo lo constituya el programa de competencia psicosocial (Ross, Fabiano y Garrido, 1990), cuyos elementos fundamentales serían los siguientes: (1) se evalúan los déficits cognitivos y de habilidades de interacción de los sujetos; (2) se trabaja con grupos reducidos (8-12 sujetos), en sesiones de 1-2 horas, 1-5 veces por semana; (3) se aplican las siguientes técnicas estructuradas: Solución de problemas cuyo objetivo es enseñar a los sujetos a reconocer situaciones problemáticas y generar soluciones a las mismas; Entrenamiento en habilidades sociales útiles para la interacción más exitosa de los individuos con su entorno social (para ello se emplean técnicas de modelado, role-play, "feedback" y práctica estructurada); Control emocional de cólera...etc., aprendiendo a anticipar situaciones y utilizar ciertas habilidades cognitivas para evitarlas; Razonamiento crítico, mediante el que se enseña a los sujetos a pensar reflexiva y críticamente sobre su propia conducta y la de los otros; Desarrollo de valores, técnica en la cual, mediante el trabajo sobre "dilemas morales", se enseña a los individuos a tomar una perspectiva social, poniéndose en el papel de los otros; Habilidades de negociación, en donde se enseña a negociar como estrategia alternativa a la confrontación; y Pensamiento creativo, programa en el que se procura desarrollar el "pensamiento lateral" o alternativo, frente a las más habituales soluciones violentas con que muchos delincuentes suelen abordar sus problemas. Son operadores corrientes de los programas cognitivo-conductuales terapeutas expertos o educadores y para-profesionales entrenados en estas técnicas.

Dentro de las terapias cognitivo-conductuales se inscriben el mayor número de programas que hemos analizado: McGurk y Newell, 1981; Hollin y

Courtney, 1983; Slot y Bartels, 1983; Weaver y Fox, 1984; Hollin et al., 1986; McDougall et al., 1987; Beljaars y Berger, 1987; Cullen, 1987; González y Gutiérrez, 1989, mayo; Garrido et al., 1989; Kury, 1989; McMurrin, 1990; Garrido y Sanchís, 1990; Hopkins, 1991; Fisher, 1991; y Bayón y Compadre, 1991.

4-Comunidad terapéutica: pretende abarcar toda la vida diaria de los sujetos; las relaciones encarcelados-personal son definidas como semejantes a las que deberían existir entre pacientes-enfermeros en un contexto terapéutico. Sus principales componentes son: (1) se eliminan los controles rígidos, o los sistemas de sanciones más habituales; (2) el control de la conducta de los sujetos se hace descansar en el propio grupo: en la comunidad, formada por encarcelados y personal; (3) tienen lugar asambleas periódicas de comunidad, para debatir los problemas existentes en la institución. Como operadores de una comunidad terapéutica suele mencionarse a todo el personal de la institución. Corresponden a esta modalidad interventiva los trabajos de Petterson et al. (1986), Robertson y Gunn (1987), Bishop et al. (1987), Cooke (1989), Sánchez y Polo (1990), Berggren y Svärd (1990, y Gustavsson (1991).

5-Terapia psicológica/psiquiátrica no conductual: incluimos en este apartado un conjunto heterogéneo de terapias fundamentadas en los modelos psicodinámico, médico, de terapia no directiva, y otros (Dünkel, 1982; Bayón 1985; Day, 1988; Rosner, 1988; Members of the Demonstration Unit, 1986; Kury, 1989; Belfrage, 1991; Cook et al. 1991; y Cooke, 1991). Como elementos comunes a todos ellos aparecen los siguientes: (1) se efectúa un diagnóstico de la problemática psicológica de los individuos; (2) la esencia de la intervención consiste en sesiones individuales o de grupo, durante períodos prolongados, dirigidos a esclarecer las eventuales problemáticas personales subyacentes en los sujetos; (3) finalmente, se valora la posible recuperación de los "delincuentes-pacientes". Sus operadores deben ser terapeutas expertos.

6-Endurecimiento regimental: esta perspectiva se basa en el modelo doctrinal clásico de la prevención especial, según el cual la sanción penal produciría por sí misma efectos reductores de la conducta delictiva futura. Sus propuestas básicas son las siguientes: (1) se establece una estricta

disciplina y supervisión de los sujetos, que afecta a toda su vida diaria; (2) se planifican actividades obligatorias: trabajo, gimnasia...etc.; (3) se aplica un sistema inflexible de sanciones. Como operadores de este sistema aparecería todo el personal de la institución. Por fortuna, hemos encontrado esta orientación en sólo uno de los trabajos revisados (Thornton, 1987).

7-Educación: aparecieron varios programas consistentes en cursos, tratamiento escolar, entrega de materiales para lectura, etc. En ellos predominaba la instrucción o enseñanza teórica de contenidos o de habilidades por encima de la puesta en práctica de los mismos. Se inscribieron en esta categoría los programas de Bovens (1987), Moreno y Battestini (1989), van Dalen (1989), McMurrin y Boyle (1990), y Singer (1991).

8-"Diversion" o derivación: programas en los que se utilizaba, como alternativa al internamiento, la libertad bajo palabra, la mediación, la reparación, la supervisión, la asistencia social, etc., generalmente en la comunidad. Tenían connotaciones de "diversion" los programas de Scholte y Smith (1987), Collins y Tate (1988), Kruissink (1990), Boother (1991), Legaz et al. (1990), y López Pina et al. (1992).

8.4. Principales problemas metodológicos de la investigación evaluativa de programas de tratamiento de la delincuencia

Thornton (1987) ha sugerido algunos de los problemas metodológicos frecuentes de la evaluación de los programas de tratamiento de la delincuencia:

1- El primero es un problema de heterogeneidad de las medidas resultado: para evaluar la reincidencia se han utilizado un gran número de indicadores diferentes: el fracaso del período de libertad condicional, nuevos arrestos, nuevas condenas, nuevos encarcelamientos, etc. Esta heterogeneidad ha dificultado la comparación válida de resultados de distintas investigaciones. Más adelante nos referiremos detenidamente al análisis de la variable reincidencia .

2- Heterogeneidad en el Tamaño de las muestras. En general existe una gran dificultad para recoger información sobre grandes muestras de delincuentes que resulten estadísticamente relevantes.

3- Dificultades de aleatorización de los grupos, lo que complica la comparabilidad de las muestras.

Para paliar este problema se han empleado técnicas alternativas a la aleatorización tales como el emparejamiento de casos entre regímenes o programas distintos a partir de tablas de predicción con variables conocidas como el número de condenas previas, la edad del primer arresto, el historial laboral o ciertas características de personalidad (Rutter y Giller, 1988).

Asimismo, Sinclair y Clarke (citados por Thornton, 1987) han propuesto la utilización del que ellos denominan diseño trans-institucional ("cross-institutional design"), que consideraría ciertas unidades de análisis tanto de las propias instituciones correccionales como de los sujetos estudiados. Para las variables teóricamente relevantes de las instituciones correccionales evaluadas así como de los delincuentes sobre los que se interviene se hallarían parámetros de reincidencia y otros, lo que permitiría establecer más apropiados procesos de correlación. Es decir, se trataría no sólo de comparar individuos entre sí, sino también instituciones correccionales diversas, programas distintos, etc., a la hora de evaluar la reincidencia.

4- Importante mortalidad experimental. A modo de ejemplo, en uno de los programas incluidos en nuestra investigación (el de Garrido et al., 1989) los grupos experimental y control sufrieron una merma de más del 50% -sobre un "n" total de 56 sujetos-, durante el corto período de tan sólo dos meses que duró la aplicación de un programa cognitivo.

Esta importante mortalidad experimental crea un grave dilema a los evaluadores correccionales, que se debaten entre la conveniencia de reflejar o no en sus resultados aquellos casos que no completaron el tratamiento, con el riesgo de sesgo que tal exclusión podría provocar en la validez del proceso evaluativo.

5- Denominación de los programas "versus" concretas acciones emprendidas. Frecuentemente, en la evaluación tradicional de programas se ha considerado ingenuamente que las meras denominaciones de los mismos (terapia de grupos, psicoterapia, economía de fichas, etc.) aprehendían y reflejaban la esencia de las técnicas aludidas en ellos. Sin embargo, pese a tener una misma denominación, programas aplicados en lugares distintos por diferentes operadores y con duraciones dispares, pueden estar consistiendo en cosas realmente muy distintas.

6- Deterioro del programa. Los programas correccionales van sufriendo, en general, un progresivo deterioro hacia formas previas de funcionamiento de las instituciones para delincuentes. Es decir, poco a poco, los nuevos modelos interventivos se van "degradando", y la cultura institucional promueve un retorno al funcionamiento anterior. Es lo que se ha denominado la erosión de la integridad del programa. Hasta el presente, la praxis ha sugerido que los programas correccionales sólo puedan ser consistentemente aplicados y evaluados durante un máximo de tres o cuatro años.

7- Finalmente, el último gran foco de dificultades se refiere a la naturaleza multidimensional de los sistemas y programas correccionales. Como ya hemos señalado con anterioridad, las organizaciones correccionales son ámbitos complejos integrados por multiplicidad de subsistemas normativos, de personal, arquitectónicos, población, etc., que pueden hacerles diferir en diversos factores y características, que influirán, sin duda, diferencialmente sobre el desarrollo y los resultados de sus propios programas.

RESUMEN

La evaluación correccional puede acometer distintas finalidades, entre las que tiene un papel destacado el descubrimiento de programas más efectivos para el tratamiento de la delincuencia.

El objetivo fundamental de este capítulo ha sido la ejemplificación del tipo de programas que son objeto de nuestro análisis. Para ello seleccionamos tres estudios "típicos", que hemos descrito brevemente. Su descripción nos ha

servido para contrastar la realidad de los informes habitualmente presentados por los autores de las evaluaciones correccionales con el "desideratum" ideal al que debería atenderse al evaluar y describir un programa. Los informes suelen ser menos sistemáticos y explícitos de lo que sería conveniente para su más adecuada comprensión y estudio. Ello nos ha conducido, finalmente, a poner de relieve las principales carencias metodológicas que suelen presentar las evaluaciones de programas de tratamiento de la delincuencia, que nosotros ahora debemos afrontar en nuestro análisis.

9. LA PLANTILLA DE VACIADO Y CODIFICACIÓN DE LA INFORMACIÓN

9.1. El proceso codificador

A la hora de acometer el proceso codificador de los estudios de una revisión, Glass et al. (citados por Gómez, 1987) han dividido las diversas características de los trabajos meta-analíticos en dos tipos: características sustantivas, que serían las propias de la cuestión analizada (ámbito clínico, programas educativos, delincuencia, etc.) y características metodológicas, o relativas al procedimiento investigador, más generales y comunes a diversos meta-análisis (muestra, número de grupos, técnicas estadísticas empleadas, etc.). La codificación de unas y otras es imprescindible con el objetivo de clarificar si los resultados de los estudios dependen prioritariamente de sus aspectos sustantivos o, si por el contrario, adquieren un peso prominente sus especificidades metodológicas (Gómez, 1987).

Ello debe conducirnos, pues, al establecimiento de un sistema de vaciado y codificación de la información contenida en los estudios que analizamos, en función de estos dos grandes ámbitos de informaciones, sustantivas y metodológicas.

Nuestras categorías de codificación deben ser sometida a una valoración de su validez y fiabilidad, en los siguientes términos (Gómez, 1987): con la valoración de la validez de las características de un estudio dirigimos nuestra atención a la concreción y operativización de sus definiciones, a la pertinencia de la información recogida, o al grado de inferencia que es asumido por el codificador. Para aumentar la validez se ha recomendado la creación de un sistema de códigos que defina y detalle las características de

los estudios. De este modo, otros investigadores podrán valorar su validez, tanto la relativa a las características conceptuales o sustantivas como a las metodológicas. En la cuestión de la validez metodológica hay quienes han propuesto descartar del análisis, sin más, aquellos estudios con defectos metodológicos serios. Sin embargo, la perspectiva más reciente, consistente con la idea originaria de Glass (1977), propone la inclusión en un meta-análisis de todos los estudios, efectuando, eso sí, una valoración de su calidad metodológica.

En lo tocante a la fiabilidad el riesgo principal provendría de la no coincidencia codificadora entre jueces. El mejor modo de promover el acuerdo inter-jueces puede constituirlo la elaboración de un "Libro de Codificación", que defina y ejemplifique las distintas características consideradas, y sirva de guía para el vaciado de la información por distintos codificadores, previamente entrenados en una homogénea interpretación del mismo. Pese a todo, si son varios los codificadores, finalmente se requerirá una comprobación de la fiabilidad inter-jueces mediante la codificación múltiple de una muestra de los trabajos.

9.2. El problema de la validez

En el campo metodológico vivimos la era de la validez. Desde hace años la necesidad de una adecuada validación de las inferencias que se realizan en cualquier medición psicológica se encuentra a la cabeza de los criterios de adecuación metodológica, por encima de la fiabilidad de los instrumentos de medición. Así lo ha puesto de relieve en nuestro país Silva (1989), quien ofrece múltiples ejemplos de voces autorizadas que han realizado el papel de la validez.

Según la concepción más clásica, la validez se refiere al grado en que un determinado instrumento de evaluación mide aquello que pretende medir (Kerlinger, 1979). Anguera (1985) recoge la definición de Heyns y Lippitt que atribuyen a la validez una doble acepción, significando, por una parte, si una puntuación mide lo que se propone medir, y, por otra, si predice alguna cosa. Este doble sentido del término lleva a esta autora (Anguera, 1985) a la diferenciación de dos significados. El primero haría referencia a

lo que tradicionalmente se ha dado en llamar la "validez interna", que apunta en la dirección de poder afirmar o no si un cierto instrumento evaluativo, cuando es utilizado, está dando cuenta de aquel concepto psicológico o constructo que supuestamente mide. Mientras que el segundo significado vendría dado ya sea por la relación que un nuevo instrumento de medida guarda con otro semejante, previamente establecido (criterio externo), o bien por el grado en que el nuevo instrumento permite anticipar o prevenir ciertas relaciones entre variables.

Actualmente, aunque se sigue admitiendo la clásica definición de validez -medir aquello que se pretende medir-, se han formulado algunas importantes matizaciones sobre la misma (Silva, 1989). Una de ellas tiene que ver con la consideración de la validez como el núcleo de las inferencias que pueden realizarse a partir de las puntuaciones obtenidas con cierto instrumento de medición. Es decir, lo que se está validando son las interpretaciones que podemos realizar sobre la base de determinadas puntuaciones obtenidas, y no -como a veces se ha afirmado- el propio instrumento evaluativo. En segundo término, la validez no se reduce a la consideración del literal de un cierto coeficiente numérico, sino que debe ser estimada, en cada caso, a partir de cualesquiera procedimientos o datos que se empleen en el proceso científico. Finalmente, la validez es un concepto unitario, que admite, eso sí, distintos tipos de evidencia (criterial, de contenido, de constructo, etc.) y no la consideración separada de diferentes tipos de validez, como suele hacerse. En suma, en términos psicométricos actuales, en un proceso experimental o evaluativo "aunque la evidencia puede acumularse de muy distintas maneras, la validez se refiere siempre al grado en que la evidencia respalda las inferencias que se hacen a partir de las puntuaciones" (AERA, APA, NCME, en cita tomada de Silva, 1989, p. 109).

Uno de los trabajos meta-analíticos del ámbito correccional que ha prestado mayor atención al problema de la validez ha sido, sin duda, el de Lösel y Köferl (1989). Estos autores analizaron, tal y como hemos descrito anteriormente, 16 programas de prisiones alemanas. Para cada trabajo codificaron hasta 39 posibles amenazas contra la validez (que hemos recogido en la Tabla 9.1, a partir de los autores), que fueron evaluadas por dos jueces independientes en una escala de 4 puntos (amenaza inexistente, baja,

media o alta). Estas amenazas contra la validez se desglosaban en validez estadística (7 ítemes), validez interna (13 ítemes), validez de constructo (10 ítemes), validez externa (4 ítemes) y validez descriptiva (5 ítemes) (véase Tabla 9.1).

Tabla 9.1. Esquema de valoraciones sobre la validez (Trascripción a partir del trabajo de Lösel y Köferl, 1989, pág. 340).

A. Validez estadística

1. Baja potencia estadística
2. Transgresión de los requisitos de las pruebas estadísticas
3. Problema en la cuantía del error
4. Fiabilidad de las medidas
5. Fiabilidad de la aplicación de los procedimientos de medida
6. Inaplicabilidades en el contexto experimental
7. Heterogeneidad aleatoria de los sujetos

B. Validez interna

1. Historia
2. Maduración
3. Pruebas
4. Instrumentación
5. Regresión estadística
6. Selección
7. Mortalidad experimental
8. Interacción con la selección
9. Ambigüedad sobre la dirección de la influencia causal
10. Difusión o limitación de los tratamientos
11. Igualación compensatoria de los tratamientos
12. Rivalidad compensatoria de los sujetos que recibieron tratamientos menos apetecibles
13. Desmoralización resentida de los sujetos que recibieron tratamientos menos apetecibles

C. Validez de constructo

1. Inadecuada explicitación preoperacional de los constructos
2. Sesgo de mono-funcionamiento
3. Sesgo de mono-método
4. Hipótesis implícitas dentro de las condiciones experimentales
5. Interrupción de la evaluación
6. Expectativas del experimentador
7. Confusión de los constructos y niveles de constructos
8. Interacción de diferentes tratamientos
9. Interacción de pruebas y tratamientos
10. Generalizabilidad restringida a lo largo de los constructos

D. Validez externa

1. Interacción de selección y tratamiento
2. Interacción de contexto y tratamiento
3. Interacción de historia y tratamiento
4. Estabilidad temporal de los efectos del tratamiento

E. Validez descriptiva

1. Descripción del concepto de tratamiento
 2. Evaluación del concepto de tratamiento
 3. Evaluación de los objetivos de tratamiento
 4. Información sobre la estabilidad temporal del tratamiento/objetivos
 5. Información sobre las medidas estadísticas
-

Los autores relatan que aunque el acuerdo entre jueces, en la codificación de la validez, fue alto, cuando no se lograba un consenso satisfactorio entre los mismos, se buscaba finalmente eliminar la discrepancia entre ellos mediante la discusión de los ítemes controvertidos.

El empleo de todos estos parámetros de validez permitió el cálculo de un índice de amenaza para cada estudio. Para ello se sumaron los diversos niveles de amenaza de cada ítem, lo que permitió obtener una amenaza promedio para cada una de las cinco categorías de validez. Lösel y Köferl (1989) afirman haber obtenido, además, un segundo índice más complejo sobre la validez, que no detallan aunque dicen que correlacionó altamente con el primero.

Todos los estudios analizados por Lösel y Köferl (1989) mostraron un cierto grado de amenazas en todas las categorías de validez, aunque con una gran variabilidad, ya que algunos estudios que manifestaron fuertes amenazas en ciertos aspectos de la validez obtuvieron bajas amenazas en otros. Generalmente, la validez interna se mostró en menor riesgo que la validez estadística y la validez de constructo. La validez descriptiva, que definiría en qué grado una tercera persona puede comprender con claridad los conceptos relacionados con el tratamiento aplicado, los procesos que intervienen en el mismo y los resultados obtenidos, mostró los mayores niveles de amenaza.

Por otro lado, los autores pudieron también establecer las relaciones existentes en su meta-análisis entre validez y tamaño del efecto. Estas fueron bajas o inconsistentes a lo largo de los diferentes modelos evaluativos

empleados (descritos con anterioridad, en este trabajo), con la excepción de la validez de constructo, que mostró una correlación positiva a lo largo de todos los modelos. Aquellos estudios con una superior validez de constructo (mayor fundamentación teórica y mejor generalizabilidad respecto del tratamiento y sus efectos) produjeron más bajos efectos que los que tenían una validez de constructo inferior.

9.3. Confección de la primera Plantilla de Codificación

En nuestra investigación, para la creación y validación de la plantilla y el manual de codificación se siguieron las siguientes fases:

1. Se diseñó un primer borrador de plantilla de codificación que incluía 146 ítemes o informaciones en cinco grandes apartados, correspondientes con los núcleos informativos más habituales en los informes de investigación: Identificación del documento, Filosofía del programa/introducción, Método, Resultados y Discusión. Para la confección de esta primera plantilla de vaciado/codificación se atendió a las consideraciones teóricas y metodológicas formuladas por Anguera y Sánchez (1982, abril), Sánchez-Meca (1986), Gómez (1987), Cooper (1989), Johnson (1989), y Sánchez-Meca y Ato (1989), así como los trabajos meta-analíticos de Lösel y Köferl (1989) y Lipsey (1992a); también se tuvo en cuenta el esquema de plantilla codificadora empleado por Pastor (1992) en una meta-evaluación sobre programas de intervención con toxicómanos.

Este primer borrador fue sometido a la consideración de mis directores de tesis, quedando su redacción inicial según se muestra en el Apéndice 4.

2. De esta plantilla originaria se derivaron dos elementos básicos para el proceso de codificación de los estudios: el primero, un pequeño programa de base de datos, en tratamiento de textos Word Perfect 5.1, que permitía la codificación directa de cada informe de una manera muy sencilla -véase Apéndice 5-; se hizo así pensando en que una muestra de estudios habría de ser codificada -por razones de fiabilidad- por dos o tres codificadores (el autor y dos más), y era necesario un proceso de copiado y fiabilización de los ítemes lo más simple posible. Posteriormente, sin embargo, tras efectuar

diversas consultas con expertos informáticos, se optó por abandonar la base de datos de Word Perfect, simple pero rudimentaria, y rediseñar nuestras bases de datos en el programa DBIV (Data Base IV), que pese a su mayor complejidad ofrecía superiores prestaciones analíticas. El segundo elemento importante creado a partir de la plantilla originaria fue un manual de codificación -recogido en el Apéndice 6-, que definía cada ítem y proponía indicaciones y ejemplos en cada caso, para auxiliar a los codificadores en su labor. Ambos instrumentos -programa de base de datos y manual de codificación- fueron igualmente sometidos a la supervisión de mis directores de tesis, cuyas sugerencias fueron introducidas en su estructura.

3. Estudio piloto: tanto la primera "Plantilla de Codificación" como el programa de base de datos y el manual de codificación derivados fueron expuestos a un pilotaje en el se incluyeron un 10% de los trabajos recogidos, seleccionados al azar. A partir de este proceso se dio la forma definitiva a estos instrumentos de codificación.

4. Análisis, mediante un proceso de rápida lectura, de todos los materiales recogidos y selección de aquéllos que eran susceptibles de meta-análisis en función de los criterios de selección a que nos hemos referido con anterioridad.

5. Vaciado, codificación y transcripción a la base de datos de todos los estudios preseleccionados.

6. Para los análisis cuantitativos, utilización del programa DSTAT (Software for the Meta-Analytic Review of Research Literatures) de B.T. Johnson (1989) de la Syracuse University (USA).

9.4. Las sucesivas Plantillas y Manuales de Codificación

El proceso de diseño y pulimento de la Plantilla y Manual de codificación de los estudios, objeto de esta meta-evaluación, y consiguientemente de las Bases de Datos de soporte informático para la misma, han pasado por una serie de laboriosas etapas, que detallamos a continuación.

Versión 1. La primera plantilla y el primer manual de codificación, integrados por 146 ítemes o informaciones sustantivas, más un amplio número de ítemes para la valoración de los niveles de inferencia asumidos en cada caso -en total 275 ítemes- fueron probados por el autor mediante la codificación de cuatro trabajos. En ellos se observaron diversas dificultades, relacionadas especialmente con la duplicidad de ciertas informaciones interventivas y metodológicas. Por ejemplo, se pedía al codificador que, por un lado, consignara una exhaustiva información relativa a las acciones interventivas emprendidas y las técnicas utilizadas para ello, y, a continuación, desde una perspectiva evaluativa, se le requería a detallar las variables independientes que habían sido utilizadas en el estudio. En la práctica, se producía una duplicación de la información solicitada, ya que en el caso de programas interventivos, acciones emprendidas y variables independientes resultaban equivalentes en la mayoría de las ocasiones.

