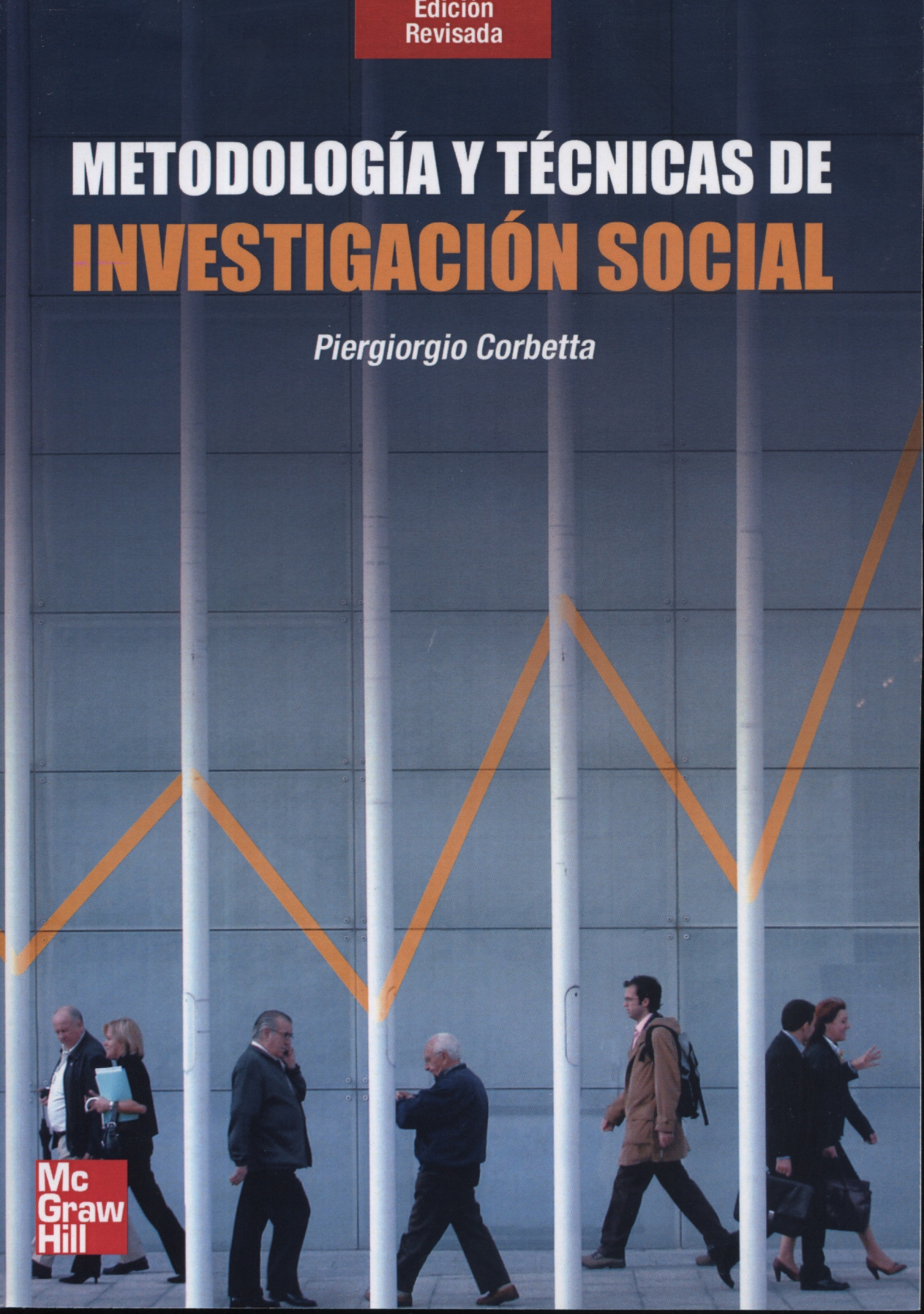
Edición Revisada

METODOLOGÍA Y TÉCNICAS DE INVESTIGACIÓN SOCIAL

Piergiorgio Corbetta



Mc Graw Hill

Edición Revisada

METODOLOGÍA Y TÉCNICAS DE INVESTIGACIÓN SOCIAL

Piergiorgio Corbetta

Este manual, elaborado por el autor a partir de su dilatada experiencia docente y extensa labor investigadora en el campo de las Ciencias Sociales, constituye una herramienta fundamental para el investigador social. El texto presenta todos los aspectos que intervienen en los métodos y técnicas de investigación social, desde los principios lógicos y las normas que los fundamentan, hasta los procedimientos específicos, su elaboración, codificación y puesta en práctica. Todo ello con una notable claridad expositiva, variedad de referencias bibliográficas y un listado completo de las fuentes relevantes. Esta edición revisada ha permitido realizar una actualización exhaustiva y profunda de todo el material disponible.

Piergiorgio Corbetta es profesor de Metodología y Técnicas de Investigación Social en la Facultad de Ciencias de la Educación de la Univesidad de Bolonia (Italia).



ISBN: 978-84-481-5610-7

9 788448 156107

www.mcgraw-hill.es

The McGraw-Hill Companies

METODOLOGÍA Y TÉCNICAS DE INVESTIGACIÓN SOCIAL

EDICIÓN REVISADA

METODOLOGÍA Y TÉCNICAS DE INVESTIGACIÓN SOCIAL
EDICIÓN REVISADA

No está permitida la reproducción total o parcial de este libro, ni su tratamiento informático, ni la transmisión de ninguna forma o por cualquier medio, ya sea electrónico, mecánico, por fotocopia, por registro u otros métodos, sin el permiso previo y por escrito de los titulares del Copyright.

DERECHOS RESERVADOS © 2007, respecto a la primera edición en español, por McGRAW-HILL/INTERAMERICANA DE ESPAÑA, S. A. U.

Edificio Valrealty, 1. ª Planta
Basauri, 17
28023 Aravaca (Madrid)

ISBN: 978-84-481-5610-7
Depósito legal: M. 4.700-2010

Traducción: Marta Díaz Ugarte y Susana Díaz Ugarte
Editor: José Manuel Cejudo
Técnico editorial: Susana Canedo
Diseño de cubierta: Juan Ramón Díaz Ruiz
Composición: Carolina Fraile Maldonado
Impreso por: Closas-Orcoyen, S. L.

IMPRESO EN ESPAÑA – PRINTED IN SPAIN

PRÓLOGO A LA EDICIÓN REVISADA

En las ciencias sociales, y por contraste con otras ciencias como la física o la biología, existen distintas maneras de enfocar y llevar a cabo una investigación empírica. Por este motivo, los estudiantes necesitan una especie de guía que les enseñe las reglas mínimas para que ellos mismos sean capaces de juzgar cuándo una investigación empírica es realmente científica. Muchas tradiciones históricas confluyen en el nacimiento y, sobre todo, en el desarrollo posterior de las ciencias sociales y todas ellas han aportado sus propias prácticas de investigación, sus propias convicciones sobre cuál es la mejor manera de estudiar un tema concreto. Así mismo, la gran variedad de temas de investigación posibles en las ciencias sociales exige a sus estudiantes y científicos la capacidad para dilucidar cuáles son el diseño y los instrumentos de investigación más adecuados para cada uno de ellos. Por ejemplo, si queremos estudiar cómo se toman las decisiones en un ayuntamiento seguramente tendremos que recurrir a técnicas de observación y producción de datos muy distintas a las que serían más adecuadas si quisiéramos estudiar la popularidad de las políticas aplicadas por el partido que gobierna en cualquier sistema democrático.

La diversidad de tradiciones y temas que se estudian en las ciencias sociales llevan a politólogos y sociólogos a reflexionar sobre el modo en que se realiza la investigación empírica, con el objetivo es dilucidar cuál de los distintos diseños de investigación posibles es más adecuado y en qué situaciones, cuáles son los errores más frecuentes, y cómo evitarlos. Por tanto, la reflexión sobre las técnicas de investigación es una reelaboración abstracta a partir de investigaciones prácticas y concretas.

Este libro puede considerarse una reflexión detallada sobre todas estas cuestiones, si bien es cierto que dedica más atención a los instrumentos de producción de evidencia empírica más frecuentes en el campo de las ciencias sociales. El manual comienza estableciendo las condiciones necesarias para poder hablar de investigación social. Para ello se remonta a los orígenes de la disciplina e indica muy detalladamente el modo en que se ha ido desarrollando a lo largo de su breve historia. A continuación, se detiene en la discusión pormenorizada de las técnicas más comunes, cuantitativas y cualitativas, de producción de datos en las ciencias sociales.

Por su contenido, este manual resulta bastante excepcional entre las publicaciones académicas en lengua española. De hecho, uno de los principales problemas que tiene que afrontar el docente de métodos y técnicas de investigación social actual es la falta de manuales que aborden la asignatura desde un punto de vista general e introductorio.

En cambio, sí abundan los libros especializados sobre una técnica concreta o un conjunto de técnicas de recopilación y análisis de evidencia empírica en el campo de las ciencias sociales. Por ejemplo, la serie de Cuadernos Metodológicos de la editorial del CIS. Muchos de ellos tratan estas técnicas a un nivel avanzado y reflejan un interesante proceso de evolución y puesta a punto de las distintas técnicas. Otros libros son más completos, pero están especializados en un tipo de técnicas específico: cuantitativas o cualitativas [1]. En definitiva, en el mercado español existen muy pocos manuales generales tan completos como éste [2].

Por lo que se refiere al contenido, en este libro el profesor Corbetta explica primero la base filosófica sobre la que se han construido y desarrollado históricamente los métodos y técnicas de investigación social. En los dos primeros capítulos, el autor ilustra de un modo ejemplar la dicotomía filosófica fundamental presente en los orígenes y el desarrollo de la investigación social. Su presentación parte de la ilusión positivista de que las ciencias sociales podían imitar a las ciencias naturales sin dificultad alguna, para desembocar en las críticas de quienes defienden que la investigación de las sociedades debe hacerse por caminos alternativos que tengan en cuenta la individualidad inherente al ser humano. Estas dos posturas opuestas se consolidaron desde principios del siglo XX y las técnicas de investigación social actuales hunden sus raíces en ellas. Por este motivo resulta de gran utilidad para el estudiante conocerlas, para comprender el espíritu que anima a las técnicas de investigación social que se explican y analizan en el resto del libro.

En la segunda parte del manual se exponen las técnicas cuantitativas, y en la tercera, las cualitativas. El autor considera que estos dos tipos de técnicas son complementarios y explica de manera convincente su postura en el curso del libro. Esta postura contiene un matiz importante en relación con la opinión predominante de quienes defienden que las diferencias entre ambos enfoques son meramente técnicas o de estilo, y que la lógica de la inferencia es la misma para los dos (por ejemplo, Bryman, 1988; Goldthorpe, 2000, y King et. al., 1994). El matiz del profesor Corbetta es relevante porque defiende que estas dos formas de hacer investigación son la expresión lógica de dos visiones epistemológicas distintas, dos paradigmas diferentes que implican modos alternativos de comprender la realidad social, los objetivos de la investigación, el papel del investigador y la instrumentación técnica.

---

[1] El número de manuales dedicados exclusivamente a las técnicas cuantitativas o cualitativas es demasiado grande como para citarlos todos en este espacio limitado.

[2] Entre los manuales publicados en España, sólo los siguientes tienen un carácter general: García Ferrando *et al.*, 1986; Manheim y Rich, 1988; King *et al.*, 2000. Los dos primeros se escribieron hace casi dos décadas y la profundidad en el tratamiento de las técnicas es desigual, probablemente porque cada capítulo lo escribe un autor distinto. Por otro lado, el manual de King *et. al.* limita su discusión a la metodología de las ciencias sociales, en concreto, se centra en el tema de la selección del mejor diseño de investigación para someter a prueba empírica cualquier hipótesis teórica, pero no discute los pormenores de las técnicas de producción de evidencia empírica de manera sistemática y detallada como el presente volumen.

Estas discusiones iniciales sirven como introducción a las siguientes partes del manual, más extensas, y dedicadas a la discusión de las técnicas de observación y producción de evidencia empírica. Suele ocurrir que el estudiante medio, poco iniciado en el tema encuentra demasiado abstracto el contenido de esta primera parte y alguna dificultad para entender parte de los conceptos que se utilizan. Mi recomendación a los estudiantes es que sigan avanzando en el manual y acepten la lógica de la inmersión. De la misma manera que para hablar un idioma no es necesario conocer a la perfección todos los detalles de la gramática, tampoco el estudiante de técnicas de investigación social tiene por qué conocer todos los engranajes de la producción y el análisis de los datos en ciencias sociales para entender, al menos en buena parte, la discusión que se presenta en la primera parte del manual.

Tal vez la virtud más destacada de este manual es que permite al lector adquirir lo que yo denomino competencia metodológica. Ésta se compone de dos aspectos que tienen el mismo grado de importancia: el saber hacer y el saber evaluar. Es decir, la competencia metodológica permite conocer y aplicar una serie de reglas para poder comprobar empíricamente una hipótesis teórica. Esto implica producir, buscar o utilizar material empírico, como, por ejemplo, el cuestionario, la entrevista cualitativa, los grupos de discusión, analizarlo, presentar los resultados de dicho análisis, y sacar conclusiones sobre la comprobación de la hipótesis teórica de la que se ha partido. Esto es lo que yo entiendo por saber hacer una investigación empírica.

La competencia metodológica también permite al estudiante evaluar el trabajo de investigación empírica de otros autores que leerá, analizará y criticará en las distintas asignaturas que estudie. En concreto, el estudiante podrá valorar hasta qué punto una investigación empírica está bien hecha o no. Es decir, si las conclusiones a las que llega están justificadas, si la manera en que presenta los resultados es adecuada, si el análisis del material empírico que ha realizado el autor es correcto.

En realidad, las dos dimensiones de la competencia metodológica están muy ligadas entre sí y si un estudiante desconoce el modo en que se lleva a cabo una investigación empírica muy difícilmente podrá juzgar el trabajo de los científicos sociales. Al mismo tiempo, leyendo la producción científica de los autores el estudiante va tomando conciencia de los dilemas a los que éstos se tienen que enfrentar, lo que le permite aprender más. Obviamente, el objetivo de este manual no es que, al terminar de leerlo, el lector sea capaz de realizar una investigación empírica por sí mismo, pero sí que sepa al menos por dónde empezar, que conozca las reglas fundamentales que hay que aplicar para comprobar una hipótesis, dónde puede conseguir datos para ello, cuáles son los fallos que se deben evitar en el diseño de la investigación.

Resulta obvio que para aprender a hacer investigación social no basta con la lectura de un manual. La única manera de aprender a usar la diversidad de técnicas que se explican en este libro es haciendo investigación: partir de una pregunta teórica, plantear hipótesis, buscar o producir evidencia empírica para poner a prueba esas hipótesis, etc. Sólo así el estudiante adquirirá una comprensión plena de las técnicas de investigación, sus ventajas y sus límites. De ahí la importancia de com-

plementar la teoría con las clases prácticas en la asignatura de métodos y técnicas de investigación social.

El lector encontrará en este libro una gran cantidad de ejemplos de investigaciones, tanto clásicas como contemporáneas, producidos en diversos campos de las ciencias sociales, como la antropología, la sociología, la ciencia política, la psicología social, la educación y la historia. Todos estos ejemplos están explicados con un estilo ameno y asequible, y con el objetivo de visualizar el contexto en el que se han utilizado las técnicas correspondientes, las preguntas teóricas que el autor trataba de responder en cada caso, las conclusiones a las que llega y la utilidad, o no, de la técnica utilizada para cada investigación.