Por otro lado, después de cada información codificable, se había introducido un ítem en el que se pedía que el codificador valorara, en una escala de Alto/Medio/Bajo, el grado de inferencia que había tenido que realizar para la codificación de dicha información. Estas valoraciones sobrecargaban aún más la ardua tarea codificadora, haciendo el proceso más engorroso e incluso irritante, ya que el codificador era permanentemente sometido a una tediosa toma de decisión sobre la adecuación/inadecuación de las informaciones que iba obteniendo.

Esta primera prueba de la Plantilla/Manual condujo a una modificación sustancial de los mismos. Se eliminaron todos los ítemes cuya información era redundante, y todos aquéllos que requerían del codificador la estresante valoración de la inferencia realizada para la codificación de cada variable.

Versión 2. Tras la primera revisión de la Plantilla/Manual de codificación, éstos quedaron integrados por 232 ítemes, distribuidos a lo largo de seis Bases de Datos (véase Manual 2 en el Apéndice 7). Mediante esta Versión 2 -de Plantilla y Manual- se efectuó la codificación de 20 trabajos que dieron lugar al estudio piloto que describiremos más adelante.

Especial atención se concedió, en esta segunda revisión, a la conceptualización del diseño de investigación de los estudios. Por lo que se

refiere a la clasificación de los diseños, se adoptó una perspectiva, en parte clásica, en parte innovadora. Por un lado, se incluyeron, para la codificación de los estudios, las categorías más habituales de diseños estándares: (1) 1 solo grupo, con medidas pre/post (o pre-experimental); (2) con grupo de control no equivalente (o cuasi-experimental); (3) con grupo de control equivalente (o experimental); y (4) diseño de caso único. A ellas se añadieron dos categorías nuevas, que pretendían dar cabida a diseños frecuentes en este ámbito, sin necesidad de tamizarlos por los patrones de la nomenclatura experimental más clásica. La primera de estas nuevas categorías se destinó a los (5) diseños conductuales (de reversión, línea base, etc.). La segunda constituyó una categoría novedosa que denominamos (6) comparación de instituciones. Bajo este epígrafe se incluirían aquellas evaluaciones consistentes en comparar programas aplicados en prisiones, centros de reforma juvenil, o instituciones diferentes, cuando afectaban a la institución como un todo e incluían muestras cambiantes de sujetos.

Generalmente, los estudios que denominaremos comparación de instituciones se refieren a intervenciones que tienen una larga duración -de años- y se aplican, y se evalúan, sobre muestras cambiantes de individuos. Este diseño es utilizado con cierta frecuencia en los programas correccionales, especialmente en las prisiones, donde existe una gran movilidad de los encarcelados, que entran en prisión y son liberados de ella, o son, con frecuencia, trasladados a otros establecimientos. Por la inexistencia de una muestra de sujetos estable a lo largo del tiempo, constituye un caso especial, no encajable en los diseños estándares. Suele corresponderse con intervenciones de amplio espectro (modificaciones legales, cambios regimentales en una prisión, transformaciones ambientales, etc.) aplicadas sobre un elevado número de sujetos. Dada su envergadura interventiva, creímos inadmisibles que diseños de este tipo fueran despachados, sin más, en alguna categoría de pre-experimental o no catalogable.

Se prestó especial interés en este momento a la valoración de la calidad del diseño. Para ello se tomó en consideración el trabajo de Durlak, Fuhrman y Lampman (1991) que valora la calidad del diseño de investigación atendiendo a seis criterios, que, con algunas modificaciones, nosotros mantuvimos para nuestra ponderación de la calidad. En suma, estos criterios son: (1) el

tamaño muestral de los grupos de tratamiento, (2) la aleatorización de los sujetos a los grupos, o la existencia de línea base estable en los diseños conductuales, (3) la tasa de mortalidad, (4) el uso de, al menos, una medida de variable dependiente normalizada, (5) el grado de control experimental ejercido, y (6) la equivalencia de las medidas reportadas en los períodos pre- y post-tratamiento. Añadimos un nuevo criterio consistente en ponderar (7) la presencia o no de medidas pre-test. Todos estos criterios eran valorados dicotómicamente como 1/0, obteniéndose una puntuación de calidad del diseño entre 0-7 puntos.

Versión 3. Atendiendo a las consideraciones formuladas, y tras la revisión individualizada de cada ítem, se confeccionó la Versión 3 de los instrumentos -Plantilla/Manual/Bases de datos-, que dotaba al proceso codificador de una mayor racionalidad y agilidad. Esta tercera plantilla revisada, que ha sido la finalmente utilizada para codificar los trabajos de nuestra investigación, incluye 195 ítemes, que son codificados en 4 Bases de Datos. Los principales ámbitos de información recogida se refieren a los siguientes apartados: identificación del documento, características de la muestra, procedimiento de muestreo, contexto de la intervención, diseño, variables extrañas, actividades interventivas, tipo de tratamiento, duración e intensidad del tratamiento, operadores de la intervención, variables dependientes medidas, resultados, conclusiones y comentarios. Todas estas informaciones pueden verse con detalle en la Versión 3 del Manual de Codificación, del Apéndice 8.

Versión 4. Una vez terminado el vaciado y codificación de estudios, sobre la base la Versión 3 de Plantilla y Manual de Codificación, se hubieron de categorizar algunos campos y variables que, hasta ese momento, habían sido recogidos en forma narrativa (por ejemplo, el "modelo teórico", la eventual existencia de alguna "alteración psicopatológica", o el "tipo de tratamiento" aplicado). Esta necesidad de nuevas categorizaciones, así como pequeños retoques en algunas definiciones, que condujo a la completa revisión de los campos informativos afectados, dio como resultado una Versión 4, y última, de los instrumentos (Plantilla, Bases de datos y Manual de codificación), que se recoge en el Apéndice 9. En la Tabla 9.2 se relacionan, en forma abreviada, las diversas informaciones recogidas en esta versión final de los instrumentos de codificación. Éstas fueron estructuradas, a efectos del vaciado de la información, en la forma más frecuente en que suelen aparecer

recogidas en los informes de investigación que eran objeto de nuestro análisis (artículos, capítulos de libros, etc.).

Tabla 9.2. Informaciones integrantes de la versión final (Versión 4) de los instrumentos de recogida de información (Plantilla, Bases de datos y Manual de codificación).

1. IDENTIFICACIÓN DEL DOCUMENTO

1.1. Aspectos previos sobre el documento

- Número de documento
- Codificador
- Idioma del documento
- ¿Es meta-analizabile?

1.2. Identificación del documento

- Autores
- Fecha
- Título
- Libro o revista en que se encuentra o localización
- Páginas
- Ciudad
- Editorial
- Número de autores
- Disciplina de los autores
- Sexo de los autores
- ¿Fueron los autores operadores del programa o sólo evaluadores del mismo?
- Fuente de localización del documento
- Identificación de documentos complementarios

2. DESCRIPCIÓN DEL DOCUMENTO

2.1. Filosofía del programa

- Literatura referenciada por los autores
- Objetivos del programa
- Modelo teórico

2.2. Método

2.2.1. Características de la muestra

- Número total de sujetos en el post-test
- Número de grupos de tratamiento
- Número de grupos de control
- Números de sujetos de los grupos en el post-test

2.2.2. Características personales y sociales de los sujetos

Edades
Sexo

2.2.3. Características delictivas de los sujetos

Tipologías delictivas
¿Existen en la muestra sujetos catalogados como violentos?
¿Existen en la muestra sujetos con psicopatologías graves?

2.2.4. Muestreo

Criterio de selección de los sujetos para el programa
Procedimiento de muestreo
Mortalidad

2.2.5. Variables del contexto de aplicación

Lugar de la intervención
Régimen de vida, si se trata de un centro
Ciudad
Provincia
País

2.2.6. Diseño

Tipo de diseño
¿Se trata de un diseño intra o inter?
Seguimiento
Calidad del diseño
 Tamaño muestral grande
 Aleatorización
 Mortalidad baja
 Existencia de al menos una VD normalizada
 Existencia de grupo control
 Equivalencia entre medidas pre y post
 Existencia de pre-test

2.2.7. Variables extrañas

Descripción de variables extrañas

2.3. Tratamiento

2.3.1. Actividades de intervención o tratamiento (hasta 8 actividades)

Descripción de cada actividad interventiva
Sesiones aplicadas
Duración de las sesiones
Frecuencia de las sesiones
¿Qué actividad resulta dominante?
Tipo de tratamiento
Duración de la intervención en su conjunto
¿Se modificó la intensidad del programa a lo largo del tiempo?

2.3.2. Profesionales de la intervención

Número
Formación

2.4. Variables criterioles, dependientes o resultado

Descripción de cada variables dependiente (hasta 8)
Tipo
Niveles
Instrumentos de medición

2.5. Resultados

Resultados cualitativos
Técnicas estadísticas utilizadas
Medidas pre-test o grupo control
Medidas post-test o grupo experimental
Dirección de los resultados
Interacción entre variables resultado

2.6. Discusión

Conclusiones de los autores

3. OTRAS INFORMACIONES RELATIVAS A LA CODIFICACIÓN

Tiempo empleado en la codificación
Fecha
Comentario

El aspecto más importante de nuestro análisis habría de consistir, finalmente, en la indagación de posibles relaciones entre las intervenciones aplicadas y los resultados obtenidos por los programas. Por ello dedicamos una especial atención al proceso de taxonomización tanto de los tratamientos como de las variables resultado medidas.

9.5. Determinación de las categorías tratamentales

Para la determinación de las categorías tratamentales, en primer lugar, se analizaron las tipologías interventivas que habían sido establecidas previamente por otros revisores. En un capítulo precedente hemos comentado las taxonomías tratamentales utilizadas por Garrett (1985), Kaufman (1985), Gottschalk, Davidson II, Gensheimer, and Mayer (1987), Whitehead y Lab

(1989), Lösel y Köferl (1989), Andrews et al. (1990), e Izzo y Ross (1990). Hemos formulado también una crítica de las mismas. Debe comentarse el hecho de que los distintos meta-análisis basaron sus taxonomías en criterios diversos, determinados, probablemente, por la propia estructura de los estudios a los que accedieron. Algunas taxonomías clasifican los tipos de tratamiento dividiéndolos fundamentalmente en técnicas, mientras que otros los clasifican más en consonancia con teorías de partida. No obstante, en la mayoría de ellas tanto el criterio técnico como el teórico están presentes.

A continuación se atendió también a una serie de consideraciones teóricas que pasamos a comentar.

Tratamientos versus modelos teóricos. Nuestra prioridad al diseñar una taxonomía de los tratamientos se dirigió, por encima de todo, a la delimitación de las acciones interventivas emprendidas. Es decir, pretendíamos que al categorizar los tratamientos/intervenciones aplicados estuviéramos contestando, para cada estudio, a la pregunta: ¿qué fue lo que se hizo? Y no a preguntas adyacentes, aunque relacionadas con la anterior, como ¿qué teoría subyacía a lo que se hizo -lo que analizaremos en otro momento-?, ¿por quiénes se hizo?, o ¿dónde se hizo?

No obstante, es necesario un comentario sobre dos de estas cuestiones: ¿qué teoría subyacía a la intervención? y ¿dónde se llevó a cabo la intervención -dónde se hizo-?, ya que la respuesta a ellas estará en muchos casos íntimamente vinculada a la contestación a nuestra pregunta nuclear: ¿qué se hizo?

Respecto de la primera cuestión -modelo teórico subyacente- nuestra primera intención inicial fue evitar referirnos a la misma, diseñando fundamentalmente una taxonomía que fuera descriptiva de la acción o acciones emprendidas. Así lo intentamos en una primera aproximación. Sin embargo, al ensayar la catalogación de los diversos tratamientos, resultó imposible, en la práctica, no tener que acudir, en múltiples ocasiones, al modelo teórico que amparaba la intervención. Por ejemplo, si pretendíamos categorizar una acción interventiva cuyo literal decía "sesiones de grupo de comunicación y habilidades sociales", la duda planteada era si se trataba prioritariamente de una psicoterapia de grupo o de un entrenamiento en

habilidades sociales. Resolver esta cuestión nos conducía ineludiblemente a recabar más información, especialmente a partir del campo "Modelo Teórico" de soporte, e incluso de otros campos informativos.

Por otro lado, es indudable la íntima vinculación existente entre muchas técnicas y ciertos modelos teóricos, lo que les hace inseparables. Piénsese, por ejemplo, en la conexión entre la técnica economía de fichas y modelo conductual. La economía de fichas es una aplicación del modelo conductual, y de ningún otro modelo teórico. Mientras que, por ejemplo, la terapia no directiva es una derivación del modelo rogeriano, y no de otros. Es verdad que no siempre se produce esta asociación indefectible entre "acción interventiva" y "modelo teórico", encontrándose ciertas acciones que resultan de la intersección de dos o más modelos teóricos.

Sin embargo, las razones expuestas nos llevaron a desistir de nuestra pretensión inicial de separar de modo rígido técnicas y modelos teóricos. Así pues, aunque formularemos por separado una taxonomía de los modelos teóricos predicados por los autores, nuestra tipología tratamental acoge ya, como podrá verse, un inevitable referente teórico.

Internamiento versus no internamiento. Debemos comentar algo también en relación con la cuestión ¿dónde se llevó a cabo la intervención? Nuestra investigación analiza distintas estrategias empleadas en Europa para el tratamiento abordaje de la delincuencia. Y la prioridad reside en conocer, por lo que se refiere a los tipos de tratamientos, qué acciones interventivas se emprendieron: ¿se trabajó con los sujetos de modo aislado?, ¿en grupo?, o ¿se transformó, por el contrario, sus contextos con afán de influir sobre ellos? El abordaje más tradicional de la delincuencia consiste en el internamiento de los delincuentes en una institución carcelaria, juvenil, etc., para cumplir una medida o pena de privación de libertad. Frente a ello, ciertas perspectivas se plantean la no-institucionalización del sujeto, efectuando o no, paralelamente, diversas intervenciones terapéuticas. Pues bien, dada la importancia que puede tener para un individuo el hecho de permanecer o no ingresado en una cárcel, etc., se nos hace evidente que, en este punto, la dimensión internamiento/no-internamiento debería prevalecer sobre la dimensión tipo de acción emprendida.

Amplitud de la intervención. La tercera dimensión considerada en nuestra tipología interventiva fue la amplitud de la intervención, que posibilita la discriminación de las intervenciones que implican cambios relativos o específicos, frente a las que suponen cambios contextuales más importantes, que afectarían a amplios sectores de la vida de un individuo (por ejemplo, en una institución correccional, cambios en el régimen de internamiento como un todo).

Dimensión teórico-práctica. Por último, al confeccionar nuestra taxonomía, también se tomó en cuenta la dimensión teórico/práctica de las acciones educativas emprendidas mediante la intervención. Además, se incluyó una categoría "otros" para aquellos tratamientos que no pudieran ser adecuadamente categorizados.

Así pues, nuestras tipologías de tratamientos tomaron en consideración las cuatro dimensiones mencionadas:

1. Acción interventiva o técnica predominante, con referencia al modelo teórico de soporte.
2. La dimensión internamiento-versus no internamiento.
3. La dimensión amplitud de la intervención.
4. La dimensión teórico-práctica de la intervención.

En base a ellas, se elaboraron tres tipologías de los tratamientos, relacionadas entre sí.

1-PRIMERA AGRUPACIÓN DE TRATAMIENTOS: tomando en consideración las dimensiones expuestas, se confeccionó la primera taxonomía de tratamientos, con las siguientes categorías:

1 (TP) Terapia psicológica/psiquiátrica no conductual: individual, de grupo, centrada en el cliente, consejo psicológico, tratamiento de una problemática concreta, etc.

2 (ED) **Educación:** cursos, tratamiento escolar, entrega de materiales educativos para lectura, etc.

3 (TC) **Terapia conductual** (más clásica, sin componentes cognitivos): exclusivamente modelos operante (economía de fichas, reforzamiento social, etc.) y clásico (sensibilización encubierta, condicionamiento aversivo, etc.).

4 (TCC) **Terapia cognitivo-conductual:** modelo de aprendizaje social (modelado, role-playing), modelos cognitivos, competencia psicosocial, entrenamiento en habilidades sociales, etc.

5 (EE) **Mero encarcelamiento/o endurecimiento del régimen de internamiento:** la intervención predominante es el propio régimen habitual de internamiento, o el endurecimiento del mismo.

6 (CT) **Comunidad terapéutica/o suavizamiento del régimen carcelario:** la intervención predominante es la conversión en comunidad terapéutica, o asimilado, del régimen de internamiento, o el suavizamiento del mismo.

7 (AC) **Programa ambiental de contingencias de carácter globalizador:** fases progresivas, programa conductual de "padres enseñantes" (Achievement Place Model), estructuración de sistemas generales de vida y reforzamiento de conducta, etc.

8 (DI) **"Diversion":** cualquier medida alternativa al internamiento, como libertad bajo palabra, mediación, reparación, supervisión o asistencia social, tratamiento ambulatorio de problemáticas diversas, etc.

9 (OT) **Otras técnicas:** programa no catalogable.

2-SEGUNDA AGRUPACIÓN DE TRATAMIENTOS. Durante el pilotaje se habían clasificado las diversas acciones interventivas emprendidas (correspondientes a 20 estudios) en terapias, educación y entrenamiento en habilidades de conducta, cambios en el régimen carcelario y alternativas al internamiento. Posteriormente, se sometió a prueba esta tipología mediante su aplicación a los tratamientos que habían sido codificados durante el

estudio piloto. Esto se hizo comprobando si todas y cada una de las acciones interventivas de los 20 estudios del pilotaje podían ser incluidas dentro de las categorías tratamentales establecidas (véase Tabla 10.2 del estudio piloto descrito más adelante).

Finalmente, tomando como base la taxonomía tratamental obtenida durante el pilotaje, se formularon las siguientes categorías interventivas, para reagrupar las establecidas en la "primera agrupación":

A. Terapias e intervenciones especializadas no conductuales sobre problemas concretos: psicológicos, psiquiátricos, drogadicción, etc.

Categorías de primera agrupación: TP (Terapia psicológica/psiquiátrica no conductual)

B. Educación/formación: predomina la enseñanza de contenidos teóricos sobre la práctica de los mismos.

Categorías de primera agrupación: ED (Educación)

C. Terapias conductuales y cognitivo-conductuales: predomina el entrenamiento y la práctica de conducta y habilidades sobre la enseñanza de contenidos teóricos.

Categorías de primera agrupación: TC y TCC (Terapia conductual y Terapia cognitivo-conductual)

D. Cambios en el ambiente/régimen institucional como un todo (prisión, hospital psiquiátrico, centro menores o residencia para jóvenes): comunidad terapéutica, sistema de fases progresivas, programas ambientales de contingencias, programas de "padres enseñantes", suavizamiento/endurecimiento del régimen institucional, etc.

Categorías de primera agrupación: EE, CT y AC (Mero encarcelamiento/o

endurecimiento del régimen de internamiento; Comunidad terapéutica/o suavizamiento del régimen carcelario; y Programa ambiental de contingencias de carácter globalizador)

E. Alternativas al internamiento: "diversion", "probation", tratamiento en el exterior, etc.

Categorías de primera agrupación: DI ("Diversion")

F. Otras técnicas: programas no catalogables

Categorías de primera agrupación: OT (Otras técnicas)

3-TERCERA AGRUPACIÓN: MODELO TEÓRICO. Finalmente, se creó una tercera taxonomía de tratamientos correspondiente al modelo teórico implícito en las intervenciones. En base a ello, y siempre a partir de las codificaciones efectuadas en la primera agrupación de tratamientos ya comentada, se determinaron las siguientes categorías de "modelo teórico":

1-TERAPIA NO CONDUCTUAL (TP)

2-EDUCATIVO-INFORMATIVO (ED)

3-CONDUCTUAL (TC y AC)

4-COGNITIVO-CONDUCTUAL (TCC)

5-TEORÍA PENAL CLÁSICA (EE)

6-COMUNIDAD TERAPÉUTICA (CT)

7-DIVERSION (DI)

8-OTROS (OT)

9.6. Determinación de las categorías de las variables de medida o resultados

Para medir y evaluar la eficacia de las acciones interventivas que se aplican en el tratamiento de los delincuentes, pueden ser concebidos multiplicidad de indicadores dentro de distintos ámbitos temáticos o áreas de comportamiento social.

Para evaluar, por ejemplo, la eficacia de programas escolares con delincuentes se han utilizado una serie de indicadores como los siguientes: número de sujetos que asisten al aula diaria, semanal o mensualmente; rendimiento académico que muestran en pruebas objetivas periódicas; frecuencia de intervenciones sobre los contenidos dados en clase que realizan los participantes en el aula; cantidad de material de ejercicios -fichas normalizadas, etc.- que los sujetos completan fuera del aula; número de matriculados en cursos oficiales de Graduado Escolar, E.G.B., UNED, etc. Parecidos indicadores podrían elaborarse para la evaluación de los programas culturales y deportivos.

En el ámbito laboral, cabrían indicadores diversos sobre: el rendimiento laboral, que podría medirse a través del número de horas trabajadas y por la producción; el grado de satisfacción de los sujetos en su actividad laboral; sus iniciativas sobre aumento y mejora en el sistema de producción; la puntualidad y el tiempo de permanencia en el puesto de trabajo; el grado de responsabilidad, medido a través del número de trabajadores que dirige o moviliza con su propia tarea; las mejoras alcanzadas en su nivel de especialización, evaluadas mediante los estándares del mundo de la empresa, etc.

Si quisiéramos evaluar el comportamiento social y de interacción podríamos emplear indicadores como los siguientes: el incremento de los contactos verbales positivos en situaciones como reuniones o entrevistas personales; el incremento de los contactos informales en situaciones de interacción; y, dentro de las instituciones de corrección, la frecuencia y cualidad de informes positivos/negativos sobre el comportamiento de los internos emitidos por el personal.

Finalmente, puede ser evaluada también la variable reincidencia, que podríamos definir como la persistencia de un sujeto -que fue condenado por la realización de ciertos delitos- en la comisión de nuevos hechos y acciones delictivas, o la continuación de su carrera delictiva. Sin embargo, sólo podemos acercarnos al conocimiento de los valores de la variable reincidencia a través de la medición de la cuantía de ciertos indicadores, como las detenciones policiales de los sujetos evaluados, los nuevos procesamientos de los mismos, las nuevas condenas que les son impuestas o sus nuevos encarcelamientos, etc.

Para la categorización de los efectos/resultados del tratamiento se siguió una lógica semejante a la empleada para la taxonomización de los tipos de tratamiento. Como primer punto se consultaron las clasificaciones realizadas en revisiones previas, que hemos descrito anteriormente. Hallamos una especial claridad y generabilidad en la categorización formulada por Garrett (1985) que había clasificado las medidas resultado de su revisión en: medidas de delincuencia/reincidencia, psicológicas, participación escolar, resultados académicos, logros profesionales, ajuste institucional, ajuste comunitario, y otras.

Así pues, tomando en consideración, de modo especial, la tipología de Garrett (1985), se construyó inicialmente la siguiente categorización de medidas resultado:

-AJUSTE INSTITUCIONAL/REGIMENTAL (I): se incluirían aquí todas aquellas medidas relativas al comportamiento de los sujetos en las instituciones, la percepción del clima ambiental, el respeto de las normas, la violencia y los delitos cometidos en prisión, actitudes hacia la organización, etc.

-AJUSTE PSICOLÓGICO DE LA PERSONALIDAD Y DE LAS ACTITUDES (P): medidas psicológicas de rasgos y factores de personalidad, teóricamente relacionados con la delincuencia (extroversión/introversión, estrés, psicopatía, etc.), así como actitudes y reacciones emocionales sobre las normas sociales/legales, y de respeto o temor frente a la justicia, el castigo, etc.

-AJUSTE ACADÉMICO/CULTURAL (A): medidas de la evolución académica y/o cultural de los sujetos (participación y rendimiento escolar, implicación en eventos socioculturales, lectura de libros, etc.).

-AJUSTE EN EL ENTRENAMIENTO/PRACTICA LABORALES (L): medidas de los progresos en la preparación e implicación de los sujetos en el mundo de trabajo, tanto dentro como fuera de las instituciones correccionales.

-REINCIDENCIA (R): medidas relativas a la comisión de nuevas acciones delictivas, nuevos contactos con la policía, detenciones, ingresos carcelarios, etc.

-AJUSTE GLOBAL (G): medida concebida como la suma de todos los resultados evaluados en cada estudio.

Una consideración importante que debe hacerse es que nuestra tipología de efectos/resultados medidos se construyó siguiendo un criterio exclusivamente temático. Esto es, se categorizaron los resultados por el tipo de variable medida -ya fuera psicológica, conductual, hábitos, etc.-, sin tener en cuenta, en la mayoría de los casos, en qué momento era medida dicha variable -dentro/fuera de la institución correccional, durante la aplicación del programa o en el período de seguimiento-. La información sobre el "cuándo" una variable era medida nos vendría dada por el diseño de investigación, informándonos sobre la correcta interpretación de los resultados. Ahora, por el contrario, sólo nos interesaba una adecuada categorización del tipo de variables resultado medidas. Una excepción a ello la constituyó la categoría (I) Ajuste institucional/regimental que, por definición, sólo se refería a variables medidas mientras los sujetos se encontraban dentro de una institución correccional.

A continuación se sometió a prueba la adecuación de la tipología de variables resultado inicialmente diseñada. Para ello, se procedió a categorizar, en la taxonomía expuesta, todas las medidas resultado (o variables dependientes) obtenidas en los 20 estudios analizados en el pilotaje (véase Tabla 9.3, a continuación).

Tabla 9.3. Variables resultado obtenidas en el estudio piloto. En la primera columna se relacionan las medidas resultado (variables dependientes) tomadas en los diversos estudios analizados en el pilotaje. Cuando la misma medida se repetía en varios estudios, se ha consignado una sola vez. En la siguiente columna, cada medida es categorizada de acuerdo con la tipología previamente establecida. En la tercera columna, se relacionan las nuevas categorías, ideadas para suplir las carencias detectadas en el proceso de aplicación resultados-categorías.

Variables resultado medidas	CATEGORÍA	N U E V A S CATEGORÍAS
Reincidencia	R	
Nuevas condenas	R	
Dureza de las nuevas condenas	R	
Tiempo entre la excarcelación- y una nueva condena	R	
Completamiento del programa de tratamiento		SEGUIMIENTO DEL TRATAMIENTO (S)
Número de sujetos agredidos	I	
Incidentes graves	I	
Número de sujetos que perciben que "no están teniendo una estancia fácil en prisión"	P	
Actos delictivos en prisión	I	
Temor a una nueva sentencia	P	
Expectativas futuras	P	
Percepción del valor terapéutico de las sesiones de grupo	P	
Aprendizaje de mecanismos de control (interno)		HABILIDADES PARA LA INTERACCIÓN SOCIAL (H)
Desarrollo social	H	
Conducta en relación con las normas del centro	I	
Nivel de ansiedad	P	
Percepción del clima social	I	
Actitudes hacia la organización	I	
Grado de prisionización ¹⁰	P	
Medidas cognitivas (capacidad de detección y definición de		

¹⁰ La prisionización se ha descrito como la asimilación que hacen los individuos, internos en una prisión, de los hábitos, costumbres y cultura carcelaria (el primer autor que se refirió a este concepto fue Clemmer, 1940). Entre los efectos más notables de la prisionización se encontrarían: un aumento del grado de dependencia de los sujetos, la disminución de su autoestima, el incremento de su dogmatismo y autoritarismo, un aumento de su nivel de ansiedad, y, finalmente, una mayor reincidencia futura en el delito (Pérez y Redondo, 1991).

problemas, toma de perspectiva social, etc.)	H
Cambio de fase o unidad, dentro de un centro o prisión	I
Variables de personalidad (autoestima, autoritarismo, etc.)	P
Condiciones de alojamiento en la comunidad	H
Trabajo	L
Relación con las personas del entorno	H
Labilidad neurótica	P
Extroversión/introversión	P
Problemas psicosociales	H
Relaciones familiares	H
Aspiraciones educativas	A
Aspiraciones laborales	L
Problemas de conducta	H
Contactos con la policía	R
Problemas en casa	H
Utilización de los servicios carcelarios	I
Ambiente prisional	I
Sanciones disciplinarias	I
Libros prestados en la biblioteca	A
Faltas de asistencia al trabajo	L
Ajuste a la vida colectiva en una comunidad terapéutica	I
Ajuste en una familia de adopción	H
Ajuste durante el servicio militar	H

Como puede verse en la Tabla 9.3, la aplicación de la taxonomía de variables resultado diseñada a las medidas obtenidas en los primeros 20 estudios, dejó fuera algunas de ellas, evidenciando que resultaban inclasificables. Por ello fue necesario crear dos nuevas categorías de variables resultado, consignadas en la tercera columna de la Tabla 9.3, que inicialmente fueron definidas como sigue:

-SEGUIMIENTO DEL TRATAMIENTO (S): se incluyen aquí todas aquellas medidas del grado de seguimiento y/o finalización del proceso interventivo por parte del sujeto.

-AJUSTE DE LAS HABILIDADES PARA LA INTERACCIÓN SOCIAL (H): este es un paquete amplio de variables, que acoge todas las medidas relativas a

eventuales mejoras de las capacidades (habilidades de vida, cognitivas, de planificación, etc.) de interacción del sujeto con su entorno (familiar y social), o a efectivas mejoras en su conducta de adaptación social.

Así pues para la Versión 3 del Manual de Codificación (Apéndice 8), la clasificación o taxonomía de las variables resultado quedó definitivamente estructurada en los siguientes ocho tipos de medidas:

(S) SEGUIMIENTO O PERMANENCIA EN EL TRATAMIENTO: se incluyen aquí todas aquellas medidas del grado de seguimiento y/o finalización del proceso interventivo por parte del sujeto. También sus valoraciones sobre el mismo.

(I) AJUSTE INSTITUCIONAL/REGIMENTAL (exclusivamente mientras los sujetos permanecen en una institución correccional): todas aquellas medidas relativas al comportamiento de los sujetos en las instituciones, la percepción del clima ambiental, el respeto de las normas, la violencia y los delitos cometidos en prisión, actitudes hacia la organización, etc. Incluir también los cambios en la clasificación interior y en grado de tratamiento.

(P) AJUSTE PSICOLÓGICO DE LA PERSONALIDAD Y DE LAS ACTITUDES¹¹: medidas psicológicas de rasgos y factores de personalidad estables, teóricamente relacionados con la delincuencia (extroversión/introversión, psicopatía, necesidad de estimulación, hostilidad, prisionización, etc.), así como actitudes y reacciones emocionales sobre las normas sociales/legales, y de respeto o temor frente a la justicia, el castigo, etc. (no incluir factores y habilidades cognitivas).

¹¹Nota importante: en algunas medidas psicológicas de VD puede darse un conflicto serio para su clasificación como (P) Ajuste psicológico de la personalidad y de las actitudes o como (H) Ajuste en las habilidades para la interacción social. Para ello debe atenderse, en primer lugar, a las definiciones y conceptos incluidos en una y otra categorías. Ante la duda, la diferencia básica entre medidas (P) y (H) estriba en el grado de estabilidad con que los autores conciben la variable: como (P) deben catalogarse aquellos rasgos o factores de personalidad o actitudinales a los que teóricamente se atribuye alta estabilidad y baja modificabilidad; mientras que se catalogarán como (H) sólo aquellos factores y variables cognitivas conceptualmente relacionados con la conducta de interacción, a los que se atribuye una baja estabilidad y, por ello, su eventual modificabilidad mediante las intervenciones.

(A) AJUSTE ACADÉMICO/CULTURAL: medidas de la evolución académica y/o cultural de los sujetos (participación y rendimiento escolar, implicación en cursos y eventos socioculturales, lectura de libros, etc.) y actitudes y conductas en este ámbito.

(L) AJUSTE EN EL ENTRENAMIENTO/PRACTICA LABORALES: medidas de los progresos en la preparación e implicación de los sujetos en el mundo de trabajo.

(H) AJUSTE DE LAS HABILIDADES PARA LA INTERACCIÓN SOCIAL: paquete amplio de variables, que acoge: (1) todas las medidas sobre eventuales mejoras del conocimiento y de las capacidades cognitivas y factores emocionales asociados para la interacción del sujeto con su entorno familiar y social (autocontrol/impulsividad, estilo cognitivo, pensamiento concreto/abstracto, solución cognitiva de problemas interpersonales, egocentrismo/empatía, valores y desarrollo moral, razonamiento crítico y 'locus of control', ansiedad, autoestima, depresión, etc.); (2) efectivas mejoras en su conducta de interacción y comunicación (habilidades sociales, habilidades de vida, etc.); (3) mejoras en higiene, limpieza y salud (reducción de adicciones, aumento de la calidad de vida en procesos de Sida, Hepatitis, etc.); y (4) medidas de progresos y funcionamiento general, sin concretar.

(R) REINCIDENCIA: medidas relativas a la comisión de nuevas acciones delictivas, nuevos contactos con la policía, detenciones, nuevos ingresos en instituciones correccionales, etc.

(O) OTRAS MEDIDAS: no clasificables en los apartados anteriores.

A ella, finalmente, se añadiría, como queda dicho, un índice de AJUSTE GLOBAL, aglutinador de todas la medidas resultado tomadas en cada estudio.

RESUMEN

A lo largo de este capítulo hemos descrito el laborioso proceso de confección de los instrumentos de vaciado y codificación de la información de los estudios. Nos hemos referido al problema de la validez de estos instrumentos, cuestión a la que algunos autores -concretamente Lösel y Köferl, 1989- han prestado una especial atención. En nuestro caso, el diseño de instrumentos ha pasado por cuatro etapas sucesivas, que han dado lugar a otras tantas versiones del conjunto de instrumentos de codificación, cuya validez fue sometida a un permanente proceso de revisión y pulimento. En esta actividad se prestó una atención especial al diseño de las categorías de "tratamientos" y "medidas resultado". Se crearon tres taxonomías interrelacionadas para clasificar los tratamientos aplicados: (1) la primera agrupación puso el énfasis en las acciones emprendidas o técnicas utilizadas; (2) la segunda consistió en una mera reagrupación de las anteriores, en función de los resultados obtenidos durante el pilotaje realizado con 20 estudios; (3) por último, la tercera agrupación tomó en consideración el "modelo teórico" de soporte de las intervenciones aplicadas.

Por su parte, las variables resultado fueron clasificadas en ocho categorías, tomando como punto de partida la taxonomía formulada por Garrett (1985), que nosotros redefinimos y ampliamos.

10. ESTUDIO PILOTO Y PROCESO CODIFICADOR

10.1. El estudio piloto

Una vez que se hubo confeccionado (1) la primera plantilla de variables o ítemes, (2) el programa de base de datos y (3) el manual que debían servir para el proceso de vaciado y codificación de estudios, se diseñó un estudio piloto para someter a prueba la adecuación de estos instrumentos. Información sobre este estudio piloto y sobre sus principales resultados fue presentada en la "Third European Conference of Law and Psychology", celebrada en Oxford en Septiembre de 1992 (véase Apéndice 10).

Para ello, de la población de estudios recogida hasta el 13 de marzo de 1992, integrada en ese momento por 213 trabajos, se seleccionaron al azar -usando una tabla de números aleatorios- 30 estudios (un 14% del total), de los que, finalmente, se escogieron 20 que respondían a los requisitos diseñados.

Todos los estudios fueron codificados por dos o tres codificadores independientes, siguiendo la plantilla y el manual de codificación establecidos.

10.1.1. Dificultades del proceso de codificación

Gran número de ítemes y dificultades de definición de los mismos. Uno de los primeros obstáculos encontradas a lo largo del primer proceso piloto de codificación fue el gran número de ítemes e informaciones que debían ser codificadas para cada trabajo -hasta un máximo de 221, según hemos comentado-, lo que hacía que el proceso resultase, en muchos casos, muy

lento y farragoso. A ello debe añadirse la dificultad de definición de muchas de las variables implicadas (por ejemplo, "modelo teórico" que subyace a un programa, "representatividad de una muestra", "procesos psicológicos desencadenados por la intervención", etc.), así como el escollo que para el codificador suponía la identificación de algunas de estas informaciones, en el marco de la descripción de un determinado programa (piénsese, por ejemplo, en el conflicto que, dada la falta de detalle de muchas de las descripciones que realizan los autores de los estudios, puede implicar al codificador la discriminación del "diseño", la "representatividad de una muestra", la "extensión", "cadencia", "formato" y "duración" de las diversas "actividades interventivas emprendidas", la discriminación entre "resultados" y "conclusiones de los autores", etc.)

Tiempo de lectura/codificación. Todo ello hizo que el proceso de lectura y codificación de los trabajos, en este primer estudio piloto, resultara sumamente lento, siendo el tiempo medio por trabajo de 58 minutos destinados a la lectura inicial de los mismos, más 116 minutos (1 h. y 56 minutos) de promedio necesarios para la codificación de cada uno de ellos. Esta cuantificación nos llevó a formular una previsión de la necesidad de invertir más de 600 horas para llevar a cabo una sola codificación de los 213 estudios hasta ese momento catalogados. Traducido en jornadas laborales supondría la dedicación de un codificador a tiempo completo durante 15 semanas de 40 horas de trabajo. Si tenemos en cuenta que, además, algunas codificaciones deberían ser dobles, a efectos de mantener un adecuado control de la fiabilidad, la inversión de tiempo necesaria para el mero proceso de codificación adquiriría parámetros inasumibles.

Estas primeras previsiones nos condujeron a plantearnos una reducción del número de ítems de codificación, lo que redundaría en un acortamiento del tiempo necesario para la misma. Por otra parte, por desgracia y por fortuna, tras la subsiguiente selección de estudios, el número de los que eran realmente analizables quedó reducido a algo más de una tercera parte de los recogidos hasta ese momento.

Diversidad de lenguas. Otra dificultad añadida a este trabajo dimana de la diversidad idiomática de los estudios incluidos en el presente análisis, frente al monolingüismo de las meta-evaluaciones anglosajonas. En aquel momento

contábamos con trabajos en ocho lenguas distintas -después reducidas a cuatro-, lo que aventuraba un especial esfuerzo tanto personal como económico para la adecuada traducción de algunos materiales. Pese a ello, para la codificación de los primeros 20 estudios del pilotaje, los codificadores, que podían leer en los idiomas correspondientes (español, catalán, inglés, francés, italiano y portugués), no invirtieron más tiempo de promedio en codificar los trabajos en lenguas extranjeras que el utilizado para los escritos en español.

10.1.2. Primeros resultados

Nuestros primeros resultados tuvieron necesariamente un cariz eminentemente descriptivo y prospectivo. En absoluto, estas primeras informaciones extraídas podían ser consideradas representativas de la población de estudios integrantes de la investigación en su conjunto, ni por su número, ni por el nivel de análisis que entonces tenían.

En primer lugar nos referiremos a dos informaciones de interés sobre los estudios, relativas a las adscripciones profesionales de los autores y a la literatura científica referenciada en los trabajos.

Disciplinas de adscripción de los autores. Una pregunta interesante en el contexto científico en que nos encontramos se refiere a qué tipo de profesionales suelen ocuparse de aplicar programas de rehabilitación o de evaluar estos programas. De nuestra muestra de estudios, en la que se incluían 29 autores, la mayor parte -12 de ellos, o sea, el 41%- eran psicólogos, siguiendo en número 3 educadores, 2 asistentes sociales, 1 oficial de probation, 1 profesor de prisiones y 1 psiquiatra (en 9 casos no constaba la disciplina de pertenencia).

Por lo que respecta a la participación de los autores como ejecutores de los programas, en un 50% de los casos los autores participaron en la aplicación, mientras que en el otro 50% de estudios sólo actuaron como evaluadores y relatores de los programas.

Literatura referenciada en los estudios. Uno de los ítemes codificados se

dirigió a conocer qué autores y trabajos previos se referenciaban en los estudios como avaladores de las técnicas o intervenciones aplicadas. En la muestra de estudios evaluada se observaron dos tendencias a este respecto: en primer lugar, el cartel de referencias era variado, y conectado en cada caso con la temática abordada, sin que se detectara la repetición de los mismos autores en diferentes trabajos; por otra parte, se observó alguna tendencia a la "endo-referencia", o a la citación por el autor de sí mismo. Quizá se justifique esta orientación por el hecho de que cada autor lo que mejor conoce es lo que él mismo hace, y lo que hizo en el pasado suele guardar una conexión importante con su trabajo actual.

Para llevar a cabo una valoración global de los programas analizados durante el pilotaje tomamos en consideración el modelo de Rossi, Freeman y Wright (1979) para la evaluación de programas sociales, que establece como principales apartados evaluativos los siguientes: (1) detección de un problema social, (2) definición de objetivos de la intervención, (3) diseño del programa de acción, (4) aplicación del programa, y (5) evaluación de sus resultados. Valoraremos independientemente cada uno de estos apartados para el conjunto de los programas de nuestra muestra de trabajos.

(1) Problema abordado. Dentro de este apartado nos referiremos a dos factores: la edad de los sujetos tratados y el tamaño de las muestras. En relación con la edad, 5 programas se dirigieron a sujetos jóvenes, con edades entre los 11-21 años, y 9 abordaron a individuos de más de 21 años (en 6 estudios no constaba esta información). Respecto del tamaño de las muestras, 1 estudio se dirigió a un solo sujeto, 8 lo hicieron a grupos de menos de 100 sujetos, 5 a grupos de menos de 200, y 3 a muestras superiores a los 500 individuos. Como puede verse, pues, aunque algunas programas abordan muestras reducidas de sujetos, otros son pretenciosos en su planteamiento y se proponen influir sobre amplias poblaciones de delincuentes.

(2) Definición de objetivos de la intervención. Los objetivos de los programas son muy variados y dispares, oscilando desde finalidades ambiciosas, encaminadas a reducir la reincidencia de los sujetos o enseñarles nuevas habilidades de interacción social, hasta objetivos menores, desde la perspectiva de su valor comunitario, como reducir la violencia carcelaria o

mejorar el clima de las instituciones.

Mayor interés tiene el elenco de modelos teóricos que subyacen a la relación que se establece entre los objetivos propuestos y las acciones interventivas emprendidas. En la Tabla 10.1 se consignan estos modelos, para aquellos trabajos en que fueron explicitados por los autores. En una primera aproximación los hemos agrupados en terapias de diversa índole (5 estudios), educación y entrenamiento de habilidades sociales (4 programas), cambios en el régimen carcelario (5 intervenciones), y alternativas al internamiento en instituciones (3 casos).

Tabla 10.1. Modelos teóricos que subyacen a los programas de la muestra, y técnicas o acciones interventivas más frecuentes, en el estudio piloto.

Modelo teórico	Número Estudios
TERAPIAS	
Socioterapia	1
Trabajo en grupo + Asistencia individual	1
Rehabilitación residencial	1
Tratamiento Psicológico en "Diversion"	1
Terapia psicoanalítica	1
EDUCACIÓN Y ENTRENAMIENTO EN HABILIDADES	
Educación a distancia ("Open University")	1
Modelo de Competencia Psicosocial	3
CAMBIOS EN EL RÉGIMEN CARCELARIO	
Endurecimiento regimental	1
Segregación de internos conflictivos	1
Suavidad regimental-menor prisionización	1
Comunidad Terapéutica	1
TV incrementa contacto de los presos con mundo exterior	1
ALTERNATIVAS AL INTERNAMIENTO	
Proyectos alternativos al encarcelamiento	1
Aprendizaje ("Teaching Family Model")	1
Actividades normalizadas-mayor ajuste social	1
No consta modelo	3
TOTAL	20

(3) Programa de acción. Las acciones interventivas emprendidas en los diversos programas de nuestra muestra resultaron generalmente consistentes con los modelos teóricos de soporte a los nos hemos referido. No obstante, como es lógico, en la mayoría de los casos se produjo el empleo de diversas técnicas dentro de un mismo programa. En la Tabla 10.2 se relacionan las técnicas o acciones interventivas de uso más frecuente en estos primeros programas analizados:

Tabla 10.2. Acciones y técnicas interventivas más utilizadas en los programas.

TERAPIAS

Psicoterapia individual

Terapia social

Discusiones, reuniones y terapias de grupo

Asistencia individual

Relaciones delincuentes-personal-pacientes-enfermero

Reuniones periódicas delincuentes-personal

Entrevista de evaluación de problemas psicológicos

Tratamiento psicológico

Relajación

EDUCACIÓN Y ENTRENAMIENTO EN HABILIDADES DE CONDUCTA

Actividades escolares

Formación profesional

"Open University" o Educación a distancia

Entrenamiento en habilidades sociales y desarrollo de capacidades cognitivas

Actividad sociopedagógica en grupo

Modelado de conducta

Extinción de conducta

Economía de fichas

CAMBIOS EN EL RÉGIMEN CARCELARIO

Privilegios y Sanciones

Entrenamiento de estilo militar

Estricta supervisión y ocupación del tiempo

Segregación de internos conflictivos

Autogobierno

Acceso a TV individual en las celdas de la prisión

Comunidad Terapéutica

ALTERNATIVAS AL INTERNAMIENTO

Libertad bajo palabra

"Diversion" o derivación

Entrevistas y terapias familiares

Asistencia a grupos de víctimas

Convivencia en residencias juveniles externas

Familias de adopción

(4) Aplicación de los programas. Existe escasisima información en los estudios evaluados sobre la integridad en la aplicación de los mismos. En ausencia de ello, hemos optado por fijarnos ahora en la duración que tuvieron las diversas intervenciones, ya que éste es, sin duda, un parámetro de interés que se relaciona con la intensidad de los programas, y, consiguientemente, con la potencia de los efectos esperables de los mismos. Sólo en 11 programas consta información relativa a la duración, cuyo promedio fue de unos 20 meses (con un rango de entre 1,5-72 meses). También aquí, sin mayor información por ahora al respecto, cabe, sin embargo, considerar que muchas intervenciones se mantuvieron durante un período temporal prolongado.

(5) Dirección de los resultados obtenidos. De esta primera muestra sólo el 50% de los estudios (o sea 10) valoraron la dirección positiva, negativa o neutra de los efectos de la intervención llevada a cabo. En estos 10 trabajos fueron evaluados un conjunto de 29 efectos o resultados. De ellos, 23 (el 79%) fueron informados como favorables, esto es, con mejores resultados para los grupos experimentales que para los controles -si existían- o para las mediciones post-intervención que para las pre-intervención. En 4 casos se constataron efectos nulos de las intervenciones, mientras que en 2 situaciones los efectos fueron negativos. No obstante, esta aparente gran abundancia de efectos exitosos debe ser tomada con gran cautela, ya que, los efectos informados fueron, ciertamente, muy heterogéneos en su relevancia. Así, mientras algunos resultados hacían referencia a disminuciones en los valores de la reincidencia, muchos de ellos se referían tan sólo a meras apreciaciones subjetivas de los propios sujetos participantes o de los evaluadores de los programas.

10.1.3. Conclusiones

A la vista de lo expuesto en este primer análisis, no es todavía el momento para extraer conclusiones sobre los programas correccionales aplicados en Europa. Nuestra única pretensión inicial fue la presentación de los primeros pasos dados en esta investigación.

Lo actuado hasta ahora en nuestro estudio piloto, nos lleva a tomar en cuenta dos consideraciones, de signo contrario, para el futuro desarrollo de esta

investigación. La primera tiene una dimensión economizadora de esfuerzos, y se refiere a la probable necesidad de reducir el número de ítemes e informaciones que son codificadas para cada trabajo, con tal de aligerar el proceso codificador.

Por lo que se refiere a la sistematización que, en el futuro, deberían tener los propios informes sobre los programas, coincidimos con Lösel y Köferl (1989) en su afirmación de la necesidad imperiosa de una descripción más cuidadosa de los contenidos y objetivos de los tratamientos, una mayor explicitación teórica de sus fundamentos, una mayor diferenciación de las evaluaciones globales, transformándolas en pequeños estudios de valoración más precisos, con análisis empíricos, y una mayor orientación evaluativa, en suma, para facilitar su integración posterior.