Los ejemplos resultan especialmente útiles para la explicación de las técnicas cualitativas, en las que no existe una manera estandarizada de realizar las investigaciones ni de aplicar las técnicas. Gracias a estos ejemplos, el manual consigue explicar cómo se hace una investigación empírica en ciencias sociales y, lo que es más importante, por qué ante determinadas preguntas teóricas se utilizan unos procedimientos y unas técnicas de investigación y no otros.

En definitiva, el lector tiene en sus manos una guía muy útil para aprender cómo se hace una investigación empírica. Frente a los típicos manuales que ofrecen una lista de los pasos a seguir para desarrollar una determinada técnica, este libro anima al lector a aprender a utilizar las técnicas de investigación, cuáles son sus ventajas y sus límites, por qué es más adecuado utilizar unas y no otras en cada caso. En ese sentido, el libro fomenta el aprendizaje adoptando un punto de vista crítico. Las reflexiones sobre la calidad de los datos que se producen en la actualidad resultan, en mi opinión, especialmente relevantes. El autor denuncia la negligencia con la que a veces se desarrolla la parte de producción de evidencia empírica en las investigaciones, en el campo de las ciencias sociales. Esto se aplica sobre todo a las investigaciones que utilizan encuestas representativas, que suelen encargar la recolección de datos a empresas privadas. En general, la falta de sensibilidad y reflexión sobre la calidad, fiabilidad y validez de la evidencia empírica conlleva el peligro de trabajar con datos de pésima calidad, que ningún análisis estadístico, por sofisticado que sea, puede mejorar.

En cuanto a la precisión y al estilo, se trata de un manual riguroso y completo en el que se presta especial atención al uso de la terminología técnica. Los contenidos son accesibles para cualquier estudiante aunque no tenga conocimientos previos sobre la investigación en ciencias sociales, dado que el autor ofrece definiciones y explicaciones detalladas de cada concepto. Además, para el lector que desee profundizar en los temas tratados, al final de cada capítulo el autor ha incluido una selección de bibliografía clásica y contemporánea que resulta muy útil. A la misma he añadido las referencias publicadas por editoriales españolas, entre las que destacan algunas de la serie Cuadernos Metodológicos editada por el CIS.

A pesar de que hasta el momento me haya referido en mi presentación a la utilidad de este manual para los estudiantes de cualquier disciplina en el campo de las ciencias sociales, este libro también puede interesar a otro tipo de lectores como, por ejemplo, quienes deseen aprender la forma de hacer investigación empírica y aplicarla después a la práctica. Es decir, para quienes realizan investigación con una motivación política (en inglés, *policy oriented*) y, por tanto, con pretensiones de intervenir en la realidad social y política. Así mismo, este volumen puede resul-

tar útil para el consumidor de investigación social no especializado, como periodistas, trabajadores sociales, técnicos de la administración pública, etc. Hoy en día, y en la mayoría de los campos de las ciencias sociales, existe una gran cantidad de datos disponibles para estudiar y analizar nuestras sociedades. Para producir y analizar dicha evidencia se requiere un cierto grado de preparación que es posible adquirir con este manual.

Para concluir, mi experiencia docente a lo largo de casi una decena de años de la asignatura de Métodos y Técnicas de Investigación Social me permite afirmar que este manual de consulta y estudio resulta imprescindible para el estudiante o todo aquel interesado en entender la esencia de las ciencias sociales. El libro fue escrito originalmente en italiano y publicado por la editorial académica *Il Mulino*. Más adelante, fue traducido al inglés y al español. El manual en inglés fue publicado en junio de 2003 por la editorial de metodología *Sage*. La primera edición en español fue publicada también en el año 2003. Para ella realicé un trabajo de adaptación de las partes referidas a las bases de datos disponibles (capítulo 5) y las fuentes estadísticas oficiales (capítulo 7) al contexto español, europeo y latinoamericano. En esta nueva edición revisada, se ha adaptado el estilo para hacerlo más asequible al lector de habla hispana, se han corregido de forma exhaustiva las erratas, y finalmente, se han actualizado los contenidos de bibliografía y bases y archivos de datos, que aparecen y se renuevan de manera constante en la disciplina de las ciencias sociales.

Madrid, marzo de 2007.

MARTA FRAILE MALDONADO
Investigadora Titular del CSIC
Unidad de Políticas Comparadas

AGRADECIMIENTOS DEL AUTOR

Quiero expresar mi gratitud a todas aquellas personas a quienes debo sugerencias, críticas e indicaciones, aquellos con quien he mantenido un diálogo científico constructivo. Varios colegas han leído partes de este libro, con otros he discutido argumentos específicos. En concreto, quiero dar las gracias de manera explícita a Fabrizio Bernardi, Massimiano Bucchi, Sergio Brasini, Mario Callegaro, Giorgio Chiari, Antonio Cobalti, Asher Colombo, Giolo Fele, Pierangelo Peri, Marilena Pillati, Maurizio Pisati, Francesca Rodolfi, Raffaella Solaini, Marco Santoro y Antonio Strati.

Quiero dar las gracias en especial a Alberto Marradi, con quien he discutido prácticamente todos los temas tratados en el manual y de quien he recibido un estímulo intelectual fundamental. También me gustaría mostrar mi agradecimiento a mis amigos del Instituto Cattaneo, Marzio Barbagli, Roberto Cartocci, Raimondo Catanzaro, Arturo Parisi, Hans Schadee, Giancarlo Gasperoni, con los que he compartido muchos años de investigación y discusión intelectual y quienes seguramente también han dejado su huella en este libro.

Doy también las gracias a mis amigos del *Survey Research Center* de la Universidad de Berkeley, y a sus directores, Percy Tannenbaum, Mike Hout y Henry Brady, donde pasé un año sabático y varios periodos de inmersión en el entorno académico estadounidense. Entre mis amigos americanos en Berkeley, deseo expresar mi gratitud a Tom Piazza y Jim Wiley por sus largas discusiones y sus valiosas sugerencias, y mi agradecimiento especial a John Stiles, cuya ayuda en la adaptación de algunas partes del libro al contexto anglosajón resultó fundamental.

CONTENIDO

Primera parte
LA LÓGICA DE LA INVESTIGACIÓN SOCIAL

1. Los paradigmas de la investigación social 3

1.1. Kuhn y los paradigmas de las ciencias 4
1.2. Tres cuestiones de fondo 7
1.3. Positivismo 9
1.4. Neopositivismo y postpositivismo 13
1.5. Interpretativismo 18
1.5.1. Los inicios 18
1.5.2. Max Weber: objetividad y orientación hacia la individualidad 20
1.5.3. Desarrollos posteriores 23
1.6. Críticas, posturas radicales y nuevas tendencias 24
1.7. Síntesis del capítulo 1 28
1.8. Lecturas complementarias 29

2. Investigación cuantitativa e investigación cualitativa 31

2.1. Paradigma neopositivista: *Crime in the Making,* de Sampson y Laub 32
2.1.1. Hipótesis 33
2.1.2. Diseño de la investigación 34
2.1.3. Observación empírica y resultados del análisis 34
2.2. Paradigma interpretativo: *Islands in the Street,* de Jankowski 36
2.2.1. Diseño de la investigación y recogida de datos 37
2.2.2. Hipótesis 37
2.2.3. Interpretación del material empírico 39
2.3. Investigación cuantitativa e investigación cualitativa: una comparación 40
2.3.1. Planteamiento de la investigación 40
2.3.2. Recopilación de los datos 47
2.3.3. Análisis de los datos 49

2.3.4. Resultados .....
2.4. Dos formas diferentes de conocer la realidad social .....
2.5. Síntesis del capítulo 2 .....
2.6. Lecturas complementarias .....

Segunda parte
LA PRODUCCIÓN DE LOS DATOS: TÉCNICAS CUANTITATIVAS

. **La traducción empírica de la teoría** .....

3.1. Estructura «tipo» de la investigación cuantitativa .....
3.1.1. Las cinco fases del proceso de investigación .....
3.2. De la teoría a las hipótesis .....
3.2.1. Teoría .....
3.2.2. Hipótesis .....
3.2.3. La transformación de la teoría en hipótesis .....
3.3. De los conceptos a las variables .....
3.4. Unidad de análisis .....
3.4.1. Los distintos tipos de unidad de análisis .....
3.5. Variables .....
3.6. Variables nominales, ordinales y cardinales .....
3.6.1. Variables nominales .....
3.6.2. Variables ordinales .....
3.6.3. Variables cardinales .....
3.7. Conceptos, indicadores e índices .....
3.8. Error total .....
3.9. Fiabilidad y validez .....
3.9.1. Fiabilidad .....
3.9.2. Validez .....
3.10. Síntesis del capítulo 3 .....
3.11. Lecturas complementarias .....

. **Causalidad y experimento** .....

4.1. Concepto de causa .....
4.2. Confirmación empírica de la relación causal .....
4.2.1. Covariación entre variables dependientes e independientes
4.2.2. Dirección causal .....
4.2.3. Control de las variables superfluas .....
4.3. Análisis de la covariación y experimento .....
4.3.1. Análisis de la covariación .....
4.3.2. Experimento .....

4.4. El experimento en las ciencias sociales .....
4.4.1. La solución científica .....
4.4.2. La solución estadística .....
4.5. Experimentos de laboratorio y experimentos de campo .....
4.5.1. Experimentos de laboratorio .....
4.5.2. Experimentos de campo .....
4.6. Experimentos y cuasiexperimentos .....
4.6.1. Experimentos verdaderos .....
4.6.1.1. Diseño «sólo después» a dos (o más) grupos .....
4.6.1.2. Diseño «antes-después» a dos (o más) grupos ....
4.6.1.3. Diseño de Solomon a cuatro grupos .....
4.6.1.4. Diseño factorial .....
4.6.2. Cuasiexperimentos .....
4.6.2.1. Diseño «antes-después» a un solo grupo .....
4.6.2.2. Diseños en serie temporal interrumpida .....
4.6.2.3. Diseño «antes-después» a dos grupos sin asignación causal .....
4.6.2.4. Diseño «ex postfacto» .....
4.7. Ventajas y limitaciones del experimento .....
4.8. Síntesis del capítulo 4 .....
4.9. Lecturas complementarias .....

5. La encuesta por muestreo .....

5.1. La encuesta por muestreo en la investigación social .....
5.2. La estandarización o invariación del estímulo .....
5.2.1. Criterio objetivista y criterio constructivista .....
5.2.2. Enfoque uniformista y enfoque individualista .....
5.2.3. El objetivo del mínimo común denominador .....
5.3. Fiabilidad del comportamiento verbal .....
5.4. Fondo y forma de las preguntas .....
5.4.1. Propiedades sociodemográficas, actitudes y comportamientos .....
5.4.2. Preguntas abiertas y preguntas cerradas .....
5.5. Formulación de las preguntas .....
5.6. Batería de preguntas .....
5.7. Modalidades de recopilación de datos .....
5.7.1. Entrevistas cara a cara .....
5.7.2. Entrevistas telefónicas .....
5.7.3. Cuestionarios autocumplimentados .....
5.7.4. Entrevistas informatizadas .....
5.8. Organización de la recopilación de datos .....
5.9. Análisis secundario y encuestas sucesivas .....
5.9.1. Análisis secundario .....
5.9.2. Encuestas sucesivas (diacrónicas) .....

5.10. Bases de datos en España y otros países 196
5.10.1. Estudios transversales repetidos 197
5.10.2. Estudios longitudinales 201
5.11. Nuevas tendencias y nuevos riesgos 202
5.12. Síntesis del capítulo 5 204
5.13. Lecturas complementarias 206

6. La técnica de las escalas 209

6.1. La operacionalización empírica de conceptos complejos 210
6.2. Preguntas de respuesta graduada: la autonomía semántica de las respuestas 213
6.3. La escala de Likert 219
6.4. El escalograma de Guttman 227
6.5. El diferencial semántico 233
6.6. El test sociométrico 238
6.7. Escalas unidimensionales y escalas multidimensionales 243
6.8. Síntesis del capítulo 6 245
6.9. Lecturas complementarias 246

7. Las fuentes estadísticas oficiales 247

7.1. Una investigación basada en estadísticas oficiales 248
7.2. Las estadísticas oficiales 251
7.2.1. La producción 252
7.2.2. La unidad de análisis 254
7.2.3. El contenido 256
7.2.4. El tamaño del registro 257
7.3. Las estadísticas oficiales en España y en otros países 258
7.3.1. Agencias oficiales 258
7.3.2. Distribución a través de organismos oficiales 261
7.3.3. Metadatos 263
7.3.4. Unidades geográficas 263
7.3.5. Normas y clasificaciones 264
7.3.6. Otros canales de distribución 265
7.3.7. Datos agregados y microdatos 266
7.4. El estudio de las estructuras sociales 267
7.5. Síntesis del capítulo 7 269
7.6. Lecturas complementarias 270

8. El muestreo 271

8.1. Población y muestreo 272
8.2. Error de muestreo 274
8.3. Tamaño de la muestra 277

8.4. Muestras probabilísticas 282
8.5. Muestras no probabilísticas 287
8.6. Problemas de muestreo en investigación social 289
8.6.1. Error de cobertura 290
8.6.2. Error de no respuesta 291
8.7. Ponderación 293
8.8. Representatividad y tamaño de la muestra 296
8.9. Síntesis del capítulo 8 298
8.10. Lecturas complementarias 299

Tercera parte
LA RECOPILACIÓN DE LOS DATOS: TÉCNICAS CUALITATIVAS

9. La observación participante 301

9.1. Observación y observación participante 304
9.2. Campos de aplicación y desarrollo de la observación participante 307
9.3. Observación declarada y encubierta: el acceso y los informadores 313
9.4. Qué observar 318
9.5. Registro de la observación 323
9.6. Análisis del material empírico 325
9.7. La observación en «microsociología» 331
9.7.1. Del estudio de «los otros» al estudio de «nosotros mismos» ... 331
9.7.2. La etnometodología 333
9.8. Limitaciones y recursos 337
9.9. Síntesis del capítulo 9 339
9.10. Lecturas complementarias 341

10. La entrevista cualitativa 343

10.1. La obtención de datos mediante preguntas 344
10.2. Entrevista cuantitativa y entrevista cualitativa 345
10.3. Tipos de entrevista 349
10.3.1. Entrevistas estructuradas, semiestructuradas y no estructuradas 350
10.3.2. Casos especiales 357
10.4. Realización de la entrevista 361
10.5. Análisis del material empírico 366
10.5.1. El análisis de tipo cualitativo 366
10.5.2. El análisis de tipo mixto (cuantitativo y cualitativo) 367

10.6. Falta de estandarización e interacción 369
10.7. Síntesis del capítulo 10 372
10.8. Lecturas complementarias 373

11. El uso de los documentos 375

11.1. Definición del objeto 376
11.2. Documentos personales 377
11.2.1. Autobiografías 378
11.2.2. Diarios 379
11.2.3. Cartas 380
11.2.4. Testimonios orales 382
11.3. Documentos institucionales 388
11.3.1. Medios de comunicación 390
11.3.2. Narrativa, textos pedagógicos, cuentos populares 392
11.3.3. Material judicial 394
11.3.4. Documentos de la política 395
11.3.5. Documentos empresariales y administrativos 397
11.3.6. Huellas físicas 399
11.4. Leer los documentos humanos 401
11.4.1. Leer los documentos personales 401
11.4.2. Leer los documentos institucionales 402
11.5. Síntesis del capítulo 11 403
11.6. Lecturas complementarias 404

Bibliografía 405

Índice analítico 419

PRIMERA PARTE

LA LÓGICA DE LA INVESTIGACIÓN SOCIAL

CAPÍTULO 1

LOS PARADIGMAS DE LA INVESTIGACIÓN SOCIAL

1.1. KUHN Y LOS PARADIGMAS DE LAS CIENCIAS 4

1.2. TRES CUESTIONES DE FONDO 7

1.3. EL POSITIVISMO 9

1.4. EL NEOPOSITIVISMO Y EL POSTPOSITIVISMO 13

1.5. EL INTERPRETATIVISMO 18

1.5.1. Los inicios 18
1.5.2. Max Weber: objetividad e individualidad 20
1.5.3. Desarrollos posteriores 23

1.6. CRÍTICAS, POSTURAS RADICALES Y NUEVAS TENDENCIAS 24

1.7. SÍNTESIS DEL CAPÍTULO 1 28

1.8. LECTURAS COMPLEMENTARIAS 29

En este capítulo se presentan los orígenes filosóficos en los que se basan los dos enfoques principales de la investigación social, que han generado respectivamente las familias de técnicas cuantitativas y cualitativas. Comenzaremos con una reflexión sobre el concepto de paradigma, es decir, la perspectiva que inspira y dirige todas las ciencias. A continuación, ilustraremos los orígenes históricos y los principios inspiradores del paradigma positivista y del paradigma interpretativo. El capítulo concluirá con una serie de reflexiones sobre las tendencias actuales en investigación social.

1.1. KUHN Y LOS PARADIGMAS DE LAS CIENCIAS

La noción de «paradigma» tiene un origen antiguo en la historia del pensamiento filosófico. Fue utilizada tanto por Platón, en el sentido de «modelo», como por Aristóteles, en el sentido de «ejemplo». En las ciencias sociales se ha abusado de su uso y resulta confusa por los múltiples y diversos significados que se le atribuyen: sinónimo de teoría, articulación interna de una teoría, sistema de ideas de orden precientífico, corriente de pensamiento o escuela, procedimiento de investigación modélico o sinónimo de método. Por ello resulta útil considerar brevemente el concepto de paradigma tal y como lo planteó a principios de los años sesenta a filósofos y sociólogos de la ciencia el estudioso Thomas Kuhn, en su célebre ensayo *La estructura de las revoluciones científicas* (1962).

La reflexión de Kuhn tiene como objeto el desarrollo histórico de las ciencias y supone un rechazo al concepto tradicional de ciencia como acumulación progresiva y lineal de nuevos logros. Según la concepción acumulativa tradicional, cada invención y descubrimiento se añadiría al cuerpo de conocimientos existente, igual que los ladrillos se superponen en la construcción de un edificio de varias plantas. Sin embargo, según Kuhn, aunque éste es el proceso de desarrollo de la ciencia en épocas «normales», existen también periodos «revolucionarios» en los que se interrumpe la continuidad con el pasado y se inicia una nueva construcción, como cuando a veces, siguiendo con la metáfora de la edificación, se demuele un edificio viejo de ladrillo para construir otro con materiales distintos, por ejemplo, un rascacielos de cristal y aluminio.

Kuhn ilustra su argumentación con una serie de ejemplos tomados de las ciencias naturales, sobre todo de la física. Cita el desarrollo de la física óptica, que en la actualidad se basa en una interpretación cuántica, según la cual la luz estaría constituida por fotones, entidades que tienen al mismo tiempo características propias de las ondas y de las partículas. Kuhn añade, no obstante, que esta teoría nació en el siglo XX. Antes de que Planck, Einstein y otros desarrollaran dicha teoría, en las universidades se enseñaba que la luz era una onda, según una teoría propuesta a principios del siglo XIX. Y aún antes, en el si glo XVII, reinaba la teoría de la óptica newtoniana, según la cual la luz estaba constituida por corpúsculos materiales.

El paso de una perspectiva teórica a otra tiene consecuencias tan drásticas para la disciplina afectada que Kuhn no duda en utilizar el término «revolución» y habla precisamente de «revolución científica». ¿Qué cambia, en una determinada discipli-

na, con una de estas revoluciones? «Cambian los problemas que se proponen a la investigación científica, y también los criterios para determinar lo que se debía considerar un problema admisible o una solución válida para el mismo» [*ibídem,* 6]. Y se realiza una reorientación de la disciplina que consiste en «transformar la estructura conceptual con la cual los científicos estudian el mundo» [*ibídem,* 102]. Esta «estructura conceptual» es lo que Kuhn llama «paradigma», y este aspecto de su teoría, más que su análisis del desarrollo de la ciencia, es lo que nos interesa ahora.

Sin un paradigma, una ciencia no tendría orientaciones ni criterios de elección: todos los problemas, todos los métodos y todas las técnicas serían válidos. El paradigma sirve como guía para la ciencia: «los paradigmas proporcionan a los científicos no sólo un modelo, sino también ciertas indicaciones necesarias para construirlo. Cuando aprende un paradigma, el científico adquiere de manera simultánea teorías, métodos y criterios, que suelen aparecer combinados de manera inextricable» [*ibídem,* 109].

Resumen 1.1. PARADIGMA

¿Qué entiende Kuhn por paradigma? Con este término Kuhn designa una perspectiva teórica:

a) común a y reconocida por los científicos de una disciplina concreta;
b) basada en logros que preceden a la propia disciplina;
c) que dirigen la investigación, mediante:

*c*1) la identificación y elección de los hechos relevantes a estudiar,
*c*2) la formulación de hipótesis para explicar el fenómeno observado, y
*c*3) la preparación de las técnicas de investigación empírica necesarias.

Nótese que hemos evitado intencionadamente la palabra «teoría» (en la definición anterior hemos utilizado el término «perspectiva teórica»). En efecto, un paradigma es más amplio y también más general que una teoría: es una visión del mundo, una ventana mental, una pauta de lectura que precede a la elaboración de la teoría. El paradigma copernicano del universo, así como el de la mecánica cuántica, representan marcos generales dentro de los que se encuadran teorías específicas sobre los movimientos orbitales o sobre la estructura del átomo.

Kuhn define *ciencia normal* como las fases de una disciplina científica en las que predomina un determinado paradigma, que es aceptado por toda la comunidad científica. Durante esta fase, es decir, hasta que el paradigma dominante se sustituye por otro con una «revolución», la ciencia avanza efectivamente según el proceso lineal y cumulativo que se suele atribuir al desarrollo científico. «La tarea de la ciencia normal no es en absoluto descubrir nuevos tipos de fenómenos [...], la investigación en el ámbito de la ciencia normal está orientada a la articulación de los fenómenos y teo-

rías derivados del paradigma» [*ibídem,* 24].

En las ciencias naturales encontramos numerosos ejemplos de paradigmas científicos. Si retomamos el ejemplo anterior, podemos hablar del paradigma corpuscular, ondulatorio o cuántico, de la física óptica. Del mismo modo podemos citar como parejas de paradigmas alternativos que se han sucedido en el tiempo la mecánica de Newton y la mecánica de Einstein, la cosmología tolemaica y la copernicana, y así sucesivamente.

¿Hasta qué punto podemos hablar de paradigmas en las ciencias sociales? Kuhn subraya que el paradigma es una característica definitoria de las ciencias «maduras». Antes de que Newton introdujera la teoría corpuscular de la luz, los científicos que la estudiaban no tenían un paradigma común, sino que existían diversas escuelas y subescuelas, cada una con su punto de vista y su teoría, por lo que, según Kuhn, «si bien quienes desarrollaban actividades en ese campo eran científicos, el resultado de su actividad no alcanzaba el rango de ciencia» [*ibídem,* 13].

Desde esta perspectiva, las ciencias sociales, por el hecho de carecer cada una de ellas por separado de un único paradigma ampliamente compartido por la comunidad científica, se encontrarían en una situación preparadigmática, salvo quizás la economía (según Kuhn, los economistas «están de acuerdo sobre la definición de economía», mientras «aún queda pendiente la cuestión de qué sectores de las ciencias sociales han alcanzado ya paradigmas como éste» [*ibídem,* 14]).

Lo dicho sobre las ciencias sociales se aplica también a la sociología. Parece difícil identificar, incluso para períodos limitados de la historia de la disciplina, un paradigma común a toda la comunidad de sociólogos.

Existe otra interpretación del pensamiento de Kuhn, que se ha propuesto precisamente para aplicar sus categorías a la sociología. Se trata de una redefinición del concepto de paradigma, en la que se mantienen todos los elementos de la definición original, es decir, visión teórica que define la relevancia de los hechos sociales, proporciona hipótesis interpretativas y orienta las técnicas de la investigación empírica, excepto uno, el consenso de la comunidad científica, al menos en periodos de «ciencia normal».

De este modo se abre la posibilidad de convivencia, dentro de una determinada disciplina, de varios paradigmas, y la sociología pasa de ser *preparadigmática,* a ser una disciplina *multiparadigmática.* Ésta es, entre otras, la lectura de Friedrichs [1970], que tras haber puesto en evidencia el paradigma que guía el funcionalismo estructural de Parsons, encuentra en el planteamiento dialéctico de origen marxista el segundo paradigma de la sociología, que desplaza el foco de atención de los conceptos de sistema y consenso, propios de la perspectiva funcionalista, a la idea del conflicto.

Esta interpretación del concepto de paradigma en términos de perspectiva teórica global es la más difundida en las ciencias sociales. Sin embargo, no es la única, sino que existen otras claramente opuestas a ella. Sin embargo, esta lectura menos rigurosa de la categoría original kuhniana realizada para adaptarla al estatus de las ciencias sociales, tampoco debe banalizarse identificando paradigma con teoría o corriente de pensamiento. De hecho, sigue siendo fundamental en el concepto de paradigma su carácter preteórico, en último término metafísico, de «visión que

orienta», de imagen del mundo, o mejor aun, como escribe Friedrichs [*ibídem,* 55], de «imagen fundamental que una disciplina tiene de su objeto», que guía y organiza tanto la reflexión teórica como la investigación empírica, y como tal precede a ambas.

En este sentido, el concepto de paradigma nos parece útil también para el análisis de los distintos marcos de referencia que han sido desarrollados y aún se consultan en el campo de la metodología de la investigación social.

1.2. TRES CUESTIONES DE FONDO

Una vez definido y delimitado el concepto de paradigma e indicada su aplicación en el campo de las ciencias sociales, conviene dejar el terreno resbaladizo de los paradigmas de la *teoría sociológica* (¿un paradigma?, ¿dos paradigmas?, ¿cien paradigmas?), para pasar a un terreno más estable, el de la metodología de la *investigación social.* Pero tampoco en este caso nos adentraremos en la compleja problemática epistemológica de cuántos y cuáles pueden ser los marcos de referencia filosóficos que orientan la investigación empírica en el campo de las ciencias sociales.

Nos limitaremos a un examen de carácter *histórico,* describiendo brevemente las perspectivas fundamentales que se han propuesto y consolidado en el curso de la evolución de la disciplina. Dado que este libro trata sobre técnicas de investigación social, resulta fundamental, en primer lugar, plantear la cuestión de los *paradigmas fundacionales de la investigación social,* de los que nacieron los primeros procedimientos operativos y que han guiado sucesivamente el avance de la investigación empírica. Sabemos que una de las funciones de un paradigma es definir los métodos y técnicas de investigación aceptables para una disciplina. Como escribe Hughes:

> Cada procedimiento o instrumento de investigación está entrelazado de manera inextricable con las interpretaciones concretas del mundo que tiene el investigador y sus modos de conocer ese mundo. Usar un cuestionario o una escala de actitudes, asumir el papel de observador participante o construir una muestra aleatoria [...] equivalen a aceptar concepciones del mundo que permitan el uso de estos instrumentos para lograr los objetivos fijados. Ninguna teoría o método de investigación [...] se justifican por sí mismos: su eficacia, su propia calificación de instrumentos de investigación [...] dependen en última instancia de justificaciones de tipo filosófico.
>
> [Hughes, 1980, 13]

¿Podemos identificar, entre las perspectivas filosóficas que han generado y acompañado en su desarrollo a la investigación social, visiones lo bastante generales, coherentes y operativas, como para poder atribuirles la categoría de paradigma? En nuestra opinión la respuesta a esta pregunta es afirmativa. Existe un acuerdo generalizado sobre la existencia de dos grandes marcos de referencia globales que han orientado la investigación social desde su nacimiento: la perspectiva «empirista» y la «humanista» (las etiquetas son muy variadas, destacando entre ellas el «objetivismo» y «subjetivismo»; aquí utilizaremos el término canónico de «positivismo» y,

aunque esté menos consolidado, el de «interpretativismo»). Se trata de dos divisiones orgánicas y claramente opuestas de la realidad social y los modos de conocerla, que han generado dos conjuntos de técnicas de investigación coherentes y muy diferenciados entre sí. Antes de describirlas, debemos conocer sus orígenes filosóficos para poder comprenderlas correctamente.

Recordemos que estos paradigmas no son teorías sociológicas, sino concepciones generales sobre la naturaleza de la realidad social, la naturaleza del hombre, y el modo en que éste puede conocer dicha realidad. Para comparar adecuadamente los dos paradigmas arriba mencionados, trataremos de entender cómo responden a las preguntas fundamentales a las que se enfrenta la investigación social, y, en general, la investigación científica. Estas preguntas se pueden expresar como sigue: ¿existe la realidad (social)?, ¿es conocible?, ¿cómo podemos conocerla? En otras palabras: *esencia, conocimiento* y *método.*

La cuestión ontológica [1]. Esta cuestión concierne a la naturaleza de la realidad social y su forma. Nos pregunta si el mundo de los hechos sociales es un mundo real y objetivo con existencia autónoma fuera de la mente humana e independiente de la interpretación que hace del mismo el sujeto. Es decir, nos pregunta si los fenómenos sociales son «cosas en sí mismas» o «representaciones de cosas». El problema enlaza con la cuestión filosófica más general de la existencia de las cosas y del mundo exterior. De hecho, la existencia de la idea en el pensamiento no nos dice nada sobre la existencia en la realidad del objeto representado, así como un dibujo no prueba la existencia de la cosa representada.

La cuestión epistemológica [2]. Esta cuestión trata sobre la relación entre el «quién» y el «qué», y el resultado de esta relación. Concierne a la posibilidad de conocer la realidad social y hace especial énfasis en la relación entre el estudioso y la realidad estudiada. «Todas las manifestaciones o los niveles de conocimiento, observar, percibir, determinar, interpretar, negar o afirmar, presuponen la relación del hombre con el mundo y son posibles sólo sobre la base de esta relación» [Abbagnano, 1971, 162]. La respuesta a esta cuestión depende, a su vez, de la respuesta dada a la cuestión ontológica precedente. Si el mundo social existe como tal, con independencia de la actuación humana, será lícito aspirar a conocerlo con distancia objetiva, sin miedo a alterarlo en el curso del proceso cognitivo. Las características, es decir, la «forma», que el conocimiento puede adoptar están estrechamente ligadas a la respuesta que demos a esta cuestión: éstas pueden ser «leyes naturales» deterministas dominadas por las categorías de causa-efecto, leyes menos imperativas, generalizaciones varias, por ejemplo, los tipos ideales weberianos, o ausencia de generalizaciones, es decir, sólo conocimientos específicos y contingentes.

---

[1] *Ontología:* parte de la filosofía que estudia el ser como tal; del griego *óntos* (ser, ente) y *lógos* (discurso, reflexión).

[2] *Epistemología:* reflexión sobre el conocimiento científico, del griego *epistème* (conocimiento cierto).

La cuestión metodológica[3]. Es la cuestión del «cómo», es decir, cómo se puede conocer la realidad social. Concierne a la instrumentación técnica empleada en el proceso cognitivo. De nuevo, las respuestas a esta cuestión están estrechamente relacionadas con las respuestas a las cuestiones precedentes. Si consideramos que la realidad social es un objeto externo que no se ve afectado por el proceso cognitivo del científico será más plausible utilizar técnicas de manipulación, por ejemplo, el experimento, el control de las variables, etc., que si consideramos que existe una interacción entre el investigador y lo investigado.