Por el momento, sin embargo, por lo que se refiere a la integración de programas correccionales europeos que llevamos a cabo, consideramos necesario adoptar un criterio extensivo en relación con la selección de los estudios que deben ser incluidos en el análisis. Esto es, nos proponemos tomar en consideración en esta meta-evaluación el mayor número posible de programas e iniciativas interventivas, aun que no posean a menudo la sistemática y elaboración metodológica que serían aconsejables. Creemos que el primer paso debe darse para conocer ordenadamente lo que existe, aunque sea irregular en su desarrollo. Más tarde, para parte de esta información, será, tal vez, posible un acercamiento metodológico más estricto y depurado.

10.2. La codificación de los estudios

No podemos finalizar este capítulo sin referirnos, aunque sea brevemente, a la que ha sido sin duda alguna la tarea más ingente y laboriosa de esta investigación y de la que, sin embargo, menor constancia ha quedado en este informe: el proceso de codificación de estudios.

En otros meta-análisis, de voluminoso número de trabajos, como en el de Lipsey (1992a), la actividad codificadora fue encomendada a terceras personas -alumnos de doctorado en el caso de Lipsey- que fueron entrenados

por el autor en el uso del instrumental evaluativo.

En nuestro caso también se pensó, inicialmente, en contar con la ayuda para la codificación de diversos profesionales expertos que se mostraron dispuestos a ello, y a recibir la instrucción adecuada para la utilización de los instrumentos. Sin embargo, una vez realizado el pilotaje y vista la dureza y complejidad inherentes a la tarea de vaciado y codificación se desistió de esta primera intención. Por un lado, el aumento del número de codificadores habría puesto en serio riesgo la validez y fiabilidad del proceso de codificación. La información requerida, que debía extraerse de los informes originales, era sobremanera amplia (finalmente, hasta un máximo de 200 ítemes) y no siempre lo suficientemente evidente, como para que no se hubiera incrementado la frecuencia de errores provenientes de un número amplio de codificadores. Por otro, el propio autor tenía el deseo de conocer, de primera mano, el mayor número posible de los programas aplicados.

Ello condujo a que se limitaran al máximo los codificadores, que trabajaron en las siguientes etapas:

1) Inicialmente dos codificadores (el Prof. Garrido y el autor) revisaron los 250 informes recogidos para efectuar la selección de aquéllos que, en principio, reunían los requisitos para el meta-análisis. En esta tarea se invirtió un promedio de 5 minutos por estudio.

2) El vaciado y codificación de los estudios fueron realizados por tres codificadores: la Prof. M.T. Anguera efectuó el vaciado y codificación de 13 trabajos; el Prof. V. Garrido de 25 informes; y el propio autor vació y codificó 50 estudios más. El tiempo neto invertido en esta actividad fue de unas 200 horas, a un promedio de 2,33 h. por trabajo. Pese a ello, finalmente, sólo se presentan resultados correspondientes a 57 estudios, mientras que, si sumamos el número de codificaciones distintas llevadas a cabo, éstas ascendieron a 88. Ello es debido a dos razones: una, a que algunos estudios fueron codificados repetidamente (por 3 codificadores) para efectuar análisis de validez y fiabilidad; la segunda, que un cierto número de estudios, inicialmente codificados, hubieron de ser a la postre eliminados de la muestra debido a carencias metodológicas o estadísticas sustantivas. Este fue el caso, por ejemplo, de todos los estudios de caso único, cuya

integración en el meta-análisis resultó inviable.

3) Paralelamente, el Prof. J. Sánchez-Meca se encargó, junto con el propio autor, de la recodificación de las variables metodológicas (diseño, tamaños muestrales, etc.), así como de la realización de todos los cálculos de tamaño del efecto, de la potencia estadística y de los análisis de resultados. A su vez, los cálculos fueron rehechos, a efectos de su fiabilización, por el Prof. F. Marín. En esta tarea se invirtió un tiempo promedio de 1,15 h. por estudio, lo que supone un tiempo total de 142 horas netas.

En síntesis, el trabajo de selección, vaciado y codificación de estudios, cálculo de los Tamaños del Efectos y de la potencia estadística ocupó un tiempo de trabajo de en torno a 362 horas, y se prolongó durante más de un año y medio. Se trata del trabajo oscuro, no vistoso, de una investigación como la presente, a cuya descripción, comentario y recuerdo nos ha parecido necesario dedicar cuando menos un pequeño apartado de este informe.

RESUMEN

Como paso intermedio en el desarrollo de esta investigación efectuamos un estudio piloto para poner a prueba la adecuación de los instrumentos de vaciado y codificación de la información. Éste se efectuó seleccionando al azar 20 estudios, que fueron codificados por dos o tres codificadores independientes. El pilotaje nos permitió, en primer lugar, detectar las dificultades más importantes con las que nos enfrentábamos: el gran número de informaciones requeridas a los codificadores y los escollos de definición de algunas de las variables; el dilatado tiempo necesario para la codificación de los trabajos; y la diversidad de lenguas de los informes. También nos ofreció unos primeros resultados, de cariz descriptivo, sobre los tamaños de las muestras tratadas, los objetivos de las intervenciones, las más frecuentes acciones de tratamiento emprendidas, así como sobre los modelos teóricos en los que éstas se fundamentaban.

En segundo término, hemos dedicado un pequeño apartado en este capítulo a describir la ingente laboriosidad y dedicación que supuso el proceso de codificación de los trabajos analizados.

11. DISEÑO DE INVESTIGACIÓN Y ANÁLISIS DE DATOS

11.1. Diseño de investigación

Ya hemos comentado que nuestra investigación se dirige principalmente a estudiar la efectividad de diversos programas de rehabilitación de delincuentes, realizados en Europa durante la pasada década. Así pues, no se trata de un experimento manipulativo diseñado por nosotros ex profeso con esta finalidad, sino de un análisis secundario de información, en el que trabajamos con datos provenientes de estudios evaluativos ya realizados por otros investigadores. La estrategia básica seguida por nosotros, mediante la técnica meta-analítica, consiste en la reordenación y sistematización de todo el amplio conjunto de información recabado, en forma tal que podamos obtener conclusiones sobre nuestras preguntas básicas de investigación: ¿Son efectivos los programas de tratamiento de la delincuencia? ¿En qué circunstancias?

Para ello hemos estructurado un sistema de recogida y ordenación de los datos, que actuará esencialmente con los siguientes factores (véase Figura 11.1): en función de variables independientes (VIs) operarán: (1) las características del tratamiento, tales como el tipo de técnica básica utilizada, el modelo teórico de soporte, y la duración, intensidad, etc., de la técnica; (2) las características de los sujetos, tales como su edad y tipología delictiva; (3) las características del ambiente o contexto de la intervención, como lugar y régimen de la aplicación, país, o fecha; y (4), finalmente, las características metodológicas del estudio, diseño de investigación, calidad del estudio, etc. Serán nuestras variables dependientes (VDs) los resultados obtenidos en los diversos estudios, que hemos categorizado en medidas de:

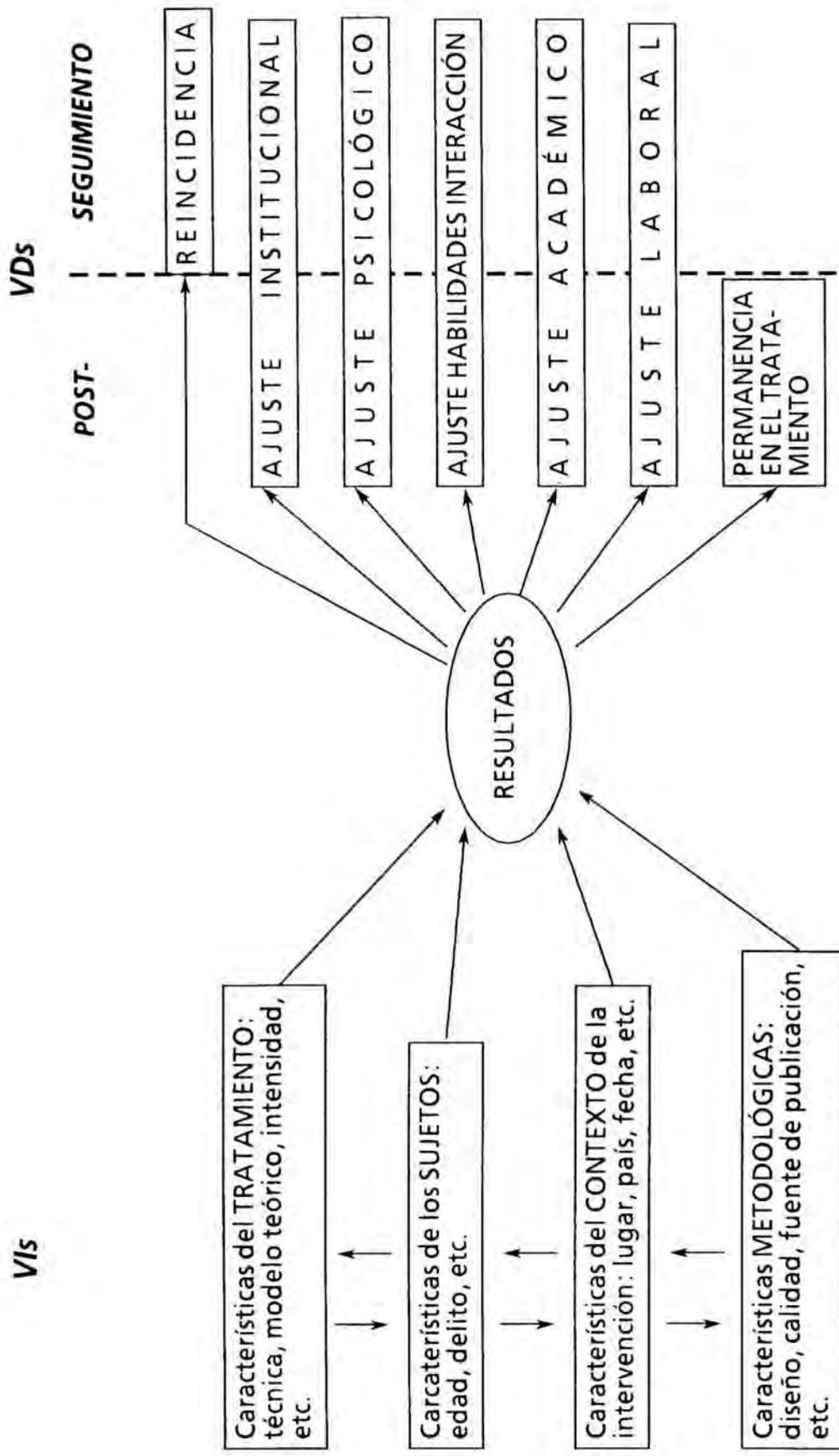


Figura 11.1 Esquema del diseño de nuestra investigación.



(1) reincidencia; (2) ajuste institucional; (3) ajuste psicológico; (4) ajuste de las habilidades para la interacción social; (5) ajuste académico; (6) ajuste laboral; (7) y permanencia en el tratamiento. Todas ellas pueden corresponder a dos momentos evaluativos diferentes: postest y seguimiento. Ambos conjuntos de variables -independientes y dependientes- serán puestos en relación a través de los diversos análisis de datos que se llevarán a cabo.

El enfoque meta-analítico aplicado en nuestro estudio es, según la clasificación de Bangert-Drowns (1986), el planteado por Hedges (1981, 1982a, 1982b; Hedges y Olkin, 1985). Básicamente, consiste en utilizar técnicas estadísticas de integración de TEs en las que se tiene en cuenta la fiabilidad de cada TE ponderando éste por la inversa de la varianza del TE. De esta forma, se asigna más peso específico a los estudios más fiables, los cuales generalmente coinciden con los que tienen tamaños muestrales mayores y una menor variabilidad en los factores estudiados.

El proceso de análisis de los datos de este meta-análisis implica diversas etapas, que pasamos a detallar.

11.2. Estudio de la fiabilidad de la codificación de las variables

Según hemos puesto de relieve, una de las dificultades más importantes de nuestra investigación estriba en la adecuada comprensión de algunas de las informaciones que son codificadas sobre cada uno de los estudios. En suma, nos referimos al problema de la calidad o fiabilidad de los datos. Con la finalidad de controlar la calidad de los datos -del proceso codificador-, nuestra prueba más importante consistirá en el cálculo del grado de acuerdo existente entre el autor -codificador principal - y dos codificadores independientes, en lo tocante a las variables más relevantes: categorías tratamentales y categorías de variables resultado.

La concordancia interjueces constituye una garantía importante respecto al riesgo de inferencia existente, el cual es mayor en observación indirecta, y por tanto en el análisis de documentos, que en el registro de conducta.

Las restricciones relativas a la concordancia entre codificadores se han intensificado en los últimos años, incidiendo tanto en el número de ellos como en la técnica seguida. El número mínimo exigido se ha elevado de dos a tres, con lo cual se acorta la distancia entre concordancia y precisión, y en cuanto a la opción por uno u otro coeficiente de acuerdo, predomina en la actualidad la elección de aquéllos que controlan el efecto del azar.

Así pues, calcularemos el grado de acuerdo entre tres codificadores aplicando la forma canónica formalizada inicialmente por Spiegelman, Terwilliger y Fearing (1953) y revisada por Krippendorff (1980), cumpliendo estrictamente los requisitos indicados. Para ello, seleccionaremos 20 estudios al azar y se pedirá a dos codificadores independientes que, siguiendo el Manual de instrucciones, codifiquen todas las variables tratamiento y resultado. La tercera codificación será la realizada por el propio autor.

11.3. Estudio de la fiabilidad de los cálculos de los TEs y de la potencia estadística

Con respecto al análisis de la fiabilidad de los cálculos de los TEs se ha realizado el siguiente proceso. Se seleccionaron aleatoriamente un subconjunto de 20 estudios que fueron analizados de forma independiente por dos codificadores. Para esta codificación paralela se utilizó un manual de cálculo de los índices del TE y de la potencia estadística (Sánchez-Meca, 1993b). Una vez obtenidos los resultados, las discrepancias se resolvieron mediante consenso. Para determinar la fiabilidad en el cálculo de los TEs y de la potencia, se obtuvo el coeficiente de correlación de Pearson entre los valores registrados por los dos codificadores. Se decidió utilizar como criterio mínimo para suponer una fiabilidad adecuada el valor de .80 recomendado por Stock et al. (1982).

11.4. Estudio del sesgo de publicación

Se conoce por sesgo de publicación el hecho bien establecido de que los trabajos publicados suelen poseer significación estadística con mayor frecuencia que los trabajos no publicados (Rosenthal, 1991). Como

consecuencia, es posible que en algunos campos de investigación los estudios publicados muestren un mayor TE que los no publicados. Por ello resultaría problemático aplicar meta-análisis sólo sobre estudios publicados: la efectividad media podría resultar una sobreestimación de la verdadera efectividad global. La ventaja de nuestro meta-análisis radica en que disponemos tanto de estudios publicados como no publicados. Así, para comprobar si el sesgo de publicación es un factor que pueda afectar a los resultados de nuestro meta-análisis, decidimos comparar la efectividad media de los estudios publicados frente a la de los no publicados.

11.5. El cálculo de la efectividad media o de los TEs

Para cada una de las variables dependientes definidas realizaremos un meta-análisis mediante el cual obtendremos los siguientes estadísticos: (a) el TE medio; (b) su correspondiente intervalo de confianza al 95%, para determinar si cada TE difiere de 0 (lo que supondría una efectividad nula); y (c) la prueba Chi-Cuadrado de homogeneidad de todos los TEs en torno al TE medio, para comprobar si éste los representa correctamente o si, por el contrario, es preciso examinar variables (o características de los estudios) que puedan explicar tal variabilidad.

Si la prueba Chi-Cuadrado de homogeneidad no resultara significativa, ello sería indicativo de que todos los estudios se encontrarían bien representados por el TE medio global y, en consecuencia, el análisis se detendría en este punto. Pero esto suele ser infrecuente. Cuando la prueba Chi-Cuadrado de homogeneidad resulta significativa, es preciso pasar a explorar aquellas variables de los estudios que pueden explicar tal heterogeneidad.

No obstante, la búsqueda de variables moderadoras debe hacerse cuando el número de TEs integrados no es demasiado pequeño. Para nuestro meta-análisis decidimos llevar a cabo esta fase de análisis sólo en aquellos casos en los que el número de TEs integrados fuera igual o superior a 10.

11.6. Búsqueda de variables moderadoras (en función de variables independientes)

Para explicar la posible variabilidad de los TEs de los estudios, procederemos a examinar la relación existente entre cada variable potencialmente moderadora y los TEs. Se tratará, pues, de realizar un análisis individual de cada variable susceptible de actuar como moderadora. Cuando la variable moderadora es cualitativa, este análisis se efectuará mediante análisis de varianza (ponderado por la inversa de la varianza de cada TE). Por el contrario, si la variable moderadora es cuantitativa, se aplicará análisis de regresión simple (también ponderado por la inversa de la varianza de cada TE).

El análisis de varianza aporta una prueba Chi-Cuadrado inter-clases, QB , con $p-1$ grados de libertad (siendo p el número de clases), calculada sobre la Suma de Cuadrados debida a los Tratamientos que refleja el grado de discrepancia entre los TEs medios de las clases o categorías de la variable moderadora. Además, este modelo de análisis incorpora una prueba Chi-Cuadrado intra-clase, QW , con $K-p$ grados de libertad (siendo K el número total de TEs integrados), calculada sobre la Suma de Cuadrados de Error que refleja el grado de homogeneidad dentro de las clases. Asimismo, la prueba QW se descompone en tantos valores QWi como clases tenga la variable moderadora, con objeto de determinar qué clases son homogéneas y cuáles no lo son; para ello, cada valor QWi se considera una prueba Chi-Cuadrado con m_i-1 grados de libertad (siendo m_i el número de TEs de la clase i).

Cuando QB resulta significativo, la variable moderadora influye sobre los TEs. Pero si, simultáneamente, QW también resulta significativo ello indica que todavía existe excesiva heterogeneidad dentro de las clases, por lo que debe haber otras variables moderadoras influyentes. Si, por el contrario QB no resulta significativo, podremos afirmar que la variable no afecta a la magnitud de los TEs.

El análisis de regresión simple, siguiendo un modelo similar al del análisis de varianza, también aporta un índice de la influencia de la variable moderadora, en concreto un valor T que se contrasta con la distribución t

de Student con $K-2$ grados de libertad (siendo K el número de TEs integrados). También incorpora un estadístico Chi-Cuadrado, QE , que evalúa la "especificación" del modelo, es decir, si pueden existir o no otras variables moderadoras de los TEs. Este estadístico tiene $K-2$ grados de libertad.

La interpretación de los resultados es similar a la del análisis de varianza. Si el valor T resulta significativo, podremos afirmar que la variable moderadora está asociada a los TEs. Si, al mismo tiempo, el valor QE también resulta significativo, ello indicará que el modelo está mal "especificado", es decir, que deben existir otras variables potencialmente influyentes sobre los TEs. Si el valor T no es significativo, tendremos que concluir que la variable no influye sobre los TEs.

11.7. Formulación de un modelo explicativo

La fase final del proceso de análisis, y la más creativa, consistirá en buscar un modelo capaz de explicar el máximo de la varianza de los TEs. Con éste se pretenderá identificar el conjunto de variables moderadoras que mejor dan cuenta de la magnitud de los TEs. Para ello, partiremos de cuatro bloques de variables atendiendo a la clasificación conceptual que, según ya hemos visto, divide las variables en (Lipsey, 1992a, 1992b): (a) variables relacionadas con las características del tratamiento o intervención; (b) variables relacionadas con las características de los sujetos receptores del tratamiento; (c) variables relacionadas con el ambiente en el que se ha aplicado el tratamiento; y (d) variables metodológicas relacionadas con el tipo de diseño y las condiciones bajo las que se han realizado los análisis estadísticos.

En función de criterios conceptuales y de los resultados obtenidos en la fase de búsqueda de variables moderadoras, seleccionaremos el conjunto de variables que se consideren más relevantes y se formulará un modelo de regresión múltiple (ponderado por la inversa de la varianza de los TEs). Para determinar la importancia relativa de cada conjunto de variables moderadoras, aplicaremos un análisis de regresión jerárquico. Siguiendo a Lipsey (1992a, 1992c), el orden de introducción de las variables en el modelo

irá condicionado por la conveniencia de incluir, en primer lugar, el conjunto de variables metodológicas con objeto de controlar su posible influencia sobre los TEs. Una vez introducido este grupo de variables, se incorporarán al modelo las variables moderadoras por bloques atendiendo a la clasificación arriba especificada. De esta forma, podremos determinar si realmente las variables relacionadas con los tratamientos, con los sujetos y con el ambiente influyen sobre los TEs, una vez eliminada la posible influencia de las variables metodológicas.

RESUMEN

Nuestro diseño de investigación estructura la recogida y el análisis de datos haciendo operar las características del tratamiento, de los sujetos, del contexto interventivo y de la metodología de los estudios como variables independientes, mientras que las diversas medidas resultado constituyen nuestras variables dependientes.

Hemos presentado las principales técnicas estadísticas que serán utilizadas en nuestros análisis. Seguiremos el enfoque meta-analítico formulado por Hedges (1981, 1982a, 1982b), consistente en emplear técnicas de integración de los TEs que toman en consideración la fiabilidad de cada TE, asignando mayor peso a aquellos estudios con tamaños muestrales mayores y menor variabilidad.

Realizaremos una serie de análisis preliminares. Para garantizar la fiabilidad del proceso codificador será calculado el grado de acuerdo entre tres codificadores, para una muestra de estudios y variables, siguiendo la forma canónica propuesta por Spiegelman, Terwilliger & Fearing (1953) y revisada por Krippendorff (1980). Paralelamente, se analizará la fiabilidad de los cálculos de los TEs y de la potencia estadística. Para ello se seleccionarán 20 estudios al azar que serán analizados por dos codificadores independientes, obteniendo a continuación la correlación entre los valores registrados por ambos. Además, se controlará el sesgo de publicación comparando los TEs promedio de los subgrupos de estudios publicados y no publicados.

El punto nuclear de nuestro análisis se dirigirá a obtener: (a) el TE medio

para cada una de las variables dependientes definidas; (b) su correspondiente intervalo de confianza; y (c) la prueba Chi-Cuadrado de homogeneidad de todos los TEs, para examinar su eventual homogeneidad o heterogeneidad. Si detectamos heterogeneidad en los TEs integrados, emprenderemos la búsqueda de variables moderadoras, analizando -mediante análisis de varianza o análisis de regresión- la relación existente entre cada una de las variables independientes consideradas y los TEs. Finalmente, ensayaremos -mediante análisis de regresión múltiple- la construcción de un modelo explicativo que pueda dar cuenta del mayor grado posible de variabilidad observada.

T E R C E R A P A R T E :
R E S U L T A D O S

12. ESTUDIO DE LA FIABILIDAD

12.1. Fiabilidad de la codificación

Con objeto de garantizar la fiabilidad del proceso codificador se llevaron a cabo tres estudios diferentes de la fiabilidad. El primero tuvo lugar durante el pilotaje, y los otros dos con posterioridad.

12.1.1. Primer estudio de fiabilidad (por consenso de expertos)

Con la Versión 2 de los instrumentos de codificación (Plantilla/Manual/Bases de datos) se llevó a cabo el estudio piloto, al que nos hemos referido, con objeto de someter a prueba la viabilidad y pertinencia de los instrumentos diseñados para la codificación de los estudios. Para ello, 20 trabajos, seleccionados al azar de entre los 215 estudios recogidos en ese momento (marzo de 1992), fueron codificados sucesivas veces por codificadores distintos, especialmente instruidos en el manejo de la Plantilla/Manual/Bases de Datos de codificación.

Toda la información codificada permitió obtener una serie de consultas en DBIV, con objeto de comparar las codificaciones repetidas. Con toda esta información se llevó a cabo una sesión de fiabilización por consenso de expertos de la Plantilla/Manual/Bases de Datos. Esta sesión tuvo lugar con mis directores de tesis, profesores María Teresa Anguera y Vicente Garrido. Posteriormente, se mantuvieron dos jornadas de trabajo con el profesor Julio Sánchez Meca de la Universidad de Murcia, experto español en meta-análisis. En estas sesiones, a la luz de la experiencia adquirida en la codificación piloto, se debatió la pertinencia de cada ítem de la plantilla, decidiendo la

eliminación o reformulación de algunos de ellos, sobre la base de los siguientes criterios:

-Se prescindió de aquellos ítemes que no habían llegado a ser codificados durante el pilotaje, o lo habían sido en con una frecuencia muy baja, debido a la inexistencia o escasez de estas informaciones en los estudios. Por ejemplo, los ítemes Nacionalidad de los autores, Representatividad de la muestra, y Raza, Status familiar, Nivel socioeconómico y Nivel académico de los sujetos, fueron eliminados de la codificación debido a la ausencia casi permanente de tales informaciones.