Por tanto, las tres cuestiones están relacionadas entre sí, no sólo porque las respuestas para cada una de ellas se influyen entre sí, sino también porque a veces es difícil distinguir los límites entre ellas. Nosotros intentaremos hacerlo para procurar la máxima claridad didáctica y expositiva. En efecto, resulta difícil distinguir las concepciones sobre la naturaleza de la realidad social de las reflexiones sobre la posibilidad de conocerla, y éstas de las técnicas para hacerlo. En realidad, esta dependencia mutua se encuentra en la propia definición de paradigma científico, que, como hemos visto, implica tanto una visión teórica, como una orientación de los procedimientos de investigación.

1.3. POSITIVISMO

La tabla 1.1 contiene un resumen que facilita la comparación entre los distintos paradigmas en relación con las cuestiones fundamentales que acabamos de presentar. Conviene notar que las columnas de los paradigmas son tres, en lugar de dos, porque hemos presentado dos versiones del positivismo. La versión original del siglo XIX, hoy probablemente en desuso incluso para los empiristas más perseverantes, y su reformulación, construida en el siglo XX para subsanar sus limitaciones evidentes. Hemos considerado oportuno presentar también el paradigma positivista original, por razones de carácter histórico, por tratarse de la perspectiva que ha acompañado el nacimiento de las ciencias sociales, y en particular de la sociología, y porque a partir de su crítica se podrán comprender mejor las características de los otros dos paradigmas.

La sociología nace, por tanto, dentro del marco del pensamiento positivista. En el momento en que, a mediados del siglo pasado, el hombre empieza a plantearse preguntas sobre la realidad social como tal y a transformarla en objeto de estudio, la nueva disciplina adopta el modelo que era el paradigma de las ciencias naturales. Los fundadores de la disciplina, en especial Comte y Spencer, tenían en común una fe ingenua en los métodos de las ciencias naturales. El paradigma positivista se

[3] *Metodología:* del griego *métodos* (vía por la cual, método). Cuestión metodológica en cuanto tiene que ver con los «métodos» de la investigación social, entendidos como cuerpo orgánico de técnicas. También la podríamos haber llamado (quizás más correctamente) «cuestión tecnológica», en cuanto tiene por objeto las técnicas, pero hemos preferido evitar este término, puesto que ya ha adoptado otro significado en el lenguaje común.

Tabla 1.1. Características de los paradigmas base de la investigación social

	Positivismo	Postpositivismo	Interpretativismo
Ontología	Realismo ingenuo: la realidad social es «real» y conocible (como si se tratara de una cosa).	Realismo crítico: la realidad social es «real» pero conocible sólo de un modo imperfecto y probabilístico.	Constructivismo: el mundo conocible es el de los significados atribuidos por los individuos. Relativismo (realidad múltiple): estas realidades construidas varían en la forma y en el contenido entre individuos, grupos, culturas.
Epistemología	Dualismo/objetividad. Resultados ciertos. Ciencia experimental en busca de leyes. Objetivo: explicación. Generalizaciones: leyes «naturales» inmutables.	Dualismo/objetividad modificados. Resultados probablemente ciertos. Ciencia experimental en busca de leyes. Multiplicidad de teorías para el mismo hecho. Objetivo: explicación. Generalizaciones: leyes provisionales, susceptibles de revisión.	Ausencia de dualismo/objetividad. El investigador y el objeto de investigación están relacionados entre sí. Ciencia interpretativa en busca de significado. Objetivo: comprensión. Generalizaciones: enunciados de posibilidad; tipos ideales.
Metodología	Experimental-manipuladora. Observación. Separación observador-observado. Predomina el método inductivo. Técnicas cuantitativas. Análisis «por variables».	Experimental-manipuladora modificada. Observación. Separación observador-observado. Predomina el método deductivo (comprobación de hipótesis). Técnicas cuantitativas, sin descartar las cualitativas. Análisis «por variables».	Interacción empática entre investigador y objeto de investigación. Interpretación. Interacción observador-observado. Inducción (el conocimiento procede de la realidad estudiada). Técnicas cualitativas. Análisis «por casos».

FUENTE: Adaptación de Guba y Lincoln [1994, 109].

puede resumir muy brevemente como sigue: el estudio de la realidad social utilizando el marco conceptual, las técnicas de observación y medición, los instrumentos de análisis matemático y los procedimientos de inferencia de las ciencias naturales.

Examinemos con más detalle el contenido de esta definición. El marco conceptual hace referencia a las categorías de «ley natural», de causa-efecto, de comprobación empírica, de explicación, etc. Las técnicas de observación y medición se refieren al uso de variables cuantitativas incluso para fenómenos de naturaleza cualitativa, los procedimientos de medición aplicados a ideologías, capacidades mentales, estados psíquicos, por ejemplo, la medición de actitudes, los tests de inteligencia, etc. Por último, los procedimientos de deducción hacen referencia al proceso que permite formular hipótesis sobre lo desconocido partiendo de lo conocido, es decir, el paso de la observación particular a la ley general, la utilización de la teoría para anticipar resultados, la generalización a toda la población partiendo de una muestra de la misma.

Gracias a Comte, el profeta del positivismo sociológico del siglo XIX, sabemos que la adopción de la perspectiva positivista constituye en cada ciencia el punto final de un camino que ha pasado previamente por las fases teológica y metafísica. No todas las disciplinas realizan este camino de manera simultánea. Primero se impuso en las ciencias de naturaleza inorgánica, como la astronomía, la física o la química; a continuación en las de la naturaleza orgánica, como la biología; y finalmente, en una secuencia que va de las disciplinas más simples a las más complejas, en el terreno más complejo por definición: la sociedad. De esta manera se constituyó una nueva ciencia, la sociología, o ciencia positiva de la sociedad. De acuerdo con esta idea, la ciencia es universal y el método científico es único. Las ciencias de la sociedad no son distintas de las ciencias de la naturaleza y el modo de pensar positivo que ha llevado a conquistas tan significativas en astronomía, física o biología, está destinado a triunfar también cuando se pasa de los objetos naturales a los sociales: la religión, la política o el trabajo.

El primer intento de aplicar esta perspectiva teórica global a la investigación empírica lo realizó Durkheim. Comte, como afirma el mismo Durkheim, «no estudió cosas, sino conceptos. Es cierto que Comte proclamó que los fenómenos sociales son hechos naturales sometidos a leyes naturales, reconociendo con ello de manera implícita su carácter de cosas [...]. Pero cuando, dejando a un lado estas generalidades filosóficas, intenta aplicar su principio y deducir del mismo la ciencia que contiene, sus objetos de estudio son las ideas» [Durkheim 1895].

Durkheim, en cambio, se esforzó por traducir los principios del pensamiento positivo en praxis empírica; él es el primer «científico social», el primer sociólogo positivista verdadero. Y su praxis empírica se fundamenta en la teoría del «hecho social». Desde las primeras páginas de las *Reglas del método sociológico* escribe que «la primera regla, que es también la más importante, impone *considerar los hechos sociales como cosas*» [*ibidem,* 56]. Para Durkheim los hechos sociales son:

> ... modos de actuar, de pensar, de sentir que presentan la [...] propiedad de existir fuera de las conciencias individuales [...]. Cuando llevo a cabo la tarea de [...] marido o ciudadano [...], cumplo con deberes que están definidos, fuera de mí y de

> mis actos, en el derecho y en las costumbres. Aun cuando éstos armonizan con mis sentimientos, y yo siento interiormente su realidad, ésta no es sin embargo menos objetiva: no los he hecho yo, sino que los he recibido mediante la educación. Lo mismo puede decirse de las creencias y las prácticas religiosas, [...] el sistema de signos que utilizo para expresar mi pensamiento, el sistema monetario que empleo para pagar mis deudas [...], las prácticas que se siguen en mi profesión.
>
> [*Ibídem,* 50-51]

Estos hechos sociales, aunque no son entidades materiales, tienen las mismas propiedades que las «cosas» del mundo natural. De ello se derivan dos consecuencias. Por una parte, los hechos sociales no están sujetos a la voluntad del hombre, sino que ofrecen resistencia a su intervención, lo condicionan y lo limitan. Por otra, precisamente como los fenómenos del mundo natural, los hechos también funcionan según sus propias reglas. Poseen una estructura determinista que el hombre, mediante la investigación científica, puede descubrir. El mundo social, como el mundo natural, está, por tanto, regulado por leyes. Y en ambos casos, estas leyes se pueden estudiar de manera objetiva. De ahí el supuesto de que existe una unidad metodológica fundamental entre el mundo natural y el mundo social, pese a que sus objetos de estudio son diferentes: ambas se pueden estudiar con la misma lógica de investigación y el mismo método, de ahí el nombre de «física social» que se suele atribuir al estudio de la sociedad.

La primera afirmación es, por tanto, que existe una realidad social fuera del individuo. La segunda dice que esta realidad social es objetivamente conocible. La tercera que se puede estudiar con los mismos métodos de las ciencias naturales. «Nuestra regla, escribe Durkheim, no implica [...] ninguna concepción metafísica, ninguna especulación sobre el fundamento de los seres, pero reclama que el sociólogo adopte la actitud de los físicos, químicos y fisiólogos que se adentran en un terreno aún inexplorado de su dominio científico [...]. Nuestro llamado positivismo no es más que una consecuencia de este racionalismo» [*ibídem,* 33].

Faltaría hablar sobre el modo de proceder de este conocimiento. En el positivismo éste es fundamentalmente inductivo, entendiendo por *inducción* «el paso de lo particular a lo universal»[4], el proceso por el cual, a partir de la observación empírica, de la identificación de repeticiones o regularidad en las fracciones de realidad estudiadas, se llega a generalizaciones o leyes universales. El procedimiento inductivo lleva implícito el supuesto de que existen un orden y una uniformidad en la naturaleza, unos principios organizadores de carácter universal. La misión del científico es, precisamente, descubrir estos principios. Esta idea ha dominado durante mucho tiempo las ciencias naturales e incluso se la ha identificado con el método experimental.

La concepción positivista de la sociedad asume plenamente la premisa de que

---

[4] Según Stuart Mill, la inducción es «la operación mental con la que deducimos que lo que sabemos que es verdad en uno o más casos aislados será verdad en todos los casos similares a los primeros por ciertos aspectos determinables» [Mill, 1843: 288].

los fenómenos de la vida social obedecen a leyes naturales inmutables: el científico social, según Durkheim, es como un explorador que «penetra en lo desconocido; debe sentir que está rodeado de hechos que obedecen a leyes insospechadas, como podían ser las leyes de la vida cuando la biología aún no existía; debe estar preparado para descubrir leyes que lo sorprenderán y lo desconcertarán» [*ibídem,* 37].

Por último, con respecto a la «forma» de este conocimiento, no hay dudas sobre la posibilidad de lograr la identificación y formulación de estas leyes de la naturaleza, su demostración y «comprobación»; leyes que en su expresión más completa adoptarán la forma de un nexo causa-efecto. «Desde el momento en que la ley de causalidad ha sido comprobada en otros ámbitos de la naturaleza y ha ampliado de manera progresiva su imperio del mundo físico-químico al mundo biológico, es lícito admitir que esa ley también se cumple en el mundo social; y hoy es posible añadir que las investigaciones realizadas sobre la base de este postulado tienden a confirmarlo» [*ibídem,* 159].

Podría afirmarse que en los métodos de las ciencias naturales son demasiado evidentes los elementos de lo que hemos llamado «fe ingenua» del paradigma positivista. En la base de las distintas articulaciones del positivismo queda siempre una especie de entusiasmo por el conocimiento «positivo» de tipo científico y la consideración de la ciencia y su método como único conocimiento válido y eficaz en todos los campos del saber humano.

1.4. NEOPOSITIVISMO Y POSTPOSITIVISMO

Dentro de la visión positivista tuvo lugar a lo largo del siglo XX un proceso continuo de revisión y adaptación, movido justamente por la conciencia de sus propias limitaciones intrínsecas y el intento de superarlas. La tranquilizadora claridad y linealidad del positivismo del siglo XIX deja libre el terreno para un positivismo del siglo XX mucho más complejo, articulado, y en algunos casos no exento de contradicciones y puntos oscuros, sin dejar no obstante de cumplir con ciertas premisas de base, como el realismo ontológico, es decir, «el mundo existe independientemente de nuestro conocimiento», y el punto de vista dominante que defiende la observación empírica como medio para conocer dicho mundo. No entraremos en los detalles de este proceso ni en sus fases históricas, y nos limitaremos a mencionar el «neopositivismo», bajo cuya etiqueta se suele catalogar el planteamiento reinante entre los años treinta y los años sesenta, y el «postpositivismo», con el que se suele identificar su evolución a partir de finales de los años sesenta [5]. Nos limitaremos, por tanto, a perfilar los cambios más importantes que se produjeron respecto a la ortodoxia positivista que acabamos de presentar.

Una de las primeras revisiones del positivismo del siglo XIX la realizó la escuela conocida con el nombre de *positivismo lógico,* que ha dado origen al neoposi-

[5] Hay que remontarse a Kuhn, Lakatos y Feyerabend para encontrar la crítica al neopositivismo denominada postpositivismo.

Resumen 1.2. LAS RESPUESTAS DEL POSITIVISMO A LAS TRES CUESTIONES DE FONDO

Ontología: realismo ingenuo. Esta posición se deriva de todo lo dicho sobre la «codificación» de la realidad social, y se puede expresar de manera sintética mediante dos proposiciones: *a*) existe una realidad social objetiva, externa al hombre, ya sea éste el estudioso o el objeto de estudio; y *b*) esta realidad es conocible en su esencia real [6].

Epistemología: dualista y objetivista; ley natural. Se afirma la posibilidad del conocimiento gracias a dos hechos: *a*) el estudioso y el objeto estudiado se consideran entidades independientes (dualismo); y *b*) el estudioso puede estudiar el objeto sin influir en él o recibir su influencia (objetividad). La investigación se produce mediante «un espejo unidireccional». El conocimiento adopta la forma de «leyes» basadas en las categorías de causa y efecto. Éstas existen en la realidad externa con independencia de los observadores («leyes naturales»), y la tarea del científico es «descubrirlas». No se teme al riesgo de que los valores del investigador puedan deformar su lectura de la realidad social, o a la inversa. Esta posición, que elimina los valores a favor de los hechos, se deriva de la noción de hecho social como *dato* externo *que no se puede modificar.*

Metodología: experimental y manipuladora. Los métodos y técnicas de la investigación positivista, así como su planteamiento de fondo, se basan en el empirismo clásico de las ciencias naturales. Es decir, se adoptan todos los principios del método empírico: *a*) el modo de proceder inductivo; y que plantea principios generales a partir de casos particulares observados, *b*) la formalización matemática, que, aunque no siempre es viable, representa la aspiración fundamental del científico positivista. La técnica idónea sigue siendo, con las limitaciones de su aplicabilidad a la realidad social, el experimento, que se basa en la manipulación y el control de las variables implicadas y en la distancia entre el observador y lo observado.

[6] Algunas cuestiones epistemológicas relativas a la posibilidad de conocer la realidad se tratarán en la discusión de la problemática ontológica, sobre la esencia de la realidad. Lo hacemos así para facilitar la comprensión del lector no familiarizado con estos conceptos, y también porque las dos problemáticas están intrínsecamente ligadas. Esto se entenderá mejor cuando estudiemos el paradigma interpretativo.

tivismo. Esta corriente surgió en torno a las discusiones de un grupo de estudiosos de distintas disciplinas que en la segunda mitad de los años veinte constituyeron el llamado «Círculo de Viena», algunos de cuyos miembros más destacados son los filósofos Schlick y Carnap, el matemático Hahn, el economista Neurath y el físico Frank. Más adelante se formaría un grupo similar en Berlín, integrado por Reichenbach, Herzberg, Lewin, Hempel y otros. La emigración a Estados Unidos de algunos representantes acreditados de esta escuela, debido a las persecuciones nazis, y la sintonía entre este planteamiento y el pragmatismo estadounidense, favorecieron notablemente la difusión del pensamiento neopositivista y su influencia sobre otras disciplinas, incluida la sociología. Fue precisamente en Estados Unidos donde surgió, a partir de los años treinta, un rico filón de investigación sociológica empírica.

Esta nueva perspectiva concede un papel central a la crítica de la ciencia, redefiniendo la tarea de la filosofía, que debe abandonar el terreno teórico para pasar al análisis crítico de lo que se elabora en las teorías de cada disciplina. De hecho, Schlick augura la llegada de un tiempo en el que no habrá libros de filosofía, sino que todos los libros estarán escritos «de modo filosófico». Por ello se rechazan las «grandes cuestiones» y las metafísicas carentes de sentido, o «pseudoproblemas», en la medida en que no se pueden demostrar. En cambio se dedica la máxima atención a los problemas metodológicos de cada ciencia, a la sintaxis de su lenguaje y sus elaboraciones teóricas, a la crítica de sus tesis y, no en último lugar, a los procedimientos de validación de las elaboraciones conceptuales mediante la verificación empírica.

Por todo lo dicho parece evidente que esta corriente de pensamiento concede una gran importancia a las cuestiones epistemológicas, y resulta, por tanto, comprensible la influencia que tuvo sobre la metodología de las ciencias, incluidas las ciencias sociales. Conviene recordar que uno de los postulados del neopositivismo es la difundida convicción de que el sentido de una afirmación se deriva de la posibilidad de formular una definición operativa a partir de ella para comprobar su validez. La siguiente frase resume muy bien este punto de vista: «el significado de una proposición es el método para su verificación».

¿Qué ha significado esta concepción de la ciencia y del conocimiento científico para la investigación social, y cuáles han sido sus consecuencias sobre los procedimientos operativos y sobre las técnicas de investigación? La consecuencia principal fue el desarrollo de una forma totalmente nueva de hablar de la realidad social, mediante un lenguaje extraído de las matemáticas y de la estadística, que Paul F. Lazarsfeld, el principal exponente de la metodología empírica neopositivista en sociología, llamó *lenguaje de las variables*. Cada objeto social, empezando por el individuo, se definía de manera analítica a partir de una serie de atributos y propiedades, las «variables», y se reducía a ellos; y los fenómenos sociales se estudiaban en términos de relaciones entre variables. La variable, con sus características de neutralidad, objetividad y operatividad matemática, se convertía así en la protagonista del análisis social.

La investigación social, por tanto, se «despersonaliza», y el lenguaje de las variables, con la medición de los conceptos, la distinción entre variables dependientes e independientes, la cuantificación de sus relaciones mutuas, la formula-

ción de modelos causales, ofrecía un instrumento formal que permitía ir más allá del «lenguaje cotidiano notoriamente impreciso [en un proceso de] clarificación y purificación del discurso [que es] fundamental para el científico social; [...] nuestro conocimiento se podía organizar de forma manipulable [...], y el sentido común se podía reformular en proposiciones verificables empíricamente» [Lazarsfeld y Rosenberg, 1955, 2, 11]. De este modo todos los fenómenos sociales se podían registrar, medir, relacionar, elaborar y formalizar, y las teorías se podían confirmar o invalidar de manera objetiva y sin ambigüedades.

Pero ya nada podía ser como antes. La concepción de la ciencia del siglo XX estaba muy lejos de las sólidas certezas del positivismo del siglo XIX, en el que predominaba la concepción «mecánica» de la realidad, la seguridad en las leyes inmutables, la fe en el progreso científico. El origen de este nuevo marco de referencia filosófico-científico está sobre todo en ciertos avances de las ciencias naturales, y en particular de la física, en los primeros años del nuevo siglo. La mecánica cuántica, la relativización del espacio y el tiempo propuesta por Einstein, el principio de indeterminación de Heisenberg, por citar sólo algunos de los cimientos de la nueva física, introducen elementos de probabilidad e incertidumbre en aspectos cruciales, como el concepto de ley causal, la objetividad-inmutabilidad del mundo exterior e incluso las categorías clásicas del espacio y el tiempo.

Las teorías acaban perdiendo la marca imperativa de las leyes deterministas para asumir la connotación de la probabilidad. El momento crucial de este cambio se encuentra en el paso de la física clásica (newtoniana) a la cuántica. Según la mecánica cuántica, en la física elemental existen procesos, los llamados saltos cuánticos, que no se pueden analizar según los mecanismos causales tradicionales, porque son hechos absolutamente imprevisibles gobernados por leyes probabilísticas. De esta forma se pierde la certidumbre de la ley, se destruye el ideal clásico de la ciencia como sistema invariable de verdades necesarias.

Las teorías científicas ya no están destinadas a explicar los fenómenos sociales mediante esquemas de naturaleza lógica restrictiva, y la ley determinista es sustituida por la ley probabilística, que implica la presencia de imprevisión, de perturbaciones y fluctuaciones. Si este supuesto de indeterminación probabilística es válido para el mundo natural, será aun más válido para el mundo social, el mundo del lenguaje, del pensamiento, de la interacción entre humanos.

Un elemento importante incorporado al pensamiento científico en su evolución a partir del modelo positivista inicial es la categoría de refutabilidad, considerada como criterio de validación empírica de una teoría o hipótesis. Ésta establece que la comparación entre la teoría y el dato empírico no puede hacerse de forma positiva, mediante la «comprobación» o verificación de que los datos confirman la teoría, sino que debe realizarse de forma negativa, comprobando que los datos empíricos «no invalidan» la teoría, es decir, que no contradicen la hipótesis, y, por tanto, son compatibles con ella. La comprobación positiva no puede servir como criterio porque los mismos datos podrían ser compatibles con diversas hipótesis.

De este planteamiento se deriva un sentido de provisionalidad de las hipótesis teóricas, que *nunca son definitivamente válidas* y siempre están expuestas al riesgo de una posible refutación. Con esto se derrumba, como escribe Popper, el ídolo de

la certidumbre. «El viejo ideal científico del *epistème,* del conocimiento absolutamente cierto, demostrable, ha resultado ser un ídolo. La exigencia de la objetividad científica hace ineludible que cada afirmación de la ciencia quede necesariamente y *para siempre en un estado de provisionalidad*» [1934: 280]. El hombre no puede conocer, sólo conjeturar. Podemos ilustrar esto con una afirmación atribuida a Einstein: «en la medida en que nuestras proposiciones son ciertas, no dicen nada sobre la realidad, y en la medida en que dicen algo, no son ciertas».

Por último, llegamos a las novedades más recientes de la orientación postpositivista. Se ha consolidado la convicción de que la observación empírica, la propia percepción de la realidad, no es una fotografía objetiva, sino que depende de la teoría, es decir, está cargada de teoría[7]. En otras palabras, incluso el mero registro de la realidad depende de la ventana mental del investigador, de sus condicionamientos sociales y culturales. Es decir, partiendo de que la realidad existe con independencia de la actividad cognitiva y la capacidad perceptiva humana, el acto de conocer está condicionado por las circunstancias sociales y el marco teórico en que se sitúa.

La afirmación de que no existe una separación clara entre los conceptos teóricos y los datos observados derroca incluso la última de las certezas positivistas, la de la objetividad del dato obtenido y neutralidad del lenguaje observador.

Hay que decir, sin embargo, que este proceso de alejamiento de la ortodoxia positivista inicial, primero con el neopositivismo y después con el postpositivismo, no conlleva una anulación del espíritu empirista. Se mantiene, también en los planteamientos más recientes, la importancia del método «científico» en la investigación social, y la analogía de fondo entre el método de las ciencias sociales y el de las ciencias naturales.

El positivismo moderno, cuando afirma que las leyes, tanto naturales como sociales, son sólo probables y susceptibles de revisión, cuando defiende la naturaleza provisional del conocimiento científico y el condicionamiento teórico sobre la propia observación, ha recorrido un largo camino desde la interpretación ingenua de las leyes deterministas del positivismo inicial. Pero, aunque ha perdido la certidumbre, no renuncia al fundamento empirista.

El nuevo positivismo redefine los presupuestos iniciales y los objetivos de la investigación social, pero los procedimientos empíricos, aunque son reinterpretados, siguen apoyándose en las bases de la operacionalización, la cuantificación y la generalización. Y dado que este manual trata sobre técnicas de investigación, éste es el punto que nos interesa. Los procedimientos operativos, las modalidades de obtención de datos, las operaciones de medición, las elaboraciones estadísticas, no sufren variaciones de fondo. Se adoptan técnicas cualitativas, pero sin restar importancia a las cuantitativas. Se adopta una mayor cautela respecto a las conclusiones, pero las técnicas utilizadas son siempre las mismas.

Podemos concluir la reflexión sobre el paradigma positivista retomando la columna de la tabla 1.1 que refleja las posiciones del positivismo moderno sobre las tres cuestiones de fondo. Éstas quedan sintetizadas en el resumen 1.3.

---

[7] La expresión procede de Hanson (1958).

1.5. INTERPRETATIVISMO

1.5.1. Los inicios

Hemos presentado dos versiones del paradigma positivista: la perspectiva inicial que apareció en el siglo XIX, y su revisión crítica, surgida en los años treinta y continuada a partir de los años setenta. Podríamos proceder de forma casi paralela con el paradigma que vamos a presentar ahora, donde el «casi» se refiere sólo al desfase temporal de la analogía. Para ello introducimos primero la «sociología comprensiva» de Max Weber, que propuso la elaboración metodológica y los primeros intentos de investigación empírica a principios del siglo XX, de manera similar al papel ejercido por Durkheim en el positivismo. A continuación vendría la reinterpretación del planteamiento original que se fue concretando a partir de los años sesenta sobre todo en la sociología estadounidense, y que dio lugar a las diversas corrientes del interaccionismo simbólico, la sociología fenomenológica y la etnometodología, diferentes entre sí pero con el elemento común de la importancia que conceden a la interacción individual.

No obstante, este paralelismo no nos parece adecuado, porque entre la visión weberiana inicial y sus posteriores desarrollos no existe la discontinuidad que hemos constatado en el paso del positivismo del siglo XIX al del siglo XX. Preferimos, por tanto, incluir bajo un único título estos dos bloques históricos de aproximación a la investigación social, clasificando bajo el término general de «interpretativismo» todas las perspectivas teóricas para las que la realidad no puede ser sólo observada, sino que debe ser «interpretada».

¿Cómo nace esta nueva perspectiva de la ciencia social? Si el positivismo tiene sus orígenes en las culturas francesa e inglesa del siglo XIX (basta citar a Auguste Comte, John Stuart Mill y Herbert Spencer), y desarrolla su vertiente sociológica sobre todo en la primera (Durkheim), la crítica más radical y sistemática a su planteamiento aparece en el contexto del historicismo alemán.

Hay que remontarse al filósofo alemán Wilhelm Dilthey para encontrar la primera crítica al cientificismo comtiano en nombre de la autonomía de las ciencias humanas, en el sentido de que no son equiparables con las ciencias naturales. Dilthey aúna en su polémica tanto el idealismo hegeliano como el positivismo comtiano, que tendrían en común «la misma fe en la historia como progreso necesario a través de fases también necesarias», a partir de un concepto de la historia entendida como «construcción del hombre, de la interrelación individual, y que devuelve por tanto al hombre su libertad, aun reconociendo la realidad del condicionamiento» [Izzo, 1974, 366].

Dilthey realiza, en su *Introducción a las ciencias del espíritu* (1883), una célebre distinción entre «ciencias de la naturaleza» y «ciencias del espíritu», basada precisamente en la relación entre el investigador y la realidad estudiada. El objeto de las ciencias naturales lo constituye una realidad externa al hombre, que se mantiene como tal en el curso del proceso cognitivo, que adopta las formas de la explicación (leyes de causa-efecto, etc.). En las ciencias del espíritu, al no existir esta separación entre el observador y la realidad estudiada, el conocimiento sólo puede producirse mediante

Resumen 1.3. LAS RESPUESTAS DEL NEO Y POST POSITIVISMO A LAS TRES CUESTIONES DE FONDO

Ontología: realismo crítico. De manera análoga al positivismo, se presume la existencia de una realidad externa al hombre, pero, a diferencia de lo defendido en ese paradigma, ésta sólo se puede conocer de un modo imperfecto, ya sea por la inevitable imprecisión de todo conocimiento humano, ya sea por la propia naturaleza de sus leyes, que tienen carácter probabilístico. Este punto de vista también se ha llamado «realismo crítico»: realismo, porque acepta que existen relaciones causa-efecto fuera de la mente humana (en la realidad), y crítico, por subrayar esa postura de sospecha continua y esa propensión a la duda que debe adoptar el científico frente a todos los logros de la ciencia.

Epistemología: dualismo-objetividad modificados; leyes de alcance limitado, probabilísticas y provisionales. Con respecto a la cuestión de la relación estudioso-estudiado, ya no se defiende el dualismo como separación y ausencia de interferencia entre las dos realidades. Se tiene conciencia de los elementos de perturbación que introduce en el objeto estudiado el sujeto que lo estudia y de la reacción que se puede derivar de ellos. La objetividad del conocimiento sigue siendo el criterio de referencia, pero sólo se puede lograr de forma aproximada. En el proceso cognitivo cobra valor el procedimiento deductivo, por el mecanismo de refutación de hipótesis. El objetivo sigue siendo formular generalizaciones en forma de leyes, aunque de alcance limitado, probabilísticas en su potencial de aplicación y provisionales en el tiempo.

Metodología: experimental-manipuladora modificada. Las fases operativas de la investigación siguen siendo fundamentalmente las planteadas por el neopositivismo, para asegurar la separación entre el investigador y el objeto estudiado (experimentos, manipulación de las variables, entrevistas cuantitativas, análisis de fuentes estadísticas, etc.), aunque no se descartan los métodos cualitativos. La comunidad científica concede importancia a la crítica de las nuevas teorías y los procesos de confirmación de las mismas (si los resultados se repiten es más probable que sean válidas las leyes generales que los explican).

un proceso muy distinto, el de la *comprensión*. Por tanto, según Dilthey, *explicamos* la naturaleza y *entendemos* la vida psíquica.

En el mismo periodo otro estudioso alemán proponía una distinción alternativa, aunque estrechamente ligada a la anterior. Nos referimos a Windelband, que intro-

duce la diferencia entre «ciencias nomotéticas», orientadas a la identificación de leyes generales, y «ciencias idiográficas», que buscan la singularidad de los fenómenos, su cualidad de únicos e irrepetibles [8].

1.5.2. Max Weber: objetividad y orientación hacia la individualidad

Esta nueva perspectiva entra en el campo de la sociología con Max Weber. Dilthey había hablado genéricamente de las «ciencias del espíritu», concediendo un papel privilegiado a la historiografía. Weber lleva el concepto de *Verstehen* a la sociología, revisando el planteamiento original de Dilthey. La preocupación de Weber es no caer en el individualismo subjetivista y el psicologismo al adoptar el principio del *Verstehen*; quiere salvar la objetividad de la ciencia social en los aspectos de «neutralidad valorativa», es decir, independencia de los juicios de valor, y posibilidad de formular enunciados con un cierto carácter general, aun partiendo de una «orientación hacia lo particular».