-Se eliminaron también aquellos ítemes cuya codificación requería un elevado grado de inferencia, ya que no suelen ser explicitados en la mayor parte de los trabajos. En este caso se encontraban, por ejemplo, los ítemes Hipótesis.

-Se refundieron algunos ítemes cuyo desglose inicial aumentaba el grado de inferencia requerido. De este modo, los ítemes originarios Objetivo1, Objetivo2 y Objetivo3 -concebidos con la finalidad de estructurar secuencialmente los objetivos del programa- quedaron reducidos a un solo Objetivo general o genérico del programa.

-Se redefinieron ciertas informaciones como el nombre del Diseño evaluativo utilizado, que inicialmente aparecía como un ítem abierto, acotándolo a una serie de tipos habituales de diseños.

-Finalmente, la reforma de los ítemes correspondientes al apartado "resultados" se pospuso al conocimiento preciso de la información requerida, en este punto, por el programa específico de meta-análisis que iba a ser utilizado, el DSTAT de Johnson (1989), al que nos hemos referido con anterioridad.

De modo semejante se procedió en lo relativo a la codificación de las variables metodológicas (tipo y calidad del diseño de investigación). En una sesión conjunta del autor con el Dr. Sánchez Meca se efectuó la revisión puntual de cada una de la codificaciones, decidiéndose las mismas por consenso.

12.1.2. Segundo estudio de fiabilidad (intra-codificador)

Mediante el procedimiento de consenso anteriormente descrito, la plantilla de codificación pareció quedar suficientemente ajustada en términos de fiabilidad. Pese a ello, se efectuó un específico estudio de fiabilidad intra-codificador en lo concerniente a la categorización de las variables dependientes o resultado. Durante el vaciado de información de cada estudio, un codificador había categorizado, al mismo tiempo, las variables resultado en el marco de la tipología establecida (S-medidas de seguimiento o permanencia en el tratamiento de los sujetos; I-medidas de ajuste institucional; P-medidas de ajuste psicológico; A-medidas de ajuste académico; L-medidas de ajuste laboral; H-medidas de habilidades para la interacción; R-medidas de reincidencia; y O-otras). Este proceso de vaciado, y codificación paralela de algunas variables, fue lento y se prolongó entre marzo de 1992 y mayo de 1993. Durante el mismo se categorizaron 173 medidas distintas. Transcurridos algunos meses -en julio de 1993- el autor -codificador principal- efectuó una nueva codificación ciega de las variables resultado previamente categorizadas. Debido al tiempo transcurrido y al gran número de variables, no cabe pensar que la eventual concordancia entre las dos codificaciones pudiera estar determinada por el factor recuerdo. Como resultado de este estudio de fiabilidad intra-codificador se obtuvo un nivel de acuerdo global, entre la primera y la segunda codificación, del 94,79% (rango 88.88% - 100%, según categorías de variables) (véase Tabla 12.1).

Tabla 12.1. Fiabilidad intra-codificación para las variables resultado. Donde Ob es la frecuencia de observaciones, Ac es el número de acuerdos entre observaciones, y %Ac es el porcentaje de acuerdo existente en cada caso.

	Ob	Ac	% Ac
S-medidas de permanencia en el tratamiento	9	8	88.88
I-medidas de ajuste institucional	21	20	95.23
P-medidas de ajuste psicológico	12	11	91.66
A-medidas de ajuste académico	15	15	100
L-medidas de ajuste laboral	11	11	100
H-medidas de habilidades para la interacción	65	59	90.76
R-medidas de reincidencia	40	40	100
TOTAL	173	164	94,79

12.1.3. Tercer estudio de fiabilidad (inter-codificadores)

Se realizó una comprobación final de la fiabilidad entre codificadores en relación con las dos variables más importante en términos sustantivos de nuestra investigación: (1) nuevamente la codificación de las variables dependientes o resultados, y (2) la codificación de los tipos de tratamiento o intervenciones. El codificador principal ya había codificado estas variables durante el proceso de vaciado de información. Los tratamientos habían sido catalogados en una taxonomía de 9 categorías (TP-terapia psicológica/psiquiátrica no conductual; ED-educación; TC-terapia conductual; TCC-terapia cognitivo/conductual; EE-mero encarcelamiento-endurecimiento del régimen de internamiento; CT-comunidad terapéutica; AC-programa ambiental de contingencias de carácter globalizador; DI-"diversion"; y OT-otras técnicas).

Para analizar la fiabilidad inter-codificadores se seleccionaron 24 estudios (el 42%) al azar y se pidió a dos codificadores expertos -mis directores de Tesis, Dra. Anguera y Dr. Garrido- que, de manera ciega e independiente, recodificaran tanto las variables resultado como los tipos de tratamiento. Para la codificación de los tipos de tratamiento se obtuvo una fiabilidad inter-codificador, para tres codificadores, de 86.2% , y para la codificación de las variables resultado la fiabilidad inter-codificadores -también para tres codificadores- fue de 78.5%. En ambos casos, la fiabilidad resulta plenamente aceptable, de acuerdo con los criterios estadísticos estándares.

Paralelamente a la taxonomía de tratamientos enunciada, se habían establecido dos taxonomías más de tratamiento, una denominada de segunda agrupación y otra modelos teóricos de las intervenciones. Como quiera que estas dos categorías consistían en una mera transformación automática de la primera taxonomía -mediante la definición de una serie de equivalencias-, la alta fiabilidad obtenida para aquélla aseguraba la fiabilidad de estas dos últimas categorizaciones.

12.2. Fiabilidad de los cálculos

Dentro de un meta-análisis, algunos estudios suelen facilitar una información estadística muy deficitaria, que nos obliga a realizar cálculos indirectos del tamaño del efecto. Por ello es conveniente que exista un libro de codificación que describa con precisión las pautas de obtención indirecta de los estadísticos, que deberá someterse a un proceso de fiabilización inter-jueces (Gómez, 1987). Así se llevó a cabo en nuestro caso. Según ya hemos comentado, se seleccionaron al azar un conjunto de 20 estudios que fueron analizados independientemente por dos codificadores. Estos 20 estudios dieron lugar a un total de 24 estimaciones del TE, ya que uno de los estudios incluía cuatro experimentos independientes y otro estudio incluía dos. El coeficiente de correlación de Pearson obtenido entre los 24 valores del TE calculados por los dos codificadores fue de 0.826, el cual resultó estadísticamente significativo ($p = .000$).

Paralelamente, se procedió a analizar la fiabilidad de los cálculos de la potencia estadística. De los 20 estudios seleccionados sólo 10 aplicaron

alguna prueba de significación, por lo que la fiabilidad de los cálculos de potencia se limita al coeficiente de correlación de Pearson obtenido entre los 10 valores, el cual fue de 0.844, resultando también estadísticamente significativo ($p = .002$).

RESUMEN

En este primer capítulo, dentro del apartado "resultados", hemos comenzado por referirnos a los estudios que fueron llevados a cabo para analizar la fiabilidad.

Durante el pilotaje previo se efectuó un primer análisis de fiabilidad inter-jueces por el procedimiento de revisión y "consenso" de todos y cada uno de los ítems de la Versión 2 de los instrumentos de codificación. Más tarde se realizaron tres estudios más sobre la fiabilidad codificadora. En uno de ellos se analizó la fiabilidad intra-codificador en lo concerniente a las variables dependientes o resultado. La fiabilidad intra-codificador promedio fue de 94.79. El segundo consistió en un análisis de la fiabilidad inter-codificadores -para tres codificaciones distintas- tanto en lo relativo a la categorización de los tratamientos -para la que se obtuvo un fiabilidad global del 86.2%- , como para las variables resultado -cuya fiabilidad resultó del 78.5%-.

Finalmente, un cuarto estudio calibró la fiabilidad de los cálculos de TE y la potencia estadística. La fiabilidad global de las estimaciones de los TEs, para una muestra de 20 estudios, obtuvo un coeficiente de correlación de Pearson de 0.826 ($p = .000$). Como fiabilidad de los cálculos de potencia estadística se obtuvo un coeficiente de 0.844 ($p = .002$).

13. ANÁLISIS DESCRIPTIVO DE LAS CARACTERÍSTICAS DE LOS ESTUDIOS

13.1. Estudios meta-analizados

En la Tabla 13.1 se recoge la relación final de estudios que han formado parte de esta integración meta-analítica.

Tabla 13.1. Relación de estudios meta-analizados. En la columna "Informes" se consignan los trabajos que fueron codificados. Los informes con asterisco (*) tuvieron que ser finalmente descartados debido a carencias severas en su metodología y/o resultados. En la columna "Núm. estudios extraídos" se reflejan aquellos informes de los que se extrajo más de un estudio.

INFORMES	NÚM. ESTUDIOS EXTRAÍDOS
Dünkel, F. (1982).	
Bayón, F. (1985).	
Bayón, F. y Compadre, A. (1991).	
Belfrage, H. (1991).	
Beljaars, I.C.M. y Berger, M.A. (1987).	
Berggren, O. y Svärd, H. (1990).	
Bishop, N., Sundin-Osborne, A. & Pettersson, T. N. (1987).	
*Boother, M. (1991).	
Bovens, R. (1987).	
Brown, B. (1985).	
Collins, S.A. y Tate, D.H. (1988).	
Cook, D.A., Fox, Ch.A., Weaver, C.M. & Rooth F.G. (1991).	2 estudios
Cooke, D.J. (1989).	
Cooke, D.J. (1991).	

Cullen, E. (1987).	4 estudios
*Daniel, C.J. (1987).	
*Daniel, C. y Dodd, C. (1989).	
Day, K. (1988).	
Fisher, M. (1991).	
Garrido, V., Redondo, S. y Pérez, E. (1989).	
Garrido, V. y Sanchís, J.R. (1990).	
*González Cieza, L. y Gutiérrez López, P. (1989, mayo).	
*Gustavsson, J. (1991).	
Hollin, C.R. y Courtney, S.A. (1983).	2 estudios
Hollin, C.R., Huff, G.J., Clarkson, F. & Edmondson, A.C. (1986).	2 estudios
Hopkins, R. (1991).	
Kravetz, Sh., Florian, V., y Nofer, E. (1990).	
Kruissink, M. (1990).	
Kury, H. (1989).	2 estudios
Legaz, F., López Pina, J.A., Sánchez-Meca, J., y Velandrino, A.P. (1990).	
López Pina, J.A., Sánchez-Meca, J., y Velandrino, A.P. (1992).	
Martín Aixelá, M. (1989).	
McDougall, C., Barnett, R.M., Ashurst B. y Willis, B. (1987).	
*McGurk, B.J. y Newell, T.C. (1981).	
*McMurrán, M. (1990).	
McMurrán, M. y Boyle, M. (1990).	2 estudios
Members of the Demonstration Unit, 1981-84, Inner London Probation Service (1986).	
Moreno, F.X. y Battestini, R. (1989).	
Petterson, T., Sundin-Osborne, A. y Bishop, N. (1986).	
Redondo, S. (1983).	
Redondo, S. (1984).	
Redondo, S., Roca, M., Pérez, E., Sánchez, A. y Deumal, E. (1990).	
Robertson, G. y Gunn, J. (1987).	
Rosner, A. (1988).	
Sánchez, M.A. y Polo, A. (1990).	
Sastriques, M. (1993).	
Scholte, E.M. y Smit, M. (1987).	
Shepherd, S. (1991),	
Singer, L.R. (1991).	
Slot, N.W. (1983).	
Slot, N.W. (1984, septiembre).	
Slot, N.W. y Bartels, A.A.J. (1983).	
Slot, N.W. y Heiner, J. (1986, mayo).	
Thornton, D. (1987).	
*Tomás Bardisa, J.L. (1990).	
van Dalen, W.E. (1989).	
Weaver, C. y Fox, Ch. (1984).	

NÚM. DE INFORMES CODIFICADOS=57
NÚM. DE INFORMES VÁLIDOS= 49

NÚM DE ESTUDIOS=57

13.2. Características del tratamiento

13.2.1. Características cualitativas

Los tipos de intervención o tratamientos fueron dimensionados, según hemos descrito con anterioridad, en tres diferentes agrupaciones relacionadas entre sí. De hecho, la segunda y la tercera agrupación se formularon mediante una transformación directa de la primera, atendiendo a una serie de criterios.

Nuestra primera agrupación de los tratamientos había tomado en consideración las dimensiones internamiento/no-internamiento de los sujetos, la amplitud de la acción emprendida y el tipo de acción o técnica (con referencia a un modelo teórico) predominante. En las categorías resultantes de ello, destaca numéricamente la mayor presencia de "terapias cognitivo-conductuales" (con un 33.33% de los estudios), seguidas de las "terapias psicológicas/psiquiátricas no conductuales" (el 15.8%) y de los "programas educativos" y "programas ambientales de contingencias" (con un 10.73% de los trabajos, respectivamente) (véanse Tabla 13.2 y Figura 13.1).

La segunda agrupación, producto de nuestro estudio piloto con 20 estudios, tiene un menor interés ya que constatamos que daba lugar a la reunión de tratamientos conceptualmente muy distintos, como es el caso de las "comunidades terapéuticas" y los "programas ambientales de contingencias". No obstante sus resultados se incluyen también en la Tabla 13.2.

Mucho mayor interés presenta, en cambio, la tercera agrupación de tratamientos: modelo teórico. Sobresale la frecuencia del "modelo cognitivo-conductual" (33.33% de los estudios), seguido del "modelo conductual" (19.30%) y de las "terapias no conductuales" (15.76%) (véanse Tabla 13.2 y Figura 13.2).

Tabla 13.2. Características cualitativas (sobre N estudios=57).

	Frec.	%
<u>Tipo de Tratamiento (Primera agrupación)</u>		
1- (TP) Terapia psicológica/psiquiátrica no conductual	9	15.8
2- (ED) Educación	6	10.53
3- (TC) Terapia de conducta	5	8.77
4- (TCC) Terapia cognitivo-conductual	19	33.33
5- (EE) Mero encarcelamiento/o endurecimiento del régimen de internamiento	1	1.75
6- (CT) Comunidad terapéutica/o suavizamiento del régimen de internamiento	5	8.77
7- (AC) Programa ambiental de contingencias de carácter globalizador	6	10.53
8- (DI) "Diversion"	5	8.77
9- (OT) Otras técnicas	1	1.75
<u>Tipo de Tratamiento (Segunda agrupación)</u>		
1-Terapias e intervenciones especializadas no conductuales sobre problemas concretos	9	15.79
2-Educación/formación	6	10.53
3-Terapias conductuales y cognitivo-conductuales	24	42.11
4-Cambios en el ambiente/régimen institucional como un todo	12	21.05
5-Alternativas al internamiento	5	8.77
6-Otras técnicas	1	1.75
<u>Modelo teórico</u>		
1-Terapia no conductual	9	15.79
2-Educativo-informativo	6	10.53
3-Conductual	11	19.30
4-Cognitivo-conductual	19	33.33
5-Teoría penal clásica	1	1.75
6-Comunidad terapéutica	5	8.77
7-"Diversion"	5	8.77
8-Otros	1	1.75
<u>Homogeneidad de la intervención</u>		
Sí	28	49.1
No	16	28.1
N/C	13	22.8

Figura 13.1. Tipología de los tratamientos.

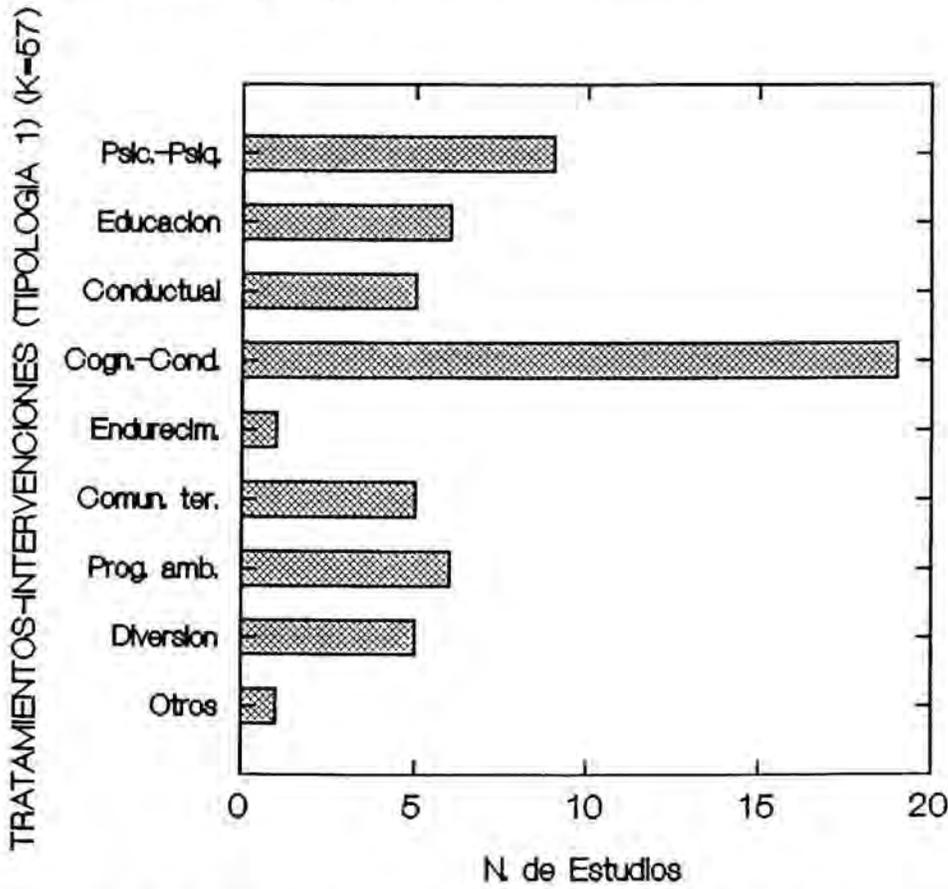
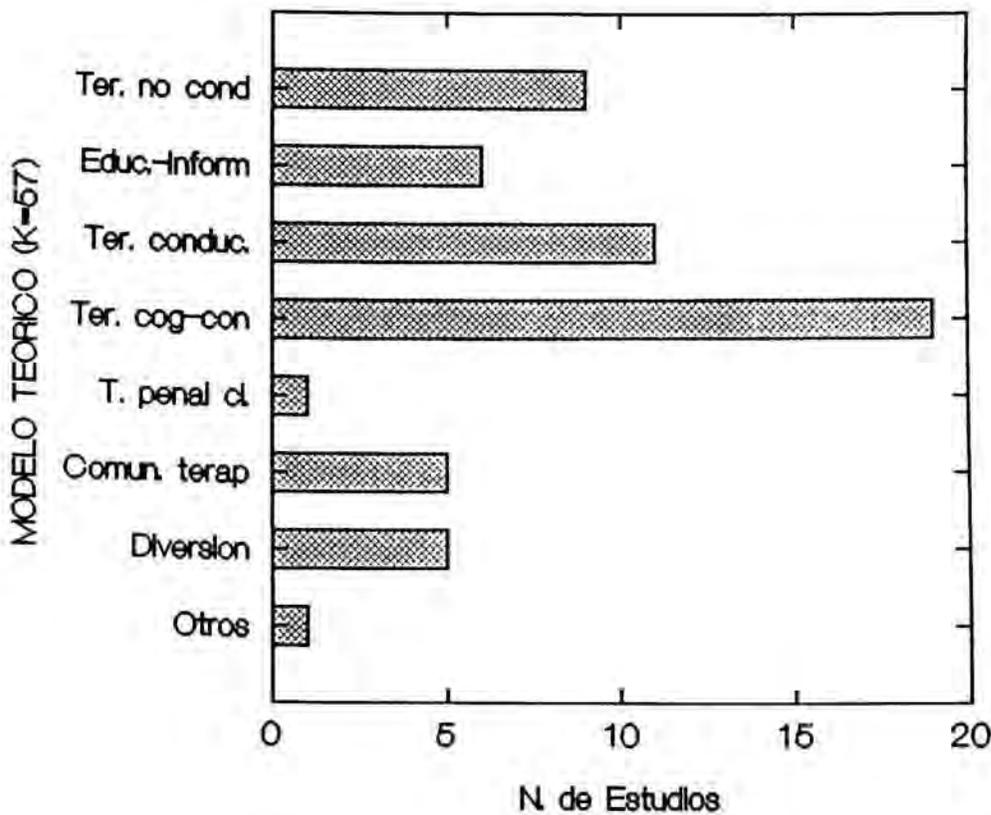


Figura 13.2. Tipología de los modelos teóricos de los tratamientos.



La variable homogeneidad de la intervención se refiere al grado en que el tratamiento fue administrado por igual -en intensidad, duración, etc.- a todos los sujetos participantes en un mismo programa. Casi un cincuenta por ciento de los programas analizados mostraron homogeneidad, mientras que en el otro cincuenta por ciento o bien no hubo homogeneidad interventiva o bien ésta no pudo ser determinada, por ausencia de datos.

13.2.2. Características cuantitativas

En este apartado se recogen una serie de variables cuantitativas relativas a la potencia interventiva que tuvieron los diversos programas analizados.

La duración promedio de las intervenciones por sujeto fue de 5.67 meses, aunque tal vez el valor más real, dada la gran variabilidad, nos lo de la mediana, que fue de 2.5 meses (véanse Tabla 13.3 y Figura 13.3). Complementariamente, la intensidad de las intervenciones, obtuvo un valor de mediana de 3.75 horas por semana y sujeto (Tabla 13.3). Finalmente, en lo tocante a la magnitud de las intervenciones el valor de mediana fue de 22.5 horas de tratamiento total por sujeto (véanse Tabla 13.3).

Por otro lado, se ponderó también la duración de los programas en conjunto -la duración de los proyectos evaluados, con independencia de los tiempos de estancia de los sujetos en los mismos-. El valor de la mediana, el más realista dada la gran variabilidad existente, fue de 14.73 meses (véanse Tabla 13.3).

Figura 13.3. Duración de la intervención en meses.