En relación con el primer aspecto, Weber insistió durante toda su vida en la necesidad de que las ciencias histórico-sociales estén libres de todo juicio de valor. Hay que decir, sin embargo, que en Weber destaca más la conciencia del problema, agudizada por su intensa actividad política y replanteada en sus últimos años de vida por las cuestiones éticas derivadas de la amenaza de conflicto mundial, que su capacidad para encontrar una respuesta unívoca para el mismo. La *neutralidad valorativa* de las ciencias histórico-sociales sigue siendo un fundamento ineludible. «La capacidad de distinguir entre conocer y valorar, es decir, entre el cumplimiento del deber científico de percibir la realidad de los hechos y el cumplimiento del deber práctico de defender los ideales propios. Éste es el principio que debemos observar con más firmeza» [Weber, 1904].

Aunque las ciencias histórico-sociales no pueden admitir supuestos de valor, según Weber, es inevitable que éstos intervengan en la elección de los problemas a estudiar, asumiendo un papel orientador en la investigación. Incluso privados de su función de valoración, los valores siguen estando presentes, realizando lo que podríamos llamar una «función de selección». Weber se adhiere en este punto a la distinción de Rickert entre «juicio de valor» y «relación de valor». Esta última «no es un principio de valoración, sino un principio de elección: sirve para establecer un campo de investigación, en el que la búsqueda se realiza de manera objetiva para lograr la explicación causal de los fenómenos [...]. La relación de valor viene a designar la dirección particular del interés cognitivo que motiva la investigación, es decir, el *punto de vista* específico en el que ésta se sitúa, delimitando su propio campo» [Rossi, 1985, 24-25].

Una vez establecida la neutralidad valorativa como primera condición para la objetividad de las ciencias sociales, quedan por definir los términos de la segun-

---

[8] Los dos términos derivan del griego. *Nomotético:* que tiene validez general (*nómos:* ley; *thétes:* que establece), *idiográfico:* que tiene rasgos individuales (*ídios:* propio; *graphos:* dibujo).

da condición, entendida como capacidad de formular principios dotados de un cierto carácter de *generalidad.* Las ciencias sociales, según Weber, se distinguen de las ciencias naturales no por el objeto, ni porque su fin sea estudiar los fenómenos sociales aisladamente, pues también buscan formas de generalización, sino por su «orientación hacia el individuo» [*ibídem,* 21].

Esta orientación afecta, en primer lugar, al método. Y para Weber el método es el de «comprender». Pero al definir lo que él entiende por *Verstehen,* Weber rechaza toda relación con la psicología. No se trata de perspicacia psicológica ni de iluminación repentina, sino de comprensión racional de las motivaciones para los actos; no se trata de intuición, sino de «interpretación»: entender el objetivo de los actos, la magnitud del propósito y la intención de la acción humana. El componente de identificación con el otro, presente en el concepto de *Verstehen,* también tiene como fin la interpretación racional: identificarse con el otro «para comprender». Comprender las motivaciones de sus actos, el significado subjetivo atribuido por el individuo a su conducta, porque toda conducta, hasta la que parece más ilógica, tiene una racionalidad propia, un «sentido» interior. En palabras de Raymond Boudon:

> Para Weber, comprender un acto de un individuo significa hacerse con medios de información suficientes para analizar las motivaciones que lo inspiraron. El observador *comprende* el acto del sujeto observado en el momento en que puede llegar a la conclusión de que, en la misma situación, él habría actuado sin duda del mismo modo [...]. Como puede verse, la *comprensión* en el sentido weberiano supone que el observador pueda *ponerse en el lugar* del que actúa, pero de ningún modo implica que la subjetividad del segundo sea evidente para el primero [...]. En realidad la noción weberiana de *comprensión* representa una actitud muy cercana a la que los manuales de lógica designan con la expresión «inducción policial», que consiste en buscar las motivaciones que no son directamente accesibles con el método de reconstrucción de los hechos a partir de información procedente de varias fuentes, con el efecto de cruzarlos entre sí.
>
> [Boudon, 1984; 31, 51]

¿Cómo se puede conseguir la objetividad con este enfoque individualista? Si el punto de partida está representado por el individuo y por el sentido subjetivo de sus actos, ¿cómo es posible llegar a un conocimiento objetivo con carácter de generalidad? Estamos ante la segunda condición para la objetividad de las ciencias histórico-sociales enunciada más arriba.

La concepción weberiana del *tipo ideal* es la solución a este problema. Para Weber los tipos ideales son formas de actuación social que pueden encontrarse de modo recurrente en la conducta de los individuos. Esas formas se reconstruyen mediante un proceso de abstracción que, aislando ciertos elementos presentes en los datos empíricos, los coordina dentro de un marco coherente. El tipo ideal es, por tanto, *una abstracción que nace del reconocimiento empírico de la uniformidad.*

> El concepto tipo-ideal, escribe Weber, se obtiene mediante la *acentuación* unilateral *de uno o de varios puntos de vista,* y la conexión de una serie de fenóme-

> nos *particulares* conocidos y comunes, presentes en mayor o menor medida [...], dentro de un marco *conceptual* unitario. Dada su pureza conceptual, este marco nunca puede encontrarse empíricamente en la realidad; se trata de una *utopía,* y la labor histórica debe encargarse de determinar en cada caso la mayor o menor distancia entre la realidad y ese marco ideal [...]. Éste tiene el significado de un concepto, *límite ideal* puro, con el que la realidad se debe *medir* y *comparar*, a fin de ilustrar determinados elementos significativos de su contenido empírico.
>
> [Weber, 1904, en Rossi, 1974, 124-125]

El tipo ideal weberiano se extiende a todos los campos de lo social y puede situarse a diversos niveles de generalidad en la escala que abarca desde el individuo hasta el conjunto de la sociedad. Weber ofreció ejemplos formulando «tipos ideales» en relación con estructuras sociales, como el capitalismo, instituciones, como la burocracia, la iglesia, la secta, las formas de poder, y el comportamiento del individuo, por ejemplo, la conducta racional.

Los «tipos ideales», escribe Weber, no se deben «confundir con la realidad [...]. Han sido construidos de un modo ideal heurístico» [Weber, 1922a]; son «ideales» en el sentido de que son construcciones mentales del hombre, y cumplen una función «heurística» en el sentido de que orientan su conocimiento. Son cáscaras vacías, «ficciones exentas de vida», como los definió Schutz; no tienen un equivalente concreto en la realidad, sin embargo, son modelos teóricos que ayudan al investigador a interpretarla. Por ejemplo, es probable que ninguno de los tres tipos ideales de poder identificados por Weber, poder carismático, tradicional y racional-legal, haya existido nunca en su forma pura. El tipo ideal es una construcción racional clara, coherente, exenta de ambigüedad. En cambio la realidad es mucho más compleja, contradictoria y desordenada: todas las formas de poder carismático que han existido en la historia, aunque se podrían clasificar bajo uno de los tres «tipos» weberianos, no representan de manera plena y exclusiva a ninguno de ellos, y tienen también elementos y aspectos de los otras dos.

Las regularidades que el investigador busca y encuentra en su interpretación de la realidad social no son las «leyes», en el sentido que les da la sociología positiva. Para Weber, «el número y el tipo de causas que han determinado un hecho concreto cualquiera, es siempre *infinito* [...], y la cuestión causal, cuando se trata de la *individualidad* de un fenómeno, no es una cuestión de *leyes,* sino una cuestión de *conexiones* causales concretas [...], la posibilidad de seleccionar entre la infinidad de elementos determinantes» [Weber 1922b]. Por tanto, no debemos hablar de leyes, sino de conexiones causales, o, utilizando una expresión de Boudon, de *enunciados de posibilidad.* («Si sucede A, entonces, en la mayoría de los casos, se produce también B» [Boudon 1984: 75]). En definitiva, el objetivo de establecer los factores determinantes de un hecho social o una conducta individual concretos resulta inalcanzable; en cambio, sí se puede conseguir el objetivo de determinar las condiciones que los hacen posibles.

A las leyes causales positivistas dotadas de *generalidad* y *obligatoriedad* (si bien atenuada por el aspecto probabilístico en la interpretación neopositivista), se

contraponen enunciados y conexiones, marcados por las características de *especificidad* y *posibilidad.*

1.5.3. Desarrollos posteriores

Hemos dedicado especial atención a Weber porque el gran sociólogo alemán se anticipó prácticamente a todos los temas desarrollados más adelante por ese rico filón de teoría e investigación sociológica que dio lugar a la sociología fenomenológica (Husserl y Schutz), el interaccionismo simbólico (Mead y Blumer) y la etnometodología (Garfinkel y Cicourel), consolidados en el contexto de la sociología estadounidense en los años sesenta. Todas estas perspectivas tienen en común los rasgos fundamentales del planteamiento weberiano: la fuerte convicción antideterminista, la oposición a la filosofía de la historia, y a toda forma de evolucionismo. De este modo, la diferencia esencial, «ontológica», entre las ciencias naturales y las ciencias sociales es la irreductibilidad de las segundas a los métodos de investigación de las primeras. La crítica a todo intento de explicar la acción humana a partir del sistema social y los factores de condicionamiento presentes en el mismo también procede de las aportaciones de Weber. En definitiva, en todos estos enfoques está presente, en términos positivos, la convicción de que en el centro de cada fenómeno social, así como de la actividad del sociólogo, se encuentra la «acción individual dotada de sentido».

Sin embargo, Weber no llevó su posición metodológica a sus últimas consecuencias. El autor elaboró estos conceptos en sus escritos metodológicos, pero tanto en su reflexión teórica como en la investigación empírica que desarrolló se movió siempre en un plano macrosociológico, en una perspectiva de historia comparada, guiado por el interés de comprender fenómenos macroestructurales como la economía, el Estado, el poder, la religión o la burocracia. En cambio, el movimiento que nace en Estados Unidos en los años sesenta desarrolla la perspectiva weberiana en la dirección microestructural. Si la sociedad se construye a partir de las interpretaciones de los individuos, y su interacción es la que crea las estructuras sociales, para entender la sociedad debemos observar dicha interacción. De aquí procede el descubrimiento y la valoración de un campo de investigación completamente nuevo para la sociología, el mundo de la vida cotidiana, cuyo estudio se había ignorado y considerado no científico hasta aquel momento.

La diferencia entre esta concepción y el marco de referencia positivista es clara. El enfoque «subjetivista» es ante todo una reacción al planteamiento «objetivista» positivista, que, al reducir la realidad social y la acción humana a algo objetivamente, silenciaba precisamente la dimensión individual, el aspecto «humano», lo que distingue el mundo de los humanos del mundo de las cosas. Precisamente lo que según el enfoque positivista alteraba la investigación «científica» y debía excluirse de ésta, los componentes individuales, la motivación, la intención, los valores, el libre albedrío, es resumen, la dimensión subjetiva no perceptible con los instrumentos cuantitativos, se convierte aquí en el principal objeto de investigación. En esta diversidad del objeto se basa también, desde el punto de vista interpretativo, su presunta superioridad con relación al método positivista.

El defensor del paradigma interpretativo no sólo corrobora la autonomía y diversidad de las ciencias histórico-sociales frente a las naturales, sino que afirma su superioridad, pues sólo el método *Verstehen* permitiría esa inteligibilidad profunda que subyace al conocimiento del mundo social y de los actos.

Las diferencias de fondo implican también una diferencia en las técnicas y los procedimientos de investigación. Este es el aspecto que más nos interesa tratar en este libro. La idea es que si la vida humana es en esencia distinta de la vida del mundo natural, entonces tendrá que estudiarse con métodos diferentes de los positivistas. El subjetivista no puede, por tanto, adoptar «el lenguaje de las variables», ni en la fase de observación empírica, debido a la importancia de los componentes intencionales y subjetivos, que por definición escapan a la cuantificación objetiva, ni en la fase de análisis de los datos, porque no puede imaginarse el análisis de la conducta humana en términos de interacción de componentes separados (las variables). La unidad del ser humano impide que el todo se pueda reducir a la suma de las partes.

El enfoque subjetivista elaborará, por tanto, procedimientos y técnicas propias de la observación y el análisis de la realidad empírica, que darán lugar al cuerpo de la denominada «investigación cualitativa». Más adelante volveremos sobre ella para tratarla con más detalle. De momento concluimos nuestra presentación del paradigma interpretativo proponiendo de nuevo una síntesis que se presenta en el resumen 1.4. Este cuadro desarrolla con más detalle la última columna de la tabla 1.1.

1.6. CRÍTICAS, POSTURAS RADICALES, Y NUEVAS TENDENCIAS

Hemos trazado las líneas maestras de lo que consideramos los dos paradigmas que han definido históricamente, al margen de diversas especificaciones, el planteamiento y las técnicas de investigación social, haciendo referencia a los conceptos fundamentales y a quienes podríamos considerar los padres fundadores de la sociología. Faltaría hablar de las críticas a los dos enfoques y de algunas posturas más radicales que han surgido recientemente.

Por lo que respecta al paradigma positivista, hemos visto que prestaba una gran atención y cuidado a la formulación y el desarrollo de técnicas y procedimientos empíricos, sobre todo en el período del neopositivismo. La radicalización de esta tendencia ha derivado en una especie de empirismo antiespeculativo, dominado por el mito del método y el dato, en el que el objetivo del científico social ya no es formular teorías y después comprobarlas empíricamente, sino recopilar y describir datos con la ingenua ilusión de «que los datos hablarán por sí mismos».

Se trata de un proceso de reducción progresiva (de ahí la acusación de «reduccionismo»), que ha atravesado varias fases. En primer lugar, se reducen los límites de la exploración teórica, destacando los problemas de la comprobación o confirmación de hipótesis en la actividad científica, es decir el *ars probandi,* en detrimento de la profundización sobre los descubrimientos, es decir, el *ars inveniendi.* En

segundo lugar, se desplaza la atención del contenido al método. En la tensión entre teoría y comprobación empírica, dar prioridad a la segunda ha significado excluir del ámbito de las consideraciones teóricas aquellos interrogantes que no fueran traducibles de manera inmediata y simple a procedimientos empíricamente comprobables, en un proceso de reducción de la complejidad teórica hasta la trivialización más extrema.

En definitiva, este proceso de reducción se concluyó desplazando la atención del método al dato, de los problemas de la operacionalización de conceptos a los problemas prácticos de la obtención y el análisis de los datos, incluso con técnicas estadísticas avanzadas. Como escribe Gallino, «los resultados inmediatos de las investigaciones fueron los que los críticos del neopositivismo sociológico podían esperar: una gran cantidad de datos, recopilados, medidos y clasificados con sumo cuidado, pero no armonizados entre sí, sin conexiones significativas entre ellos, incapaces de proporcionar un conocimiento adecuado del objeto al que hacen referencia» [Gallino, 1978, 457].

El enfoque del paradigma que hemos denominado interpretativista tampoco está exenta de problemas. La crítica se ha centrado no tanto en los planteamientos weberiano iniciales, como en sus intérpretes y seguidores, que llevaron la «orientación hacia el individuo» weberiana original a sus últimas consecuencias. En Weber aún está muy presente la tensión por la superación de la subjetividad: él no excluye la posibilidad de alcanzar formas de generalización cognitiva, es decir, los tipos ideales, y una parte importante de sus exposiciones metodológicas está orientada a conciliar causalidad y comprensión; por último, aunque parte del individuo, no descuida la problemática sistémica ni la dimensión institucional de la sociedad.

Las nuevas corrientes de la reflexión sociológica desarrolladas a partir de los años sesenta, también conocidas como sociología «neocomprensiva», en cambio, hicieron hincapié en el carácter subjetivista del planteamiento weberiano inicial. Así mismo, desplazaron la atención al mundo de la vida cotidiana y la interacción intersubjetiva. Este desplazamiento ha dado más fuerza a las dos críticas del paradigma interpretativo formuladas por sus oponentes.

La primera crítica sostiene que el subjetivismo extremo excluye la posibilidad misma de que exista la ciencia, y en particular la ciencia social. El hecho de que la acción humana tenga siempre un elemento de singularidad, su irreductibilidad a componentes sociales externos al individuo, o incluso la afirmación de que la realidad es una construcción subjetiva, todo ello niega la posibilidad de llegar más allá del individuo, de formular generalizaciones supraindividuales, y niega, por tanto, la objetividad del conocimiento. La ausencia de separación entre el estudioso y lo estudiado también contribuye a esta negación. La imposibilidad para el investigador de trascender al objeto de investigación significa la imposibilidad del conocimiento objetivo, así como del control intersubjetivo, que subyace al propio concepto de ciencia, es decir, el hecho de que otro investigador pueda llegar al mismo resultado elaborando los mismos datos u otros.

En segundo lugar, se acusa al enfoque interpretativo, siempre por la atención que presta al individuo, de excluir de sus intereses lo que debería ser el objeto de la reflexión sociológica por excelencia: las instituciones. Con ello corre el peligro de

Resumen 1.4. LAS RESPUESTAS DEL INTERPRETATIVISMO A LAS TRES CUESTIONES DE FONDO

Ontología: constructivismo y relativismo (realidad múltiple). «Constructivismo»: el mundo susceptible de ser conocido es el del significado atribuido por los individuos. La perspectiva constructivista radical excluye la existencia de un mundo objetivo, dado que cada individuo produce una realidad propia. La posición más moderada no se plantea el problema de la existencia de una realidad externa, al margen de las construcciones individuales, pero afirma que sólo se pueden conocer éstas últimas. «Relativismo»: estos significados, estas construcciones mentales, varían entre los individuos; e incluso cuando no son individuales, sino que son comunes a grupos de individuos, varían entre las diferentes culturas. No existe una realidad social universal válida para todos los hombres («realidad absoluta»), sino que existen múltiples realidades («realidad múltiple»), como múltiples y diversas son las perspectivas con las que los hombres perciben e interpretan los hechos sociales.

Epistemología: no dualismo y no objetividad; tipos ideales, enunciados de posibilidad. Tiende a desaparecer la separación entre estudioso y objeto de estudio, así como la separación entre ontología y epistemología. Por contraste con la perspectiva positivista, la investigación social se define como «una ciencia experimental en busca de leyes» [Geertz, 1973: 5], en la que las categorías centrales son las de valor, significado y finalidad. Para lograr su objetivo, que es la comprensión de la conducta individual, la ciencia social puede servirse de abstracciones y generalizaciones: los tipos ideales y los enunciados de posibilidad.

Metodología: interacción empática entre el estudioso y lo estudiado. La interacción entre el estudioso y lo estudiado en las fases empíricas de la investigación ya no se considera negativa, sino que, por el contrario, representa la base del proceso cognitivo. Si el objetivo es comprender el significado atribuido por el sujeto a la propia acción, las técnicas de investigación sólo pueden ser cualitativas y subjetivas, entendiendo por subjetivas variables según la forma que adopte la interacción estudioso-estudiado. El conocimiento se produce mediante un proceso de inducción, o sea de «descubrimiento en la realidad», por parte de un estudioso que se acerca a ella libre de prejuicios y de teorías preconcebidas.

dejar fuera de su ámbito de estudio aspectos de la sociedad que, aunque también tienen su origen en la interacción social, han adquirido autonomía respecto de los individuos y sus elecciones. La misma crítica de fondo se formula de manera lige-

ramente distinta acusando a la sociología fenomenológica, a la etnometodología y al interaccionismo simbólico de haberse ceñido a lo «microsociológico», de haber limitado sus propios intereses a la interacción, a las relaciones interpersonales, de no querer o no poder aplicar sus planteamientos a problemas que trasciendan a los hechos de la vida cotidiana.

Todo lo expuesto hasta ahora se basa en las grandes corrientes del pensamiento sociológico que han fundado la disciplina, han dado forma a sus métodos y técnicas de investigación y han guiado la investigación sociológica desde su nacimiento hasta mediados de los años setenta, aproximadamente.

El último cuarto del siglo XX representó un periodo de desafíos en la historia de la investigación social, después de una década tumultuosa en la historia de las sociedades occidentales, los años sesenta, con los movimientos por los derechos civiles, las protestas estudiantiles, las revueltas urbanas con trasfondo racial, la lucha contra la pobreza y las desigualdades, el desarrollo y la difusión del movimiento feminista. En este contexto, la teoría sociológica y la investigación social asumieron una relevancia y una popularidad hasta entonces desconocidas, y la sociología, por dedicarse a la reflexión sobre estas transformaciones, pareció renovar su «misión» social. Se consolidaron nuevas perspectivas teóricas, surgieron los enfoques neomarxistas, neoweberianos, se desarrolló una teoría crítica y nuevas perspectivas radicales opuestas a la complaciente alianza entre neopositivismo y funcionalismo que había dominado hasta ese momento la elaboración teórica y la investigación social.

En el mismo periodo, además de estas perspectivas macro, empezó a desarrollarse la denominada «microsociología», que representaba a diversas escuelas de pensamiento y visiones teóricas, como la sociología fenomenológica, el interaccionismo simbólico, la etnometodología, la hermenéutica, que tenían en común su interés por los hechos de la vida cotidiana, las microinteracciones entre individuos, las dinámicas interpersonales, en lugar de centrarse en las grandes transformaciones históricas y los procesos relativos a la sociedad.

Este abandono de las grandes perspectivas teóricas y las explicaciones de carácter general trajo consigo los primeros brotes de una crítica generalizada a cualquier explicación teórica y el cuestionamiento del carácter científico de la sociología. Esta tendencia se radicalizó hacia los años noventa, en una corriente intelectual de características variadas y poco precisas, que a veces se resume bajo la etiqueta de «postmodernismo».

Podemos definir de manera escueta esta corriente haciendo referencia a aquello que rechaza. El modernismo, entendido como heredero directo de la ilustración es un ejercicio crítico de la razón en torno al hombre, la naturaleza y la sociedad, se caracteriza por su confianza en la ciencia, basada en el orden y la racionalidad, en la simplicidad de la explicación científica y en el carácter cumulativo del saber. El *postmodernismo* implicaría «ir más allá», y en contra, de los logros del modernismo, en una crítica que podemos resumir brevemente en cuatro puntos: *a*) *rechazo de las teorías generales,* universales, acusadas de totalitarismo uniformador, de imperialismo cultural, de negación y represión de las diferencias entre las sociedades con el fin de perpetuar las aspiraciones hegemónicas del mundo y la cultura occidentales, a favor de planteamientos y lenguajes teóricos diversos, en

una exaltación de la fragmentación y de la no unidad de la explicación científica; *b*) *rechazo de la racionalidad,* la linealidad y la simplicidad del conocimiento científico, a favor de la paradoja, la contradicción, la opacidad y la idea de que existen múltiples facetas alternativas e irreconciliables; *c*) *exaltación de las diferencias,* la diversidad, la multiplicidad de verdades particulares y contextuales, rechazo del carácter cumulativo de la ciencia; y *d*) exaltación de lo «otro», lo distinto, las minorías, identificación con los oprimidos, consideración del «poder» como categoría explicativa y origen de todas las relaciones y estructuras sociales.

No profundizaremos más en este breve y simplificado esbozo de las nuevas tendencias de la ciencia social contemporánea y sus posibles paradigmas. Nuestro único interés es describir los que han sido o son los paradigmas fundamentales de las ciencias sociales que han influido y dado forma, generado y nutrido las estrategias de investigación empírica, sus métodos y sus técnicas.

Desde este punto de vista, las nuevas tendencias y perspectivas que han agitado el mundo de la ciencia social y de la sociología en el último cuarto de siglo no han provocado grandes revoluciones en las técnicas de investigación social, a excepción de una mayor legitimidad y aceptación de las técnicas de investigación cualitativas y un uso mucho más frecuente de las mismas, sin añadir, en ningún caso, grandes novedades a las técnicas empleadas. De este tema nos ocuparemos en el capítulo siguiente.

1.7. SÍNTESIS DEL CAPÍTULO 1

1. Toda ciencia «madura» tiene, en cada momento de su historia, un paradigma propio. Éste representa una especie de «visión que orienta», una perspectiva teórica aceptada por la comunidad de científicos que dirige la investigación, determinando qué estudiar y formulando hipótesis explicativas de los fenómenos observados.
2. En las ciencias sociales, los dos paradigmas que han orientado la investigación desde sus inicios han sido el *positivismo* y el *interpretativismo.* A fin de compararlos, nos hemos preguntado cómo ha respondido cada uno de ellos a las tres cuestiones de fondo de la investigación social: la cuestión ontológica (¿existe la realidad social?), la cuestión epistemológica (¿se puede conocer?) y la cuestión metodológica (¿cómo podemos conocerla?).
3. El paradigma positivista se consolidó en la investigación social en el siglo XIX, tras el gran éxito de las ciencias naturales. El positivismo sostenía que la realidad social se debía estudiar con la misma lógica investigadora y el mismo método utilizados en las ciencias naturales; de ahí el nombre de «física social» atribuido al estudio de la sociedad.
4. En el curso del siglo XX la inspiración positivista sufre ciertas transformaciones para superar sus propias limitaciones. Según los paradigmas neopositivista y postpositivista, las teorías sociales no se deben expresar en forma de leyes deterministas, sino en términos probabilísticos. Toda afirmación teórica asume

por tanto una connotación de provisionalidad y queda siempre sujeta a la posibilidad de ser desmentida. Además, en la comunidad científica se desarrolla la conciencia de que cualquier observación empírica no es una descripción objetiva de la realidad, sino que está «cargada de teoría», en el sentido de que hasta el simple acto de registrar una realidad se ve afectado por los esquemas mentales del investigador. Esta revisión del positivismo no renuncia sin embargo a sus fundamentos empíricos ni a su fe en la cuantificación y en la posibilidad de generalizar los resultados, por lo que dedica mucha atención a las técnicas de investigación empírica, al denominado «lenguaje de las variables», procedente de las matemáticas y la estadística.

5. Según el interpretativismo, existe una diferencia «epistemológica» fundamental entre las ciencias sociales y las naturales, ya que la realidad social no puede ser simplemente observada, sino que se debe «interpretar». En las ciencias naturales el objeto de estudio es una realidad externa al investigador y así se mantiene durante toda la investigación, por ello el conocimiento adopta la forma de «explicación». En las ciencias sociales no existe esta separación entre el observador y lo observado, y el conocimiento sólo puede alcanzarse a través de un proceso totalmente distinto, el de la comprensión *(Verstehen)*. Esta diferencia fundamental conlleva procedimientos y técnicas de investigación distintos. El enfoque subjetivista, al no poder adoptar el «lenguaje de las variables», ha tenido que desarrollar sus propias técnicas de observación y modalidades de análisis del material empírico, que constituyen el cuerpo de la denominada «investigación cualitativa».
6. La radicalización de ambos enfoques puede tener graves consecuencias. Si el planteamiento positivista se lleva hasta sus últimas consecuencias podría producirse una reducción progresiva del alcance de la investigación, que quedaría centrada en el dato empírico y limitada a una mera descripción del mismo. Por su parte, el subjetivismo extremo pone en duda la propia existencia de la ciencia social, excluyendo la posibilidad de realizar generalizaciones supraindividuales y afirmando que la realidad es una construcción subjetiva. Un desarrollo reciente del paradigma interpretativo ha sentado las bases del denominado «postmodernismo», una suerte de rechazo del concepto tradicional de ciencia, entendida como orden y racionalidad, simplicidad y generalización, a favor de la paradoja, la contradicción y la exaltación de las diferencias.

1.8. LECTURAS COMPLEMENTARIAS

Una selección de ensayos que permiten profundizar en las perspectivas teóricas que han orientado los métodos de investigación social es la antología de G. Ritzer y B. Smart, *Handbook of Social Theory* (Sage, 2001, 552 pp.). En particular, para profundizar en los temas que hemos tratado en este capítulo, recomendamos los siguientes ensayos: M. Gane, *Durkheim's Project for a Sociological Science;* P. Halfpenny, *Positivism in Twentieth Century;* S. Whimster, *Max Weber: Work and Interpretation;* K. L. Sanstrom, D. D. Martín y G. A. Fine, *Symbolic Interactionism at the End of the*

Century; S. Crook, *Social Theory and the Posmodern.*

Una presentación introductoria sobre el planteamiento paradigmático de la tradición de investigación cuantitativa y cualitativa puede encontrarse en el primer capítulo de A. Tashakkori y C. Teddlie, *Mixed methodology: Combining Qualitative and Quantitative Approaches* (Sage 1998, 185 pp.); una guía más completa sobre las diferentes respuestas de las escuelas de pensamiento clásicas y contemporáneas a los dilemas fundamentales de la investigación social, se encuentra en N. Blaikie, *Approaches to Social Inquiry* (Polity Press, 1993, 238 pp.).

Para enmarcar las perspectivas teóricas de las que derivan las diversas corrientes de la investigación cualitativa actual, recomendamos el ensayo de Y. S. Lincoln y E. G. Guba, *Paradigmatic Controversies, Contradictions, and Emerging Confluences,* de Denzin y Lincoln (2000). En la misma línea y con algo más de profundidad se sitúa la obra de J. F. Gubrium y J. M. Holstein, *The New Language of Qualitative Method* (Sage, 1997, 244 pp.), en la que los autores identifican cuatro «idiomas» (naturalismo, constructivismo social, emocionalismo, postmodernismo), de los que se derivan los diversos enfoques de la investigación cualitativa reciente. Para una reflexión sobre las tendencias actuales de la investigación social, pero desde el punto de vista del paradigma neopositivista, que no rechaza el uso de técnicas cualitativas, remitimos al libro de J. Goldthorpe, *On Sociology: Numbers, Narratives, and the Integration of Research and Theory* (Oxford U. P., 2000, 337 pp.).

CAPÍTULO **2**

INVESTIGACIÓN CUANTITATIVA E INVESTIGACIÓN CUALITATIVA

2.1. EL PARADIGMA NEOPOSITIVISTA: *CRIME IN THE MAKING*, DE SAMPSON Y LAUB 32

2.1.1. Hipótesis 33
2.1.2. Diseño de la investigación 34
2.1.3. Observación empírica y resultados del análisis 34

2.2. EL PARADIGMA INTERPRETATIVO: *ISLANDS IN THE STREET*, DE JANKOWSKI 36

2.2.1. Diseño de la investigación y recopilación de datos 37
2.2.2. Hipótesis 37
2.2.3. Interpretación del material empírico 39

2.3. INVESTIGACIÓN CUANTITATIVA E INVESTIGACIÓN CUALITATIVA: COMPARACIÓN 40

2.3.1. Planteamiento de la investigación 40
2.3.2. Recopilación de los datos 47
2.3.3. Análisis de los datos 49
2.3.4. Resultados 52

2.4. DOS FORMAS DIFERENTES DE CONOCER LA REALIDAD SOCIAL 59

2.5. SÍNTESIS DEL CAPÍTULO 2 62

2.6. LECTURAS COMPLEMENTARIAS 63

En este capítulo se presentan dos ejemplos típicos de investigación cuantitativa y cualitativa, a fin de ofrecer una perspectiva general de ambos enfoques. Después de ello, estudiaremos las distintas fases de una investigación social, comparando brevemente el modo en que las abordan los enfoques cuantitativo y cualitativo. Concluiremos el capítulo con una consideración sobre la complementariedad de las dos técnicas de investigación como fuentes de conocimiento social.