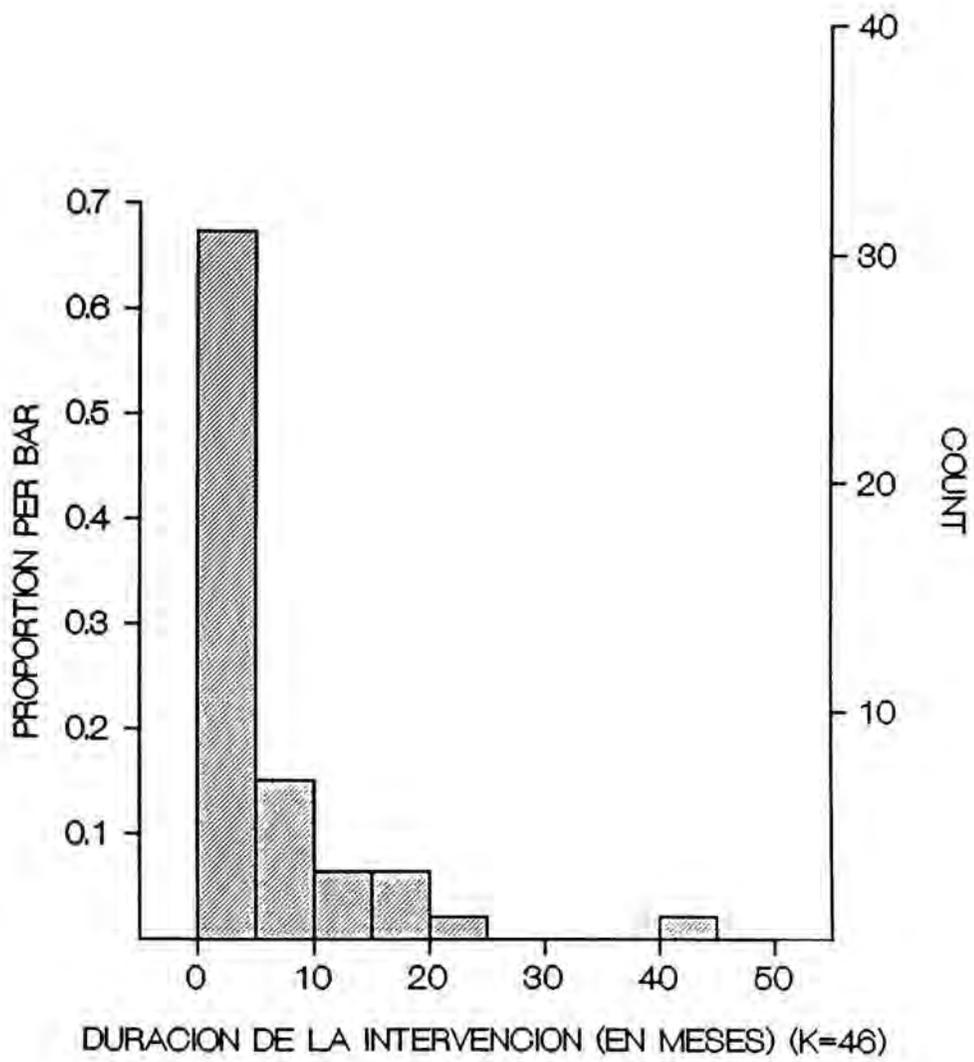


Tabla 13.3. Características cuantitativas.

	k	Mín	Máx	Me	s	Md
-Duración intervención por sujeto (en meses)	46	0.067	41.0	5.67	7.63	2.5
-Intensidad intervención (n horas/semana)	46	0.75	56.0	17.04	23.52	3.75
-Magnitud intervención (n total horas/sujeto)	45	1.5	9840.0	687.6	1691	22.5
-Duración Programa en conjunto (en meses)	52	1.7	276.0	31.23	47.73	14.73

13.3. Características de los sujetos

En términos de edad, la mayoría de los programas analizados se aplicaron a sujetos adultos, mayores de 21 años, en un 43.86% de los casos. Un volumen importante, del 31.48% de los programas, se dirigió a jóvenes delincuentes entre 16 y 21 años. La edad promedio de los estudios fue de 21.7 años. En lo relativo al sexo de los sujetos, la inmensa mayoría fueron programas dirigidos a hombres (el 93.51%). Pocos programas detallaron si sus sujetos tenían o no alguna alteración psicológica diagnosticada (sólo un 7.02%). Entre los sujetos tratados predominaron los delincuentes por delitos contra la propiedad (el 31.58% de los estudios) y los grupos mixtos, integrados por delincuentes de diversas tipologías delictivas (en un 28.07 de los trabajos (véanse Tablas 13.4 y 13.5, y Figuras 13.4, 13.5 y 13.6).

Tabla 13.4. Características cualitativas (sobre N estudios = 57).

	Frec.	%
<u>Tipo de sujetos (por edad)</u>		
0-Adolescentes (pre-delincuentes)	9	15.79
1-Jóvenes (16-21 años)	18	31.58
2-Adultos (>21 años)	25	43.86
3-Mixto (sujetos varias edades)	5	8.77
<u>Alteración psicológica diagnosticada</u>		
Sí	4	7.02
No	53	92.98
<u>Tipo de delito predominante (>60%)</u>		
1-Contra la propiedad	18	31.58
2-Contra las personas	2	3.51
3-Delitos sexuales	4	7.02
4-Tráfico de drogas	2	3.51
5-Relacionados consumo alcohol	7	12.28
6-Otros	0	0.00
7-Mixto	16	28.07
N/C	8	14.04

Tabla 13.5. Características cuantitativas.

	k	Mín	Máx	Me	s	Md
-Sexo (% hombres)	52	0	100	93.51	20.86	100
-Edad (años)	39	12.5	39.3	21.7	6.9	19.0

Figura 13.4. Tipología de los sujetos en función de la edad.

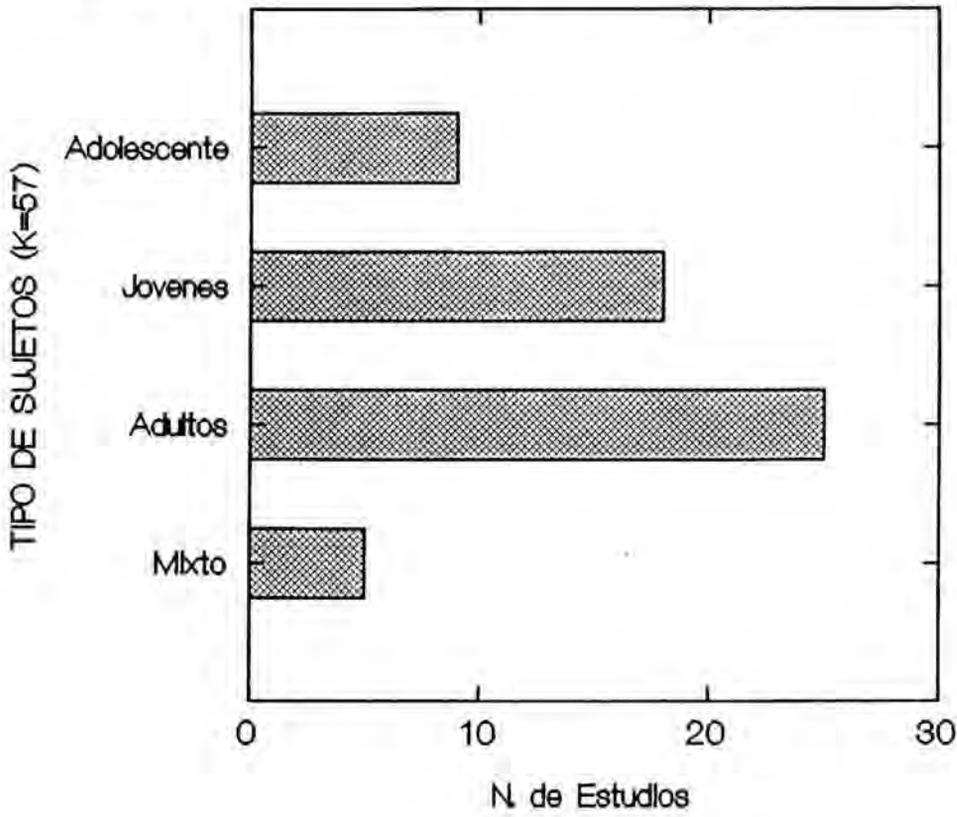


Figura 13.5. Edad media de los sujetos.

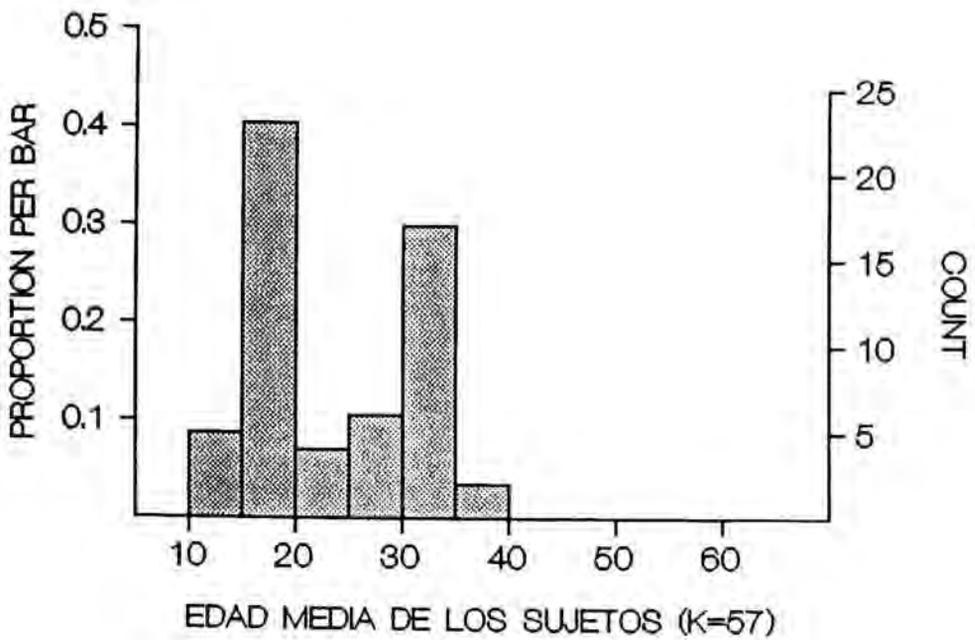
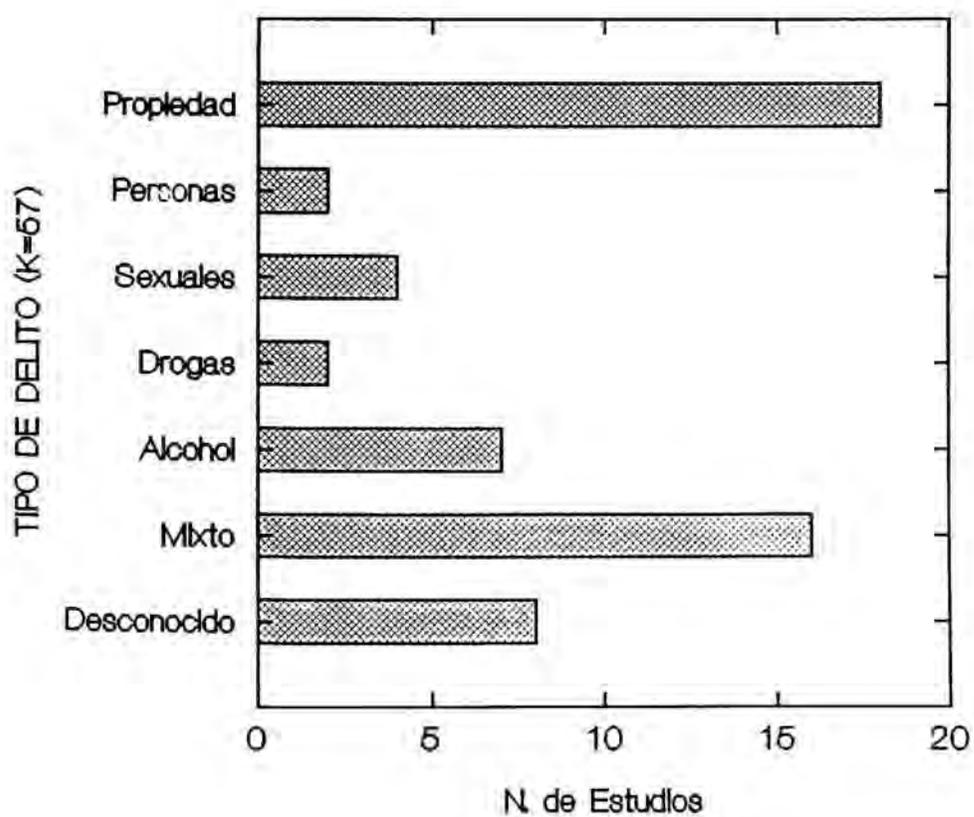


Figura 13.6. Tipología delictiva de los sujetos.



13.4. Características del ambiente/contexto de la intervención

Si analizamos los trabajos en función del lugar de la intervención, la mayor proporción de estudios tuvo lugar en prisiones de adultos (el 29.82%), prisiones de jóvenes (el 28.07%), y en la comunidad (el 24.56%). En correspondencia con ello, prioritariamente acontecieron en contextos institucionales cerrados (en un 59.65% de los casos) (véanse Tabla 13.6 y Figuras 13.7 y 13.8).

El país de donde hemos recogido más trabajos ha sido Gran Bretaña (con 26), seguido de España (con 13), de Holanda (con 9), de Alemana y Suecia (con 4, respectivamente), y de Israel (1 trabajo) (Figura 13.9).

Se analizó la variable disciplina de los autores del estudio o documento -en concreto, del primer autor del estudio-. En algo más de una cuarta parte de los trabajos no pudo extraerse esta información. Entre los estudios que aportaron este dato destaca claramente la presencia de psicólogos (el 57.89% de los primeros autores), seguidos a gran distancia de otros profesionales, como psiquiatras, profesionales de la educación, oficiales de "probation", y trabajadores sociales (Figura 13.10). En lo concerniente a las fechas de publicación o emisión del documento -téngase en cuenta que hay documentos no publicados- sobresale el número de trabajos correspondientes a 1.987 (con un 19.30% del total) y años siguientes, mientras que se recogieron menos documentos datados entre 1982-1986 (Figura 13.11).

Un apartado, sin duda, muy interesante lo constituyente el número y la cualificación profesional de quienes aplicaron los programas. En lo relativo al número de profesionales implicados, muchos trabajos no informaron con suficiente claridad sobre ello. Se observa, por lo demás, una gran dispersión en el número de operadores -que va desde 1 a más de 15-, sin que destaque ningún número específico. En lo relativo a las profesiones/formación de los operadores de los programas, sobresalen nuevamente los psicólogos, que aparecen mencionados en 30 estudios, seguidos de los funcionarios de prisiones, con 15 referencias. A cierta distancia se muestran psiquiatras, educadores, trabajadores sociales, agentes de "probation", etc. Debe mencionarse la presencia explícita de policías en 4 programas.

Tabla 13.6. Características cualitativas (sobre N estudios = 57).

	F r e c .	%
<u>Lugar de aplicación de la intervención</u>		
1-Centro de reforma juvenil	3	5.26
2-Prisión de jóvenes	16	28.07
3-Prisión de adultos	17	29.82
4-Comunidad (probation, grupo terapia...)	14	24.56
5-Otros	4	7.02
N/C	3	5.26
<u>Régimen de la intervención</u>		
1-Cerrado	34	59.65
2-Semi-abierto	2	3.51
3-Abierto	11	19.30
4-Otros	7	12.28
N/C	3	5.26
<u>País de la intervención</u>		
1-Gran Bretaña	26	45.61
2-España	13	22.81
3-Alemania	4	7.02
4-Holanda	9	15.79
5-Suecia	4	7.02
6-Israel	1	1.75
<u>Disciplina de los autores del estudio (primer autor)</u>		
1-Psicología	33	57.89
2-Educación	2	3.51
3-Psiquiatría	3	5.26
4-"Probation officer"	2	3.51
5-Trabajo social	1	1.75
N/C	16	28.07
<u>Fecha de publicación del estudio</u>		
1.982	1	1.75
1.983	5	8.77
1.984	3	5.26
1.985	2	3.51
1.986	5	8.77
1.987	11	19.30
1.988	3	5.26
1.989	8	14.04
1.990	8	14.04
1.991	9	15.79
1.992	1	1.75
1.993	1	1.75
Total	57	100.00

Número de profesionales participantes en la aplicación del programa

1	4	6.06
2	9	13.63
3	7	10.60
4	8	12.12
5	4	6.06
8-15	7	10.60
más de 15	4	6.06
N/C	23	34.84
TOTAL		66 100

Profesiones/formación de quienes que aplican los programas¹²

1-Psicólogos	30	28.57
2-Funcionario penitenciarios	15	14.28
3-Psiquiatras	8	7.61
4-Educadores	8	7.61
5-Trabajadores sociales	8	7.61
6-Profesores enseñanza	6	5.71
7-Agentes de <u>probation</u>	5	4.76
8-Personal sanitario	4	3.80
9-Policías	4	3.80
10-Otros (personal administrativo pedagogos, directivos prisiones, representantes fiscalía y tribunales)	10	9.52
TOTAL REFERENCIAS PROFESIONALES	105	100

¹²La variable profesión/formación se ha ponderado tomando en consideración si una determinada profesión aparece/no aparece en cada estudio, con independencia del número de profesionales de ella que sean mencionados. Así pues, esta tabla contestaría a la pregunta ¿cuántos estudios mencionan de modo explícito la participación de cada categoría profesional?

Figura 13.7. Lugares en que se llevaron a cabo las intervenciones.

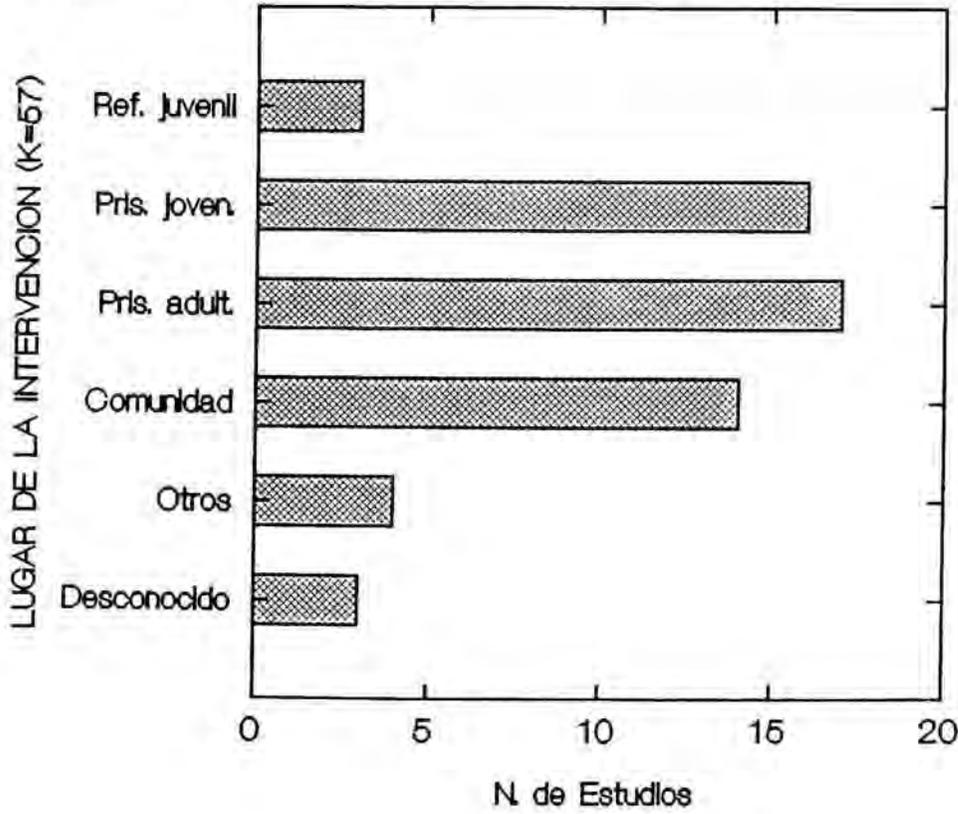


Figura 13.8. Régimen de vida en que se aplicaron las intervenciones.

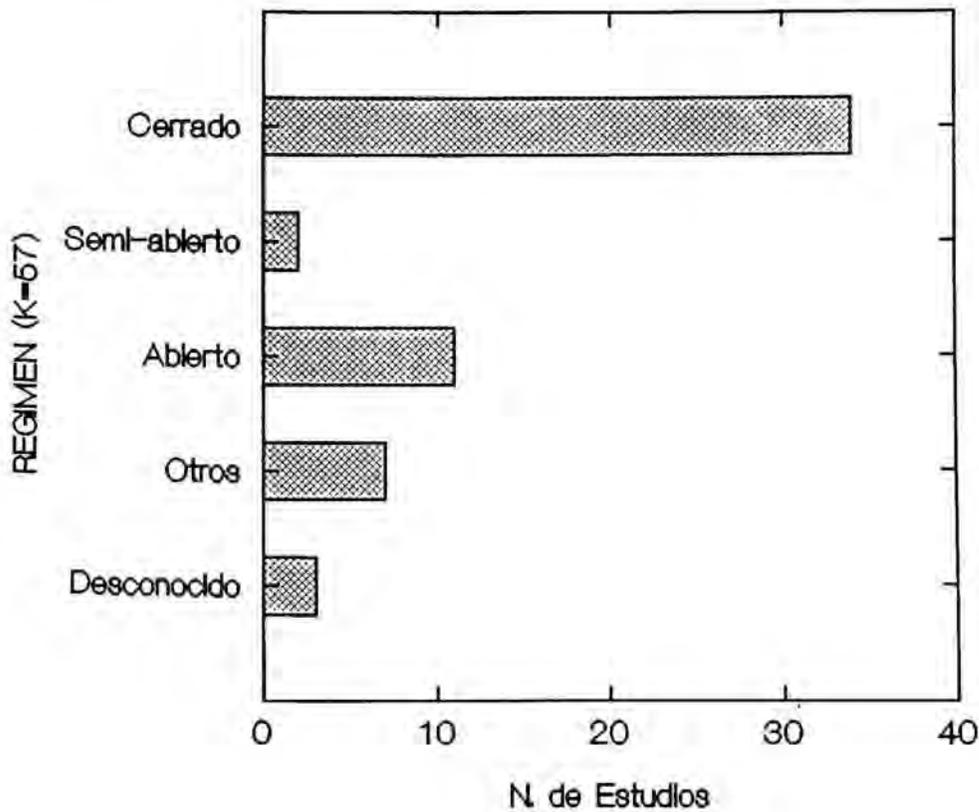


Figura 13.9. País en que se realizaron las intervenciones.

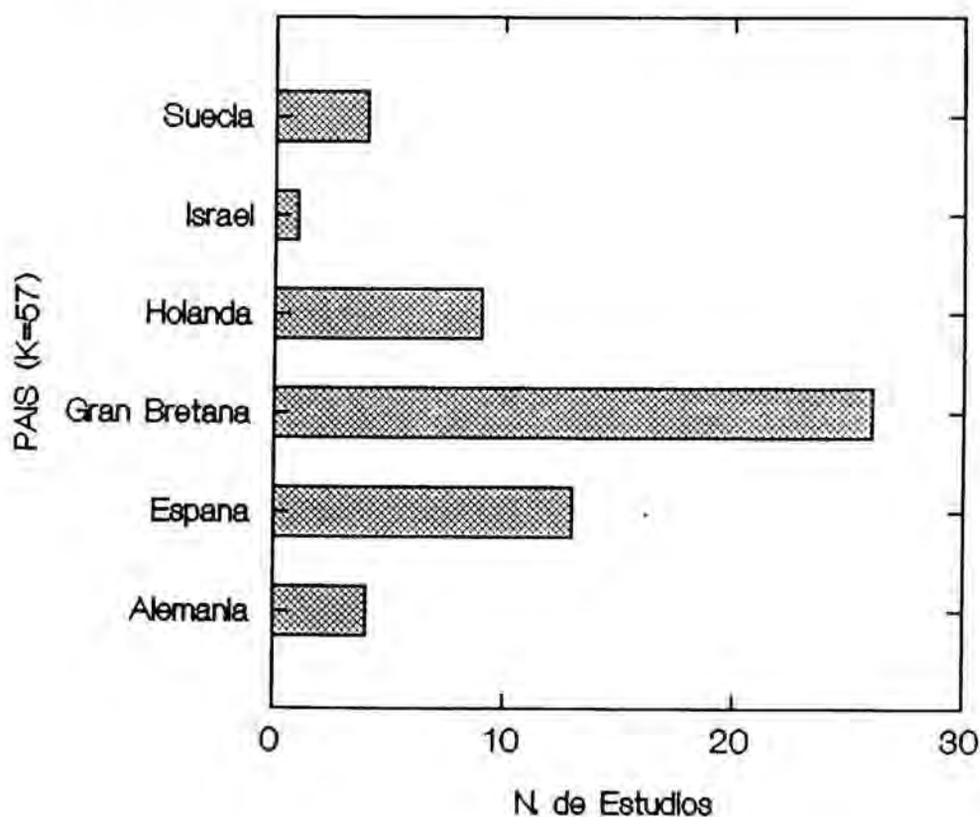


Figura 13.10. Disciplina del primer autor de las intervenciones.

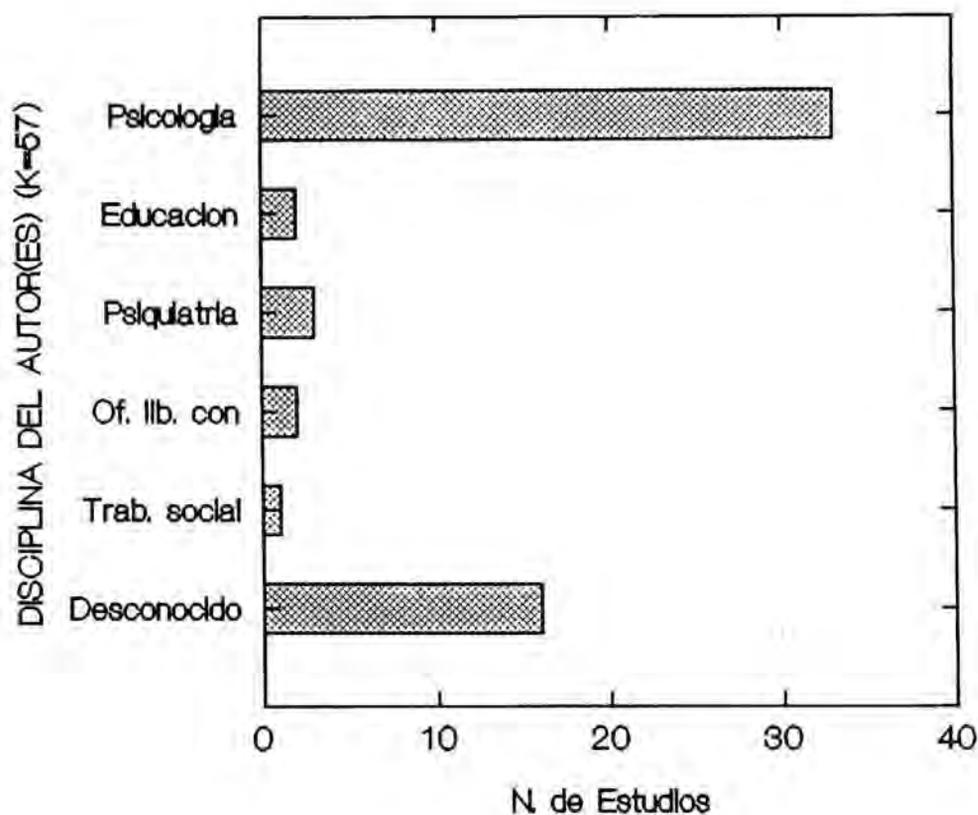
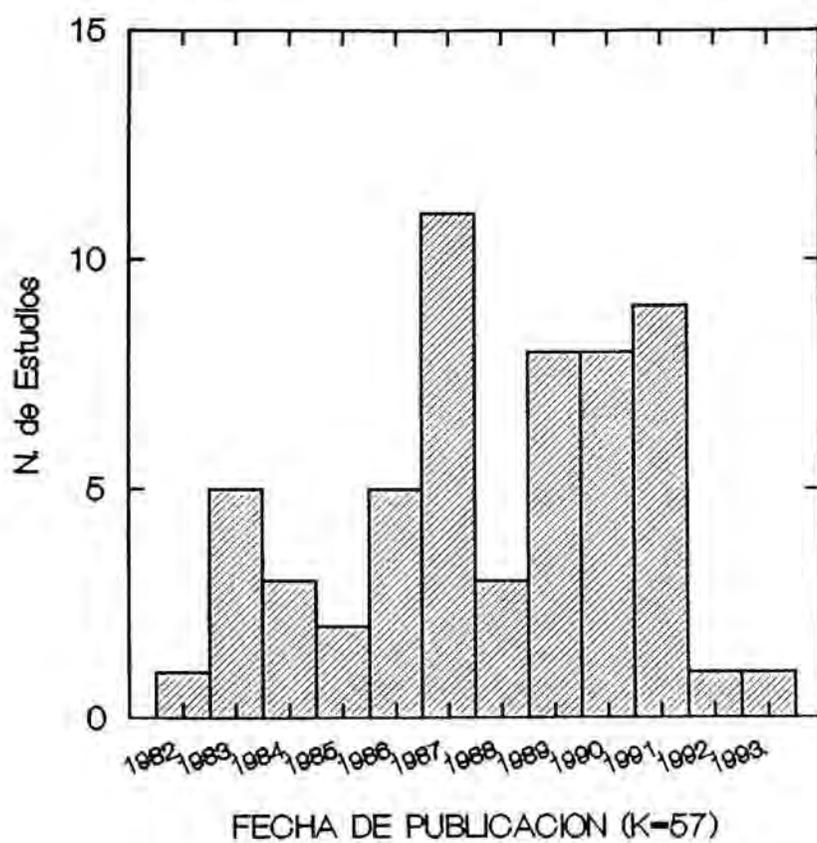


Figura 13.11. Fecha de publicación de los estudios.



13.5. Características metodológicas

13.5.1. Características cualitativas

Desde una perspectiva metodológica clásica, destaca la presencia de diseños de investigación cuasi-experimentales (el 43.86%) y pre-experimentales (el 21.05%), considerados globalmente, sobre los experimentales (que sólo suponen el 26.32%). Los diseños conductuales -de línea base múltiple o regresión- sólo aparecen en 3 casos, y los que hemos denominado de comparación de instituciones en 2 (véanse Tabla 13.7 y Figura 13.12). Pese a ello, la mayoría de los estudios utilizan diseños inter, es decir incluyen algún grupo de comparación, aunque no sea experimentalmente equivalente (ello ocurre en el 68% de los casos).

Se establecieron una serie de criterios de ponderación de la calidad de los diseños: tamaño muestral grande, asignación aleatoria, mortalidad experimental pequeña, empleo de alguna variable dependiente normalizada, de grupo control, medición de resultados sobre todas las variables dependientes preestablecidas, y existencia de medidas pre-test. De estos criterios el más atendido en los estudios fue la utilización de alguna VD normalizada, objetiva o conductual ciega (en un 82.4% de los casos); por el contrario, el requisito menos presente fue la asignación aleatoria de los sujetos, que sólo se dio en un 29.8% de los trabajos. Los antedichos criterios, ponderados en forma dicotómica, dieron lugar a la confección de una escala de 0-7 puntos de calidad global de los diseños. Esta escala parece, en principio, adecuada para valorar la calidad de la muestra de diseños, a la vista de la distribución de estudios en ella. Como puede verse, esta escala -de 0 a 7 puntos- acoge la mayor parte de diseños en su parte central -de 3 a 4 puntos-, disminuyendo el número de estudios en los extremos. En suma, parece ajustarse de modo adecuado a una distribución normal (véanse Tabla 13.7 y Figura 13.13).

Tabla 13.7. Características cualitativas (sobre N estudios = 57).

	Frec.	%
<u>Tipo de diseño</u>		
1-Diseño Pre-Experimental	12	21.05
2-Diseño Cuasi-Experimental	25	43.86
3-Diseño Experimental	15	26.32
4-Diseño Conductual	3	5.26
5-Diseño de Comparación de Instituciones	2	3.51
<u>Diseño inter-intra</u>		
Inter	39	68.42
Intra	18	31.58
<u>Calidad del diseño</u>		
1-Tamaño muestral > 30	30	52.6
2-Asignación aleatoria a los grupo	17	29.8
3-Mortalidad exptal total < 20	17	29.8
4-Utiliza alguna VD normalizada	47	82.4
5-Utiliza grupo de control ¹³	38	66.7

¹³En teoría el número de estudios catalogados como utiliza grupo control (n=38) y los catalogados como diseño inter (n=39) deberían coincidir. Sin embargo, en uno y otro caso se han utilizado criterios distintos para la codificación, atendiendo a intereses más generales. Por un lado, para dicotomizar los diseños en inter/intra se ha considerado la mera presencia o no de algún tipo de grupo control o de comparación. Sin embargo, a la hora de ponderar la "calidad del diseño" se ha sido estricto y amplio a la vez al puntuar el ítem utiliza grupo de control bajo los siguientes parámetros: para los diseños tradicionales (experimentales), se ha sido estricto, puntuando positivamente sólo a aquellos estudios que habían incluido un grupo de control "a priori", y no a posteriori, formándolo a base de los sujetos en lista de espera, bajas, etc.; por otra parte, se ha puntuado positivamente en este ítem a los diseños conductuales que pese a carecer de grupo de control, atendían, en cambio, al requisito metodológico conductual de poseer

6- Informa resultados sobre todas las Vds prefijadas	35	61.4
7-Toma medidas pretest	34	59.6

Calidad global del diseño (suma puntuaciones de calidad: 0-7 puntos)

0.	0	0.00
1.	2	3.51
2.	5	8.77
3.	17	29.82
4.	18	31.58
5.	8	14.04
6.	7	12.28
7.	0	0.00

Criterio de selección de la muestra

1-Todos los sujetos de una unidad	7	12.28
2-Por previsión de estancia suficiente para completar programa	5	8.77
3-Sujetos de ciertas tipologías delictivas	8	14.04
4-Sujetos violentos en la institución	9	15.79
5-Sujetos con necesidades específicas	28	49.12

Tipo de Publicación

1-Artículo revista	27	47.37
2-Capítulo de libro	14	24.56
3-Libro/monografía	1	1.75
4-Comunicación	5	8.77
5-Tesis doctoral	0	0.00
6-Manuscrito no publicado	10	17.54

reversión o línea base múltiple. Finalmente, debe también tenerse en cuenta que en algunos casos nuestro análisis nos ha llevado a tomar en consideración el mismo grupo de control para estudiar más de un programa de tratamiento. Este es el caso cuando un trabajo presenta información sobre dos tratamientos distintos con un solo grupo de control: dentro de nuestro análisis aparecerán dos estudios: uno, en el que se compara el primer grupo tratado con el único grupo de control; y otro, en el que se compara el segundo grupo tratado con el mismo único grupo de control.

Fuente de publicación

Publicado	42	73.68
No publicado	15	26.32

Idioma del documento (sobre N estudios = 65)

1-Español	14	21.53
2-Inglés	48	73.84
3-Catalán	2	3.07
4-Sueco	1	1.53

Fuente de localización de los documentos (sobre N estudios = 66)

1-Posesión previa (en revistas y libros propios u obtenido de los autores en Congresos, etc.)	30	45.45
2-Referencia obtenida de consulta bases documentales: Dialog System, Pascal System y Psychological Abstracts	14	21.21
3-Remitido por los autores ante mi comunicación por correo	3	4.54
4-Referencia facilitada por Centros de Documentación extranjeros ante mi comunicación por correo	9	13.63
5-Biblioteca Centre d'Estudis Jurídics (por revisión directa de revistas y libros)	10	15.15

Figura 13.12. Tipología de los diseños de investigación.

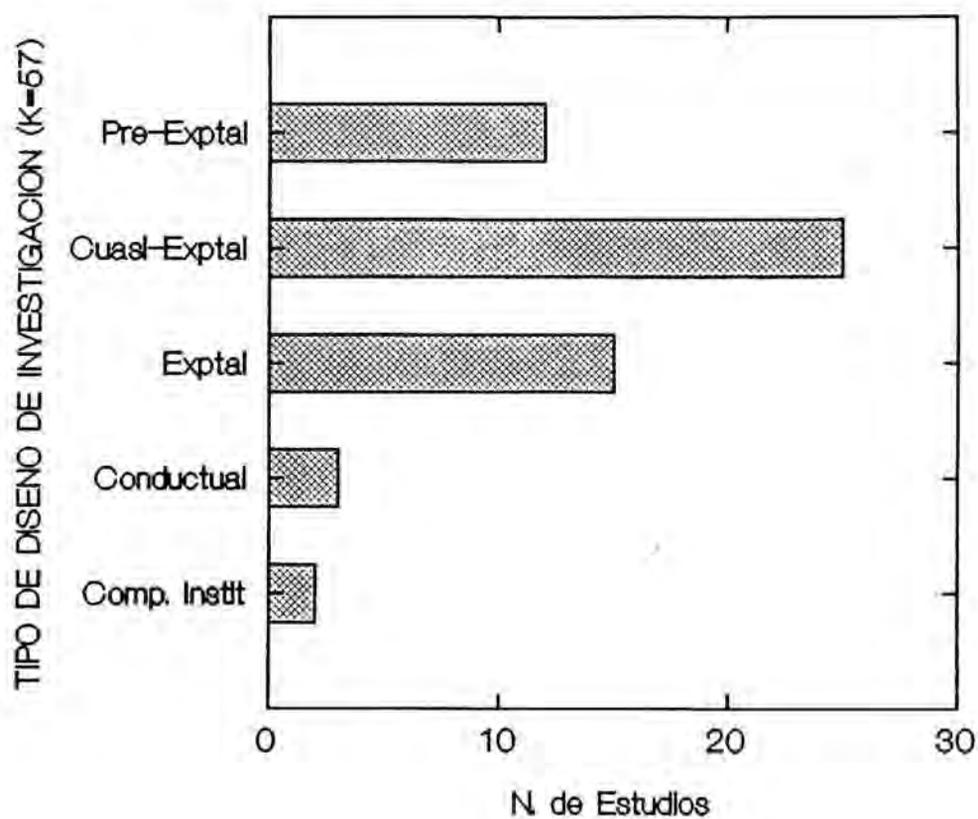
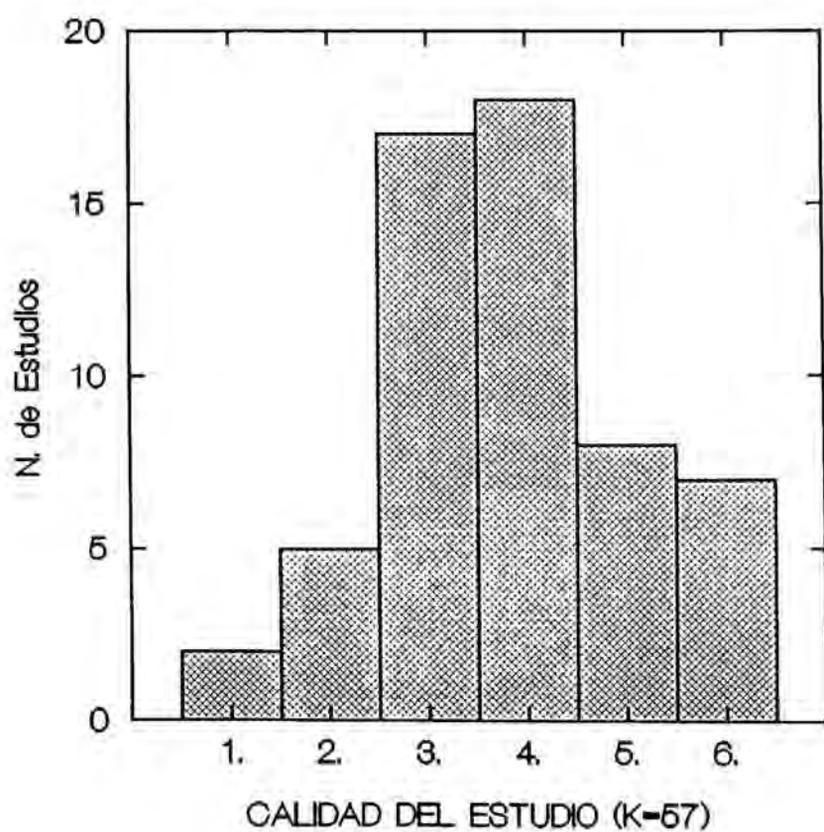


Figura 13.13. Calidad de los estudios.



Se analizó también cuáles habían sido los criterios de selección de las muestras de los estudios. En la mayoría de los casos se trataba de sujetos elegidos para el programa por sus específicas necesidades de tratamiento (un 49%), repartiéndose el resto de estudios entre las restantes categorías (todos los sujetos de una determinada unidad o centro, que existiera una cierta previsión de permanencia suficiente para completar el programa, sujetos de determinadas tipologías delictivas, o sujetos violentos en las instituciones) (véanse Tabla 13.7 y Figura 13.14).

Con objeto de poder controlar un eventual sesgo de publicación en nuestros resultados, se tomó también en consideración la modalidad de publicación o divulgación del estudio. La mayoría de los estudios habían sido publicados en revistas (el 47.37%) o libros (el 24.56%), aunque una significativa proporción de documentos era material no publicado (comunicaciones y manuscritos) (el 26.32%) (véase Figura 13.15).

Figura 13.14. Criterio de selección de los sujetos para los programas.

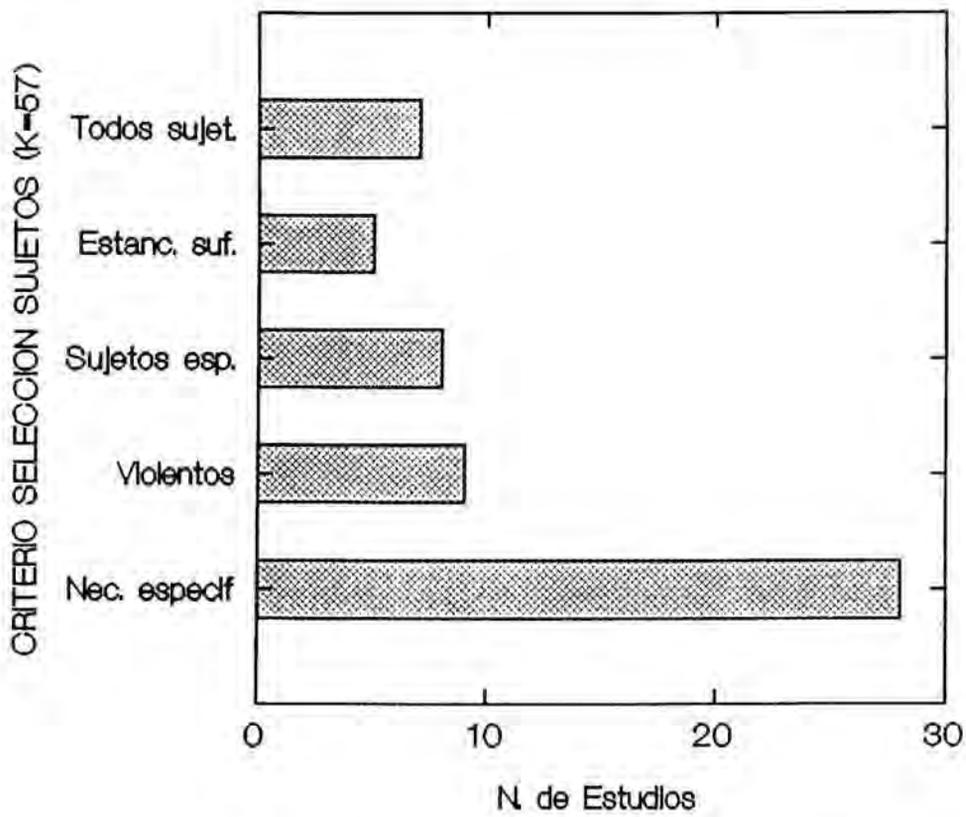
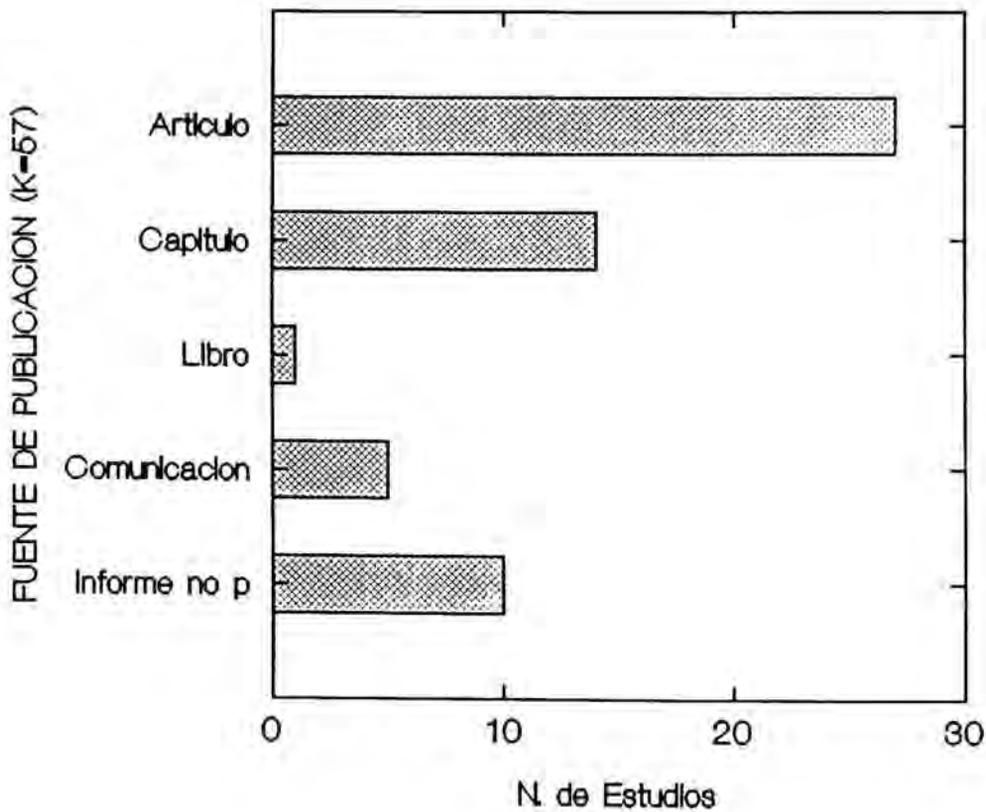


Figura 13.15. Fuente de publicación de los estudios.



En relación con el idioma, la mayoría de nuestros materiales son trabajos en inglés (el 73.68%) y español (el 21.53%). También contamos con 2 documentos en catalán y 1 en sueco, que incluía un resumen en inglés.

En los comienzos de esta investigación se dedicó un esfuerzo considerable, durante más de un año, al proceso de localización de trabajos. Los resultados consignados en fuentes de localización de los documentos cuestionan en parte la utilidad de esta laboriosa tarea de búsqueda y localización de materiales nuevos. Como puede verse, finalmente casi la mitad de los documentos analizados ya se encontraban en poder del autor desde el principio (30 estudios) o fueron localizados por revisión directa de materiales de biblioteca (10 estudios). Otras fuentes, de gran laboriosidad -envío de cartas a investigadores y a centros de investigación internacionales, y búsquedas en bases documentales-, no resultaron tan efectivas como el autor en principio esperaba.

13.5.2. Características cuantitativas

De gran importancia resulta el conocimiento de los tamaños muestrales. La variabilidad es grande en este punto. Una vez descartados los estudios de caso único, los tamaños muestrales de los grupos experimentales oscilan entre 4 - 1.000 sujetos y los de los grupos controles entre 5-889. Habida cuenta de esta gran variabilidad, probablemente las medianas, 25 y 19.5, son las puntuaciones que mejor representan los tamaños muestrales promedio de los grupos experimentales y controles, respectivamente. Los tamaños muestrales totales fueron los siguientes: 4.407 sujetos fueron tratados, en total, en los 57 programas analizados, mientras que 1.989 formaron parte de los grupos de control, en los 32 estudios en que hubo tal grupo de control (Tabla 13.8).

En 41 estudios (el 71.9% del total) se produjo una mortalidad promedio de 20 sujetos (valor de mediana) en el grupo experimental. Mientras que en 25 estudios (el 78.1% de los que poseían grupo de control) los grupos controles

experimentaron alguna merma, difícilmente representable por la media, dada la gran dispersión muestral existente, ni, en este caso, por la mediana, que es 0 (Tabla 13.8 y Figuras 13.16 y 13.17).

En 43 estudios hubo seguimiento, cuya duración mediana fue 12 meses (rango 0.5-120 meses) (Tabla 13.8 y Figura 13.18).

Tabla 13.8. Características cuantitativas.

	k	Mín	Máx	Me	s	Md	Total
<u>Tamaño muestral</u>							
Tamaño muestra total	57	6	1212	112.5	213.2	36	6414 ¹⁴
Tamaño muestral G.E.	57	4	1000	77.3	153.3	25	4407
Tamaño muestral G.C.	32	5	889	62.1	154.7	19.5	1989
<u>Mortalidad experimental</u>							
-En el G.E.	41	0	82.1	23.7	22.3	20	
-En el G.C.	25	0	71.4	13.8	19.9	0	
Duración seguimiento (en meses)	43	0.5	120	19.7	23.2	12	
Potencia estadística	28	0.13	0.999	0.6	0.266	0.589	
<u>Tiempo de lectura, vaciado y codificación de los estudios (sin considerar la codifi- cación de resultados) (en minutos)¹⁵</u>							
	65	50	285	134(2,23 h)		$\frac{T.Total}{8734'}$	(145,5h)

¹⁴Los tamaños muestrales corresponden al post-test, una vez eliminada la mortalidad experimental. Si sumamos el tamaño muestral total del G.E. y el del G.C. obtenemos un tamaño muestral conjunto total de 6.396. Sin embargo el tamaño muestral conjunto, en datos reales, es de 6.414 sujetos, algo superior. Esta discrepancia es debida al hecho de que tanto en el G.E. como en el G.C. no siempre todos los sujetos son evaluados en todas las variables consideradas. Por ello, en estas situaciones ha sido necesario promediar tamaños muestrales, lo que introduce estas pequeñas diferencias numéricas.

¹⁵Ya hemos comentado que el tiempo total empleado para la codificación es muy superior al aquí consignado. En total, el proceso codificador no supuso un tiempo menor a las 362 horas.

Figura 13.16. Mortalidad experimental en los grupos tratados.

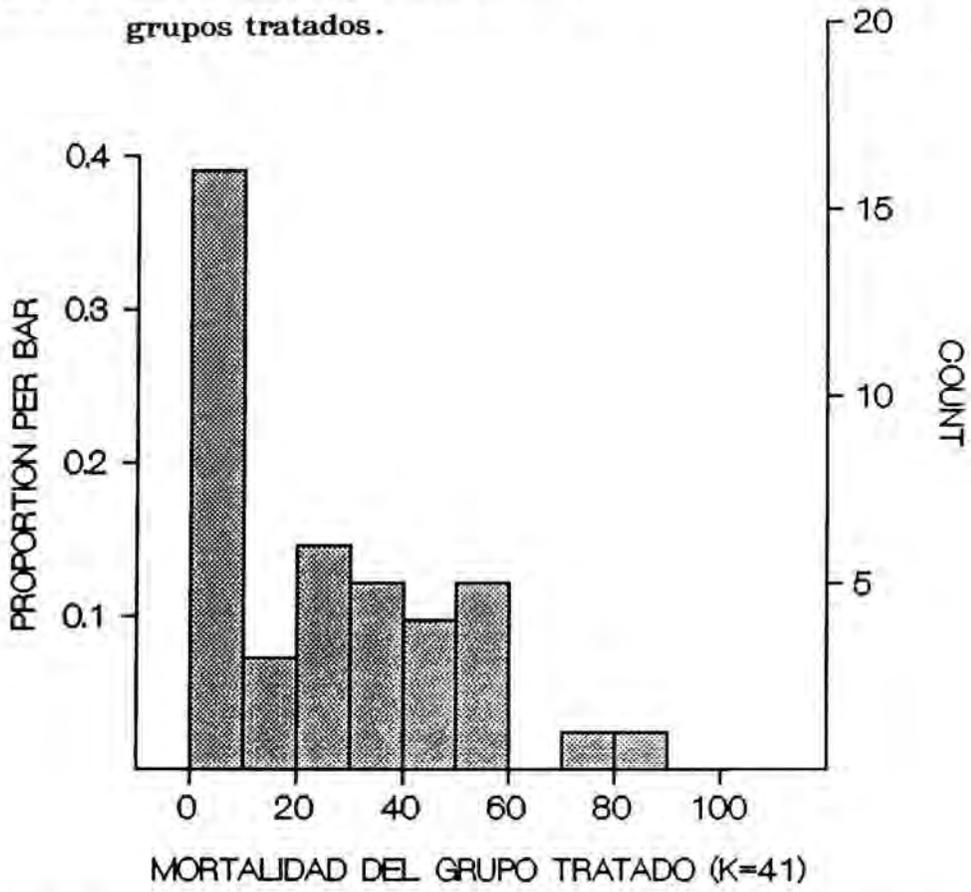


Figura 13.17. Mortalidad experimental en los grupos de control.

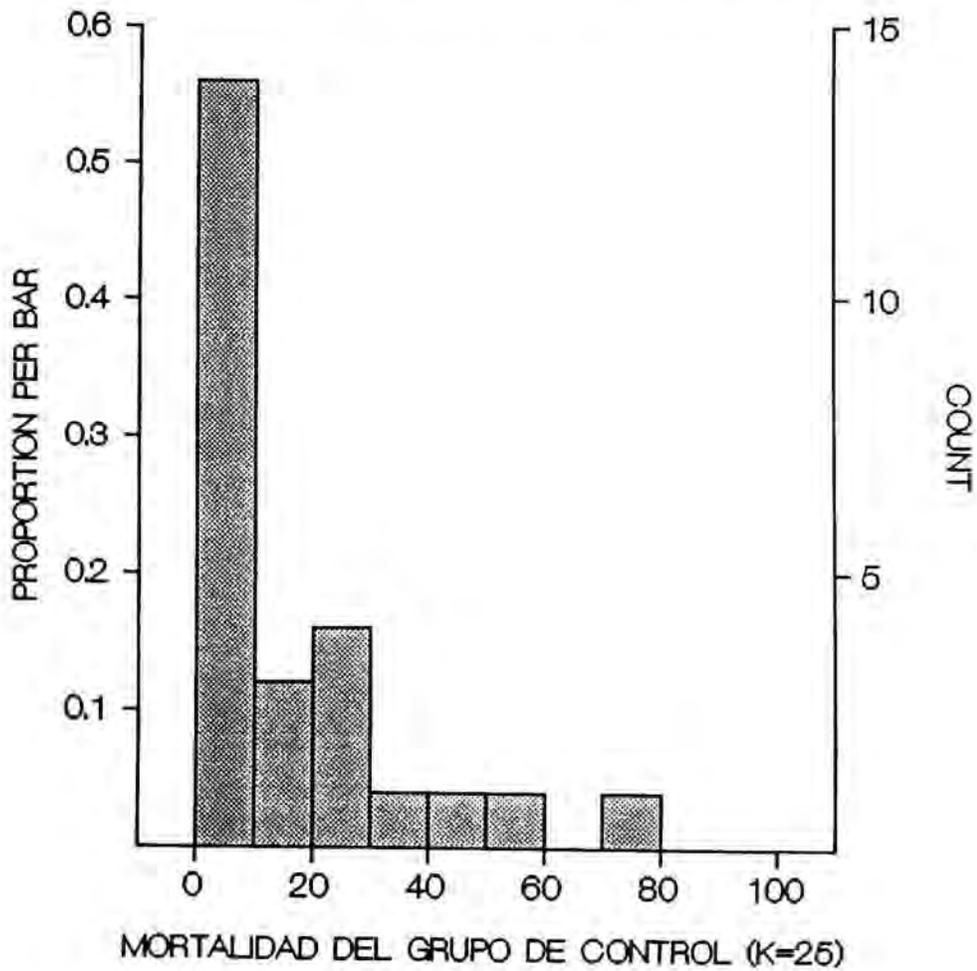


Figura 13.18. Duración del periodo de seguimiento.

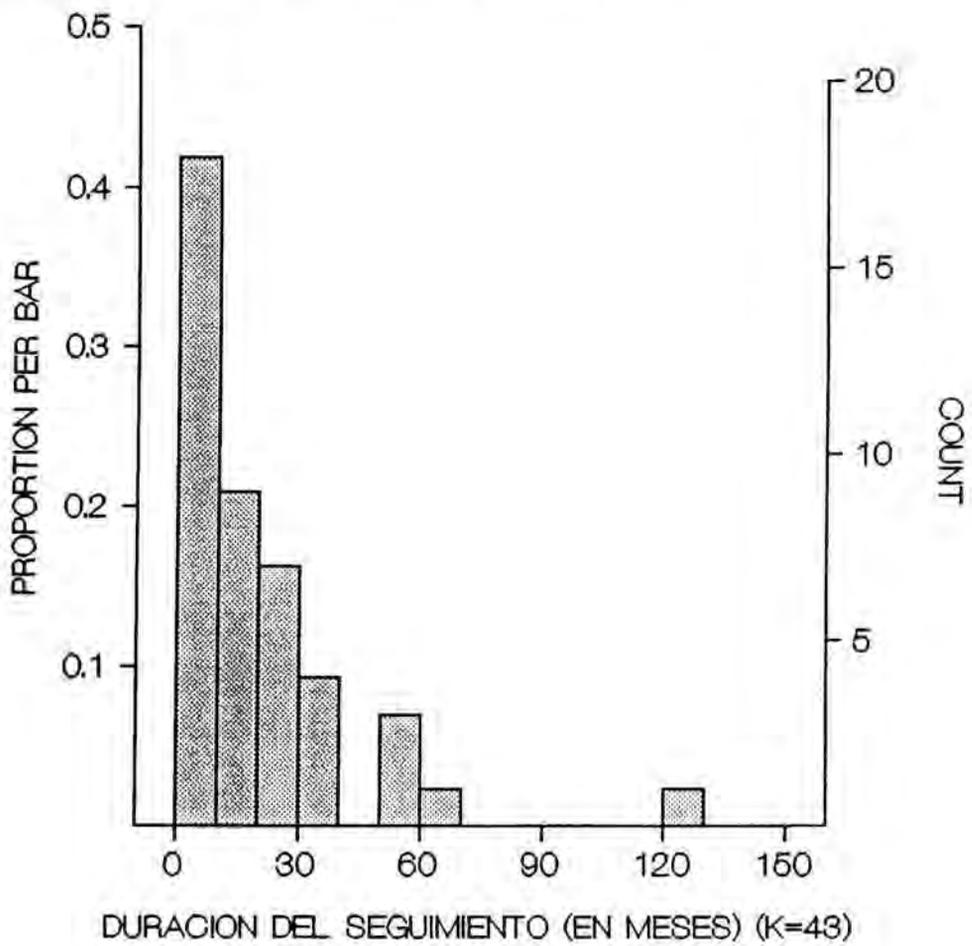
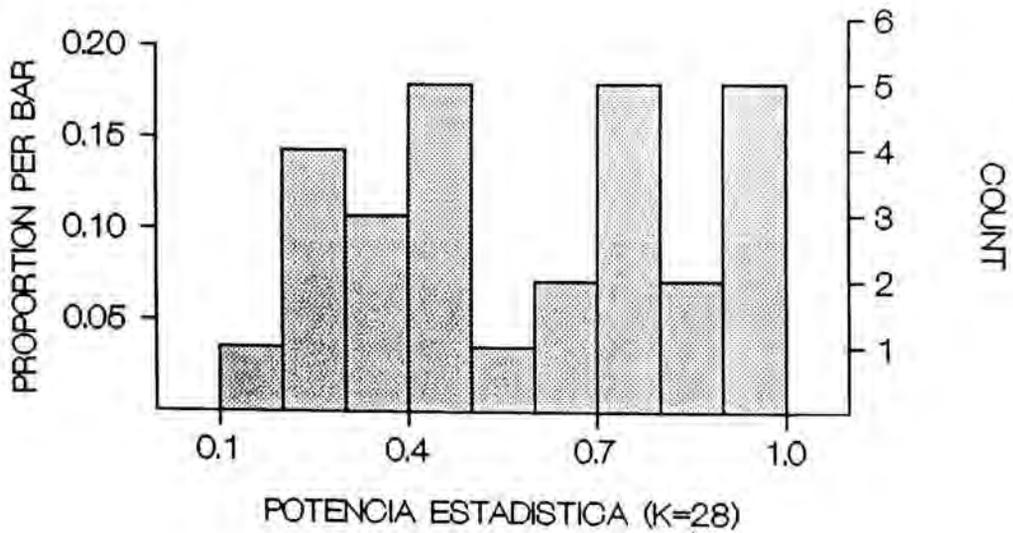


Figura 13.19. Potencia estadística de los estudios.



Hemos considerado importante el análisis de la potencia estadística. La potencia estadística es un parámetro que pondera la probabilidad existente de poder afirmar la hipótesis alternativa; esto es, de poder encontrar, si existen, unos resultados positivos -de afirmar que el tratamiento fue efectivo-. La potencia estadística dependerá básicamente del tamaño de los resultados obtenidos -en nuestro caso del tamaño del efecto- y del tamaño muestral. Su cálculo sólo adquiere sentido para aquellos estudios que han utilizado pruebas inferenciales (sólo 28 estudios). En un sugerente símil debido al Dr. Julio Sánchez Meca (Sánchez-Meca, 1993a), la potencia estadística sería a la obtención de los resultados "verdaderos" de una investigación lo que la luminosidad existente en una habitación es a la localización de un objeto extraviado dentro de ella. Nuestras posibilidades de encontrarlo aumentan, sin duda, a medida que el objeto sea más grande, pero también a medida que aumente la iluminación existente -la potencia estadística en nuestro caso-. Para un N de 28 estudios hemos obtenido una potencia estadística promedio de 0.6 (Md = 0.59), que se sitúa por debajo del desideratum estadístico de 0.80 propuesto por Cohen, pero algo por encima del promedio de estudios psicológicos en general, que ha sido cifrado entre el 0.47 y 0.67, según diversos análisis de potencia (Sánchez-Meca, Valera, Velandrino y Marín, 1992; Valera, Sánchez-Meca, Velandrino y Marín, 1993, Julio) (véanse Tabla 13.8 y Figura 13.19).

RESUMEN

Para comenzar, se han descrito las principales características de los estudios analizados, en cuatro grandes bloques: las peculiaridades de los tratamientos aplicados, los sujetos tratados, los ambientes o contextos de las intervenciones, y, finalmente, las circunstancias metodológicas de los trabajos.

En relación con los tratamientos, predominaron las "terapias cognitivo-conductuales" y las intervenciones tuvieron una duración promedio de 2,5 meses. Los sujetos tratados fueron mayoritariamente varones adultos -aunque un tercio de los programas se dirigieron a jóvenes entre 16 y 21 años- que habían cometido delitos contra la propiedad. La mayoría de las intervenciones habían sido realizadas en prisiones, ya sea de adultos o de

jóvenes, siendo sus más frecuentes operadores y evaluadores los psicólogos.

Por último, en lo tocante a la metodología de los estudios analizados, predominaron los diseños evaluativos cuasi-experimentales con grupos de comparación, de una calidad metodológica razonable. Destacaron los documentos publicados en inglés, aunque también se habían recogido trabajos no publicados. Los tamaños muestrales de los grupos experimentales y controles tuvieron un promedio de 25 y 19.5 sujetos, respectivamente. Se obtuvo una potencia estadística promedio de 0.6, que responde a los parámetros habituales de la investigación psicológica.

14. ANÁLISIS DE LOS EFECTOS: META-ANÁLISIS GLOBAL

14.1. Análisis del sesgo de publicación

Se ha mencionado que una de las más frecuentes fuentes de sesgo que acechan en un estudio meta-analítico es el sesgo de publicación. Esto es, la posibilidad existente de que si sólo, o preferentemente, hemos recogido trabajos que fueron publicados, éstos nos informen de resultados más favorables que los que realmente se obtendrían si el muestreo de estudios fuera más exhaustivo e incluyera, en semejante proporción, trabajos no publicados. Se ha argumentado la mayor tendencia editorial a publicar estudios con resultados positivos que aquellos otros que han obtenido resultados nulos o negativos. Por ello, nuestro primer análisis se dirige a someter a prueba esta posibilidad en nuestro caso.

En la inicial recogida de materiales realizamos un esfuerzo especial para recolectar trabajos de todo tipo, publicados o no. Ello dio como resultado la selección de un importante número de estudio no publicados, que ascendió a 15 trabajos (el 26.3%) de la muestra. En la Tabla 14.1 hemos analizado los resultados diferenciales de los estudios publicados y no publicados.

Tabla 14.1. Resultados de los estudios publicados versus no publicados (análisis del sesgo de publicación).

FUENTE	k	d+	95% CI	r+	Qw	&	p
Publicado	42	+0.2860	+0.2371 / +0.3349	+0.1416	140.5815		.0000000
No publicado	15	+0.3786	+0.2786 / +0.4786	+0.1860	28.0320		.0213701
OVERALL:	57	+0.3039	+0.2599 / +0.3478	+0.1502	171.2699		.0000000

Note: QB = 2.656 (p = 0.1031311).

95% CI: Intervalo de confianza al 95% en torno al TE medio, d+.

K: Número de estudios.

d+: TE medio.

Qw: Prueba Chi-Cuadrado de homogeneidad de los TEs.

p: Valor de probabilidad asociado a la prueba de homogeneidad.

r: Transformación a coeficiente de correlación de Pearson de d+.

Este análisis previo, relativo al sesgo de publicación, era condición necesaria para poder continuar con nuestro análisis de resultados. Como puede verse, no se observan diferencias significativas entre los estudios publicados y no publicados en cuanto a la efectividad exhibida por los programas aplicados (QB = 2.656, p = 0.1031). Por tanto, es posible afirmar con cierta seguridad que el fenómeno del sesgo de publicación no constituye una amenaza seria a los resultados de este meta-análisis. Hubiera supuesto una amenaza contra la validez de los resultados del meta-análisis el hecho de haber obtenido una mayor efectividad en los estudios publicados frente a los no publicados. Sin embargo, en nuestro caso incluso los estudios no publicados presentan un TE superior (d = 0.3786) que los publicados (d = 0.2860).

14.2. Exploración de posibles estudios meta-analíticos

Como hemos ya tenido ocasión de describir, todos los programas que hemos analizado consistieron en distintos tipos de intervención, aplicadas sobre grupos de sujetos, con la pretensión de alterar diversas variables de su conducta, psicológicas, etc., relacionadas con la delincuencia. Éstas fueron medidas inicialmente en grupos experimentales y en controles o en un período

PRE. Posteriormente, se midieron en un período POST -inmediatamente después de haber sido aplicada la intervención-, y/o en un período de SEGUIMIENTO -una vez transcurrido algún tiempo, después de la intervención-. Así pues, el análisis de nuestros resultados seguirá esta misma lógica: para el conjunto de estudios analizados, se han explorado todas las posibilidades de análisis, comparando entre sí cada grupo de medidas -PRE, POST, SEGUIMIENTO- para todas y cada una de las categorías de variables dependientes evaluadas -reincidencia, ajuste institucional, ajuste psicológico, habilidades sociales, ajuste académico, ajuste laboral, y seguimiento del tratamiento-. Esta exploración inicial se presenta en la Tabla 14.2, en la que se incluye, para cada comparación de medidas, el TE o diferencia media tipificada (d^+), para un determinado intervalo de confianza (L.I./L.S.) y grados de libertad (GL). En cada caso se presenta también información, mediante la prueba Chi-Cuadrado (Q_w), sobre la homogeneidad de los TEs, con un probabilidad determinada (p). Para cada puntuación d^+ o TE se ha obtenido también su correspondiente coeficiente de correlación (r), al tratarse de un estadístico frecuentemente utilizado en los meta-análisis, lo que nos permitirá en su caso la comparación con otras investigaciones que lo hayan empleado (véase Tabla 14.2).

Tabla 14.2. Tabla-resumen de la efectividad media y de la homogeneidad de los distintos meta-análisis posibles.

		K	d+	I.C. 95%		GL	Qw	p	r
				L.I.	L.S.				
GLOBAL		57	0.3039	0.26	0.35	56	171.270	.0000	.15
GLOBAL POSTEST		26	0.2490	0.18	0.32	25	101.402	.0000	.12
GLOBAL SEGUIMIENTO		45	0.2633	0.22	0.31	44	153.123	.0000	.13
Reincidencia	Postest	--	--	--	--	--	--	--	--
	Seguimiento	32	0.2430	0.20	0.29	31	123.479	.0000	.12
Ajuste Institucional	Postest	11	0.1055	0.03	0.19	10	55.577	.0000	.05
	Seguimiento	7	0.9011	0.53	1.27	6	9.372	.1537	.41
Ajuste Psicológico	Postest	6	0.0751	-0.01	0.16	5	5.027	.4126	.04
	Seguimiento	3	0.6703	0.07	1.27	2	0.815	.6653	.32
Habilidades Sociales	Postest	15	0.4555	0.33	0.58	14	24.349	.0416	.22
	Seguimiento	13	0.4018	0.29	0.52	12	32.326	.0012	.20
Ajuste Académico	Postest	7	0.6644	0.47	0.86	6	17.162	.0087	.31
	Seguimiento	3	0.5714	0.27	0.87	2	31.281	.0000	.27
Ajuste Laboral	Postest	0	--	--	--	--	--	--	--
	Seguimiento	7	0.5914	0.39	0.79	6	4.605	.5953	.28
Seguimiento Tratamiento	Postest	1	-0.121	--	--	--	--	--	--
	Seguimiento	--	--	--	--	--	--	--	--

I.C. 95%: Intervalo de confianza al 95% en torno al TE medio, d+.

L.I.: Límite Inferior del intervalo.

L.S.: Límite Superior del intervalo.

K: Número de estudios.

d+: TE medio.

GL: Grados de Libertad de la prueba de homogeneidad.

Qw: Prueba Chi-Cuadrado de homogeneidad de los TEs.

p: Valor de probabilidad asociado a la prueba de homogeneidad.

r: Transformación a coeficiente de correlación de Pearson de d+.

En la Tabla 14.2 la columna **k** refleja el número de estudios que incluye cada una de las comparaciones (d+) efectuadas. Como puede verse, sólo el análisis global se refiere al conjunto de los 57 estudios analizados. Los restantes análisis van paulatinamente decayendo en el número de estudios considerados. Obviamente, para "Ns" muy reducidos no es oportuno el meta-análisis. Se ha establecido el límite de $N > 10$ como criterio necesario para la consideración meta-analítica de una determinada comparación. Este criterio lo cumplen siete comparaciones, lo que nos ha conducido a la realización de SIETE META-ANÁLISIS DISTINTOS, que hemos consignado en negrita en la propia tabla 14.2 y que corresponden a los siguientes: (1) global, (2) global-postest, (3) global-seguimiento, (4) reincidencia (obviamente en seguimiento), (5) ajuste institucional en el postest, (6) habilidades sociales en el postest, y (7) habilidades sociales en el seguimiento. Nos referiremos al meta-análisis global en el presente capítulo y a los seis restantes en el capítulo siguiente.

14.3. Meta-análisis global

El meta-análisis global es, sin duda, debido a su generalidad, el más importante de los siete que hemos efectuado. En él se incluyen todos los 57 estudios seleccionados y el conjunto de todas las medidas resultado tomadas en ellos, ya sea en el postest ya sea durante el seguimiento. Nos permite contestar a las preguntas básicas de esta investigación: ¿En conjunto, resultaron los tratamientos aplicados efectivos?; ¿obtuvieron los grupos tratados mejores resultados que los controles?; ¿en qué grado?

Para cada estudio se ha calculado un TE promedio, a partir de todas las variables resultado que fueron evaluadas en el mismo, sin considerar la categoría de éstas. Ello quiere decir que los efectos aquí obtenidos se refieren a todas las categorías de variables incluidas en nuestra taxonomía, a saber: (S) SEGUIMIENTO O PERMANENCIA EN EL TRATAMIENTO (por ejemplo, opinión de los sujetos sobre el funcionamiento del programa,

asistencia al programa, etc.); (P) AJUSTE PSICOLÓGICO DE LA PERSONALIDAD Y DE LAS ACTITUDES (p.e., niveles de impulsividad y necesidad de estimulación, extraversión-neuroticismo-psicoticismo, prisionización, etc.); (A) AJUSTE ACADÉMICO/CULTURAL (p.e., asistencia y rendimiento en clase, participación en tareas educativas, motivación académica, etc.); (L) AJUSTE EN EL ENTRENAMIENTO Y PRÁCTICA LABORALES (p.e., calidad del trabajo realizado, aprovechamiento de la instrucción laboral, obtención de un puesto de trabajo a la salida de la institución, etc.); (H) AJUSTE DE LAS HABILIDADES PARA LA INTERACCIÓN SOCIAL (p.e., habilidades para promover la relación con otras personas, toma de perspectiva social, habilidades en la vida familiar, autocontrol sexual, nivel de desarrollo moral, medidas de ansiedad en situaciones de interacción, etc.); y (R) REINCIDENCIA (p.e., revocación de la libertad condicional, nuevos delitos, nuevas condenas, etc.).

En la Tabla 14.3 hemos recogido los estadísticos de TE global -d y r- obtenidos para cada uno de los 57 estudios analizados (véase la Tabla 14.3).