2.1. EL PARADIGMA NEOPOSITIVISTA: *CRIME IN THE MAKING*, DE SAMPSON Y LAUB

El debate entre investigación cuantitativa y cualitativa ha sufrido diversas vicisitudes: tras la viva y fructífera confrontación de los años veinte y treinta, cuando ambos criterios generaron ideas de gran valor y contribuyeron al avance de la disciplina, pensemos, sobre todo en el marco de la investigación cualitativa, en la denominada «Escuela de Chicago», el debate entró en una fase de latencia, con el dominio en los años cuarenta, cincuenta y sesenta, de la perspectiva cuantitativa. Durante todos estos años la investigación cualitativa se consideró una especie de hija ilegítima de la ciencia social, y el etnógrafo apenas superaba la imagen de un buen periodista, negándosele la categoría de científico social.

En los años sesenta la discusión se replanteó a partir de una serie de contribuciones teóricas importantes [Goffman, 1959; 1967; Schutz, 1967; Glaser y Strauss, 1967; Blumer, 1969][1]. Pero fue a partir de los años ochenta, y durante los noventa, en un proceso que se ha prolongado hasta hoy, cuando la investigación cualitativa se desarrolló de manera activa avivando el debate metodológico, con la elaboración de nuevas técnicas de recopilación y análisis del material empírico, la creación de nuevas revistas científicas dedicadas específicamente al enfoque cualitativo, y una producción de reflexiones, propuestas, investigaciones y manuales sin precedentes.

---

[1] Como prueba del cambio de clima experimentado con respecto a la investigación cualitativa entre los años sesenta y ochenta podemos citar las dos conclusiones de una investigación, difundida en 1962 y publicada de nuevo veinte años después. En la primera edición del estudio, realizado mediante observación participante en un suburbio italoamericano de Boston, el autor Herbert Gans casi parece excusarse al definir su trabajo como «un estudio no científico, en tanto que no proporciona lo que Merton ha llamado la indiscutible evidencia de las hipótesis [...]. La validez de mis resultados se apoya exclusivamente en mi juicio personal [...]. Este estudio es una exploración inicial de una comunidad que pretende dar una visión global de la misma, pero no puede documentar plenamente todos los resultados [...]. Muchas de las hipótesis formuladas se podrán confirmar con investigaciones sociales más sistemáticas» [Gans, 1962, 350]. En la segunda edición del libro, que actualiza y amplía la anterior, el autor siente la necesidad de añadir una posdata sobre las «virtudes de la observación participante», donde afirma que si tuviera que reescribir el apéndice escrito en 1962, «eliminaría la conclusión a modo de disculpa [debida al hecho de que] entonces escribía en un periodo en el que aún se veneraba la «ciencia» y los científicos sociales pensaban que el método de las ciencias naturales se debía aplicar también a la realidad social [...]. En aquella época yo era particularmente consciente de los límites de la observación participante, pero hoy sé que ningún método de investigación puede alcanzar los ideales que tanto respeto me inspiraban» [Gans, 1962, 2ª ed., 1992, 414]. Y concluye afirmando la superioridad del método de observación participante, por lo menos en temas como los tratados en su estudio.

En este capítulo desarrollaremos la parte inferior de la tabla 1.1 del capítulo anterior: tras presentar lo que hemos definido como los «paradigmas fundadores» de la investigación social, queremos ilustrar ahora sus aportaciones en el terreno de la metodología. Uno de los objetivos de ello es mostrar al lector no experto en investigación social los resultados de los dos planteamientos en términos de investigación empírica.

Empezaremos el capítulo con la descripción de dos investigaciones, una inspirada en el paradigma neopositivista y la otra en el paradigma interpretativo. Para destacar mejor las diferencias entre ambos enfoques, hemos elegido dos investigaciones sobre la misma temática, la delincuencia juvenil, con interrogantes bastante similares. A continuación, analizaremos las diferencias entre los dos enfoques.

En esta comparación entre las perspectivas cuantitativa y cualitativa hablaremos de dos «tipos ideales» de investigación, subrayando, en parte por motivos didácticos, las diferencias entre ambos puntos de vista. Reconocemos que entre los dos extremos, o entre los dos tipos «puros», existen posiciones intermedias y que, sobre todo en el criterio cuantitativo, es posible incorporar elementos de técnicas cualitativas. Por otra parte, el enfoque estrictamente interpretativo, en el que el investigador trata de «ver el mundo con los ojos del sujeto estudiado», exige una identificación y una implicación mayor, y conlleva un radio de acción más limitado comparado con el enfoque cuantitativo.

La primera investigación se puede atribuir a la inspiración y las técnicas de la corriente neopositivista, representada por la obra de Robert J. Sampson y John H. Laub, *Crime in the Making. Pathways and Turning Points Through Life,* publicada en 1993 en Estados Unidos.

El origen de esta investigación contiene una curiosa coincidencia. Se trata del hallazgo, en los sótanos de la biblioteca de la Law School de la Universidad de Harvard, de unas 60 cajas que contenían el material original de una impresionante investigación longitudinal, recopilado durante 24 años, entre 1939 y 1963, por el matrimonio de Sheldon y Eleanor Glueck, y utilizado sólo parcialmente en sus publicaciones, entre las que se encuentra el clásico *Unraveling Juvenile Delinquency,* publicado en 1950. Ante este inesperado hallazgo, Sampson y Laub deciden volver a analizar los datos, mediante lo que suele definirse como un *análisis secundario*, para responder a las nuevas cuestiones que había planteado entre tanto el desarrollo de la teoría y la investigación a los estudiosos de las conductas desviadas de los menores.

2.1.1. Hipótesis

Los dos autores lamentan el hecho de que, dado que la mayoría de los delitos eran cometidos por adolescentes, los estudios de sociología criminal se hayan concentrado sobre todo en esta edad, descuidando por una parte la infancia, en la que según algunos habría que buscar el origen del comportamiento antisocial, y por otra la edad adulta, en la que momentos cruciales de la vida, como la incorporación al mercado laboral o el matrimonio, pueden conllevar cambios radicales en la actitud social del individuo. Este punto de vista implica la necesidad de dejar atrás los

estudios sincrónicos o «transversales» *(cross sectional)*, que «fotografían» a un grupo de individuos en un momento determinado, para pasar a estudios diacrónicos o «longitudinales», que siguen a un grupo de individuos durante un periodo de tiempo determinado, obteniendo información y datos sobre los mismos en distintos momentos de sus vidas.

A partir de aquí, Sampson y Laub discuten las tesis de aquellos que han estudiado la conducta criminal desde una perspectiva de ciclo de vida, y trazan las líneas fundamentales de una posible «teoría del control social informal por edades», en la que se discuten, para cada edad, las variables que se han venido considerando causantes de la conducta desviada (pobreza, familia desestructurada, infancia antisocial, etc.), y los mecanismos informales de control social operativos en ese momento del ciclo de vida. Todo ello, con una visión global que intenta superar «el reducido marco de las perspectivas sociológicas y psicológicas hasta ahora dominantes en criminología, en parte por limitarse a datos de adolescentes y a un solo periodo de sus vidas» [Sampson y Laub, 1993, 23], para orientar criminología hacia una perspectiva de «trayectoria vital».

2.1.2. Diseño de la investigación

La investigación de Sampson y Laub es un análisis de los datos recopilados por Sheldon y Eleanor Glueck más de treinta años antes. Estos autores reunieron información sobre 500 jóvenes, varones, blancos, que habían cometido algún delito y que cuando comenzó la investigación, en 1939, tenían entre 10 y 17 años, y sobre 500 jóvenes «normales» (que no había cometido ningún delito). Los primeros procedían de dos correccionales de Massachussets. Los segundos procedían de escuelas públicas de la misma zona, y fueron seleccionados según una cuidadosa correspondencia: por cada uno de los 500 jóvenes que habían cometido algún delito se eligió un joven normal con las mismas características de edad, origen étnico, barrio y cociente intelectual. La investigación comenzó en 1939; se hizo un seguimiento sistemático de los jóvenes entre 1939 y 1948, con entrevistas a ellos, sus familias y sus profesores (o empleadores). También se recopiló información de vecinos, trabajadores sociales, policías y jueces, y se consultaron documentos judiciales oficiales para registrar los posibles delitos cometidos.

2.1.3. Observación empírica y resultados del análisis

Para cuantificar los resultados los autores crearon el «índice de conducta desviada no oficial». Se registraron todas las conductas, no sólo las delictivas (robo, hurto, juegos de azar, daños, etc.), sino también las que podían considerarse «malas conductas» (fumar, beber, escaparse de casa, hacer novillos, etc.), referidas por los propios sujetos, sus padres y sus profesores; se recopiló información procedente de las distintas fuentes y se construyeron índices de desviación para cada conducta, así como un índice de desviación global (con valores comprendidos entre 1 y 26). Este índice representaba la «desviación no oficial», mientras que la «desviación oficial», determinada por los delitos denunciados ante la autoridad judicial,

estaba representada por una variable dicotómica (autor de delito / no autor de delito), aplicada a la muestra de 500 + 500 sujetos. Estos dos indicadores de desviación son las variables dependientes de la investigación.

La exposición de los resultados de la investigación de Sampson y Laub ocupa cinco capítulos del libro. Estos tratan sobre los siguientes temas: contexto familiar y delincuencia juvenil, papel de la escuela, el grupo de compañeros y hermanos/as, continuidad temporal de la conducta, relaciones sociales adultas y cambio en la conducta delictiva, y modelos comparados de delitos y desviación. Todos los capítulos tienen la misma estructura lineal, que contempla los siguientes puntos: *a*) marco teórico; *b*) observaciones empíricas; *c*) resultados del análisis, y *d*) vuelta a la teoría.

Para que el lector pueda comprender los procedimientos empleados para el análisis de los datos, revisaremos el primero de esos capítulos. Los autores distinguen entre «variables fundamentales» (o estructurales) y «variables de proceso» (o procesales). Las primeras son las variables clásicas (pobreza, disgregación familiar, casos de delincuencia en los padres, etc.), que se suelen utilizar en este tipo de estudios; las segundas hacen referencia a las «relaciones informales» (en este capítulo con la familia, en capítulos siguientes con la escuela, el trabajo, etc.), que para Sampson y Laub desempeñan un papel fundamental en el proceso que conduce a la desviación. A partir de esta distinción, los autores construyen un modelo teórico, según el cual las variables estructurales influirían sobre la conducta desviada de modo indirecto, a través de las variables de proceso, representadas por el vínculo-control familiar.

Una vez definido el marco teórico, los autores pasan a describir las *variables.* Distinguen nueve variables estructurales: «situación familiar», que comprende tres categorías, y toma el valor 1 cuando el joven ha crecido en una familia en la que uno o los dos padres estaban ausentes por divorcio, separación, abandono o muerte; «dimensión de la familia», expresada en el número de hijos; «nivel socioeconómico», con tres categorías, acomodado, marginal, dependiente (de ayuda externa); «nacimiento en el extranjero», «movilidad residencial», «trabajo de la madre», «desviación del padre», «desviación de la madre», etc. Identifican cinco «variables procesales familiares», ligadas a la relación afectiva con los padres, la posible aplicación de castigos físicos, la presencia/ausencia de supervisión materna y al rechazo, abandono u hostilidad por parte de los padres. La variable dependiente del modelo es la conducta desviada; como ya se ha indicado, según los casos, se tratará de desviación «oficial» o «no oficial».

Llegados a este punto, podemos pasar a los *resultados del análisis.* Los autores emplean el procedimiento estadístico de la regresión múltiple. Clasifican las variables en tres grupos: las instrumentales básicas, las procesales familiares y las dos dependientes (conducta desviada). Relacionan los tres grupos entre sí de dos en dos, y en todos los casos encuentran correlaciones elevadas: entre las variables básicas y las variables procesales (lo que significaría que las condiciones estructurales de la familia influyen en los vínculos afectivos y la relación pedagógica), entre las variables básicas y la desviación (la situación familiar precaria, pobre, etc., favorece la conducta desviada), entre las variables procesales y la desviación (el debilitamiento de los lazos familiares favorece la última).

Todas las relaciones empíricas encontradas entre las variables son resultados

esperados que concuerdan con la teoría. Pero quizás lo más importante de los resultados de estos autores es que cuando se analiza el modelo completo, es decir, las variables estructurales básicas y las variables procesales familiares, y la conducta desviada, el efecto de las variables estructurales básicas prácticamente desaparece. ¿Qué significa esto? Significa que las variables estructurales no tienen un efecto *directo* sobre la conducta desviada, sino que su acción está mediatizada por las variables procesales. Por ejemplo, una situación de disgregación familiar favorece el abandono por parte de los padres (ausencia de control, etc.), y éste, a su vez, facilita la aparición de conductas desviadas. Pero cuando hay control y cuidado de los padres hacia los hijos, cesa la influencia de la disgregación familiar sobre la desviación potencial de los hijos. Los autores calculan que un 73 por 100 del efecto de las variables estructurales sobre la delincuencia está mediatizado por las variables procesales familiares.

La última parte de la fase empírica es la *vuelta a la teoría*. Los autores llegan a la conclusión de que «los procesos familiares de control informal tienen un importante efecto inhibidor sobre la delincuencia de los adolescentes [...]. Es sorprendente, a la luz de estos resultados, que muchas explicaciones sociológicas del delito no tengan en cuenta a la familia [...], lo que ha generado una clara divergencia entre el sentido común popular [...] y los científicos sociales» [Sampson y Laub, 1993, 85, 97]. Estos resultados vienen a apoyar su «teoría integrada de los controles sociales informales», en lo que respecta a la primera fase, la relativa al control familiar.

En los capítulos siguientes, Sampson y Laub aplican un patrón de análisis relativamente similar al papel de la escuela, el grupo de compañeros, los hermanos, el trabajo, el matrimonio, y, por último, reformulan minuciosamente el modelo propuesto al principio, esta vez, comprobado empíricamente. De este modo llegan a lo que ellos denominan «modelo teórico dinámico del delito, la desviación y el control social informal a lo largo de la vida», en el que dividen los primeros 45 años de vida en 5 fases (infancia, de 0 a 10 años, adolescencia, de 10 a 17 años, transición a la adultez joven, entre 17 y 25 años, adultez joven, entre 25 y 32, transición a la madurez, de 32 a 45 años), y para cada fase destacan el papel de los factores que favorecen la aparición (o el mantenimiento) de la conducta desviada, y de los factores que, por el contrario, tienden a inhibirla.

Este modelo permite responder a los interrogantes planteados al principio, al formular las hipótesis. En particular, al reconocer la importancia de las «variables estructurales» y «procesales», aclaran que las segundas son las que en última instancia explican la mayor parte de la variación, tanto en la desviación en la adolescencia, como en el proceso de abandono de la delincuencia, a una edad posterior.

2.2. EL PARADIGMA INTERPRETATIVO: *ISLANDS IN THE STREET,* DE JANKOWSKI

En las conclusiones de la obra anterior, Sampson y Laub escriben: «Este libro ha querido enfrentarse al siguiente reto: ¿conseguiremos desarrollar y comprobar

empíricamente un modelo teórico capaz de *explicar* la conducta delictiva y la desviación en la infancia, en la adolescencia y en la edad adulta? [*ibídem,* 243]. En las conclusiones del libro que presentamos ahora (Martín Sánchez Jankowski, *Islands in the Street. Gangs and American Urban Society,* publicado en 1991 en Estados Unidos), el autor escribe: «Nosotros, en las ciencias sociales, así como en las políticas públicas, nunca hemos *comprendido* totalmente a las bandas urbanas. Para empezar, hemos fracasado en nuestro intento de *comprender* adecuadamente a los individuos que forman parte de estas bandas [...]; además, el hecho de que las bandas no hayan sido *entendidas* como organizaciones ha impedido la *comprensión* de su comportamiento» [Jankowski, 1991, 311, 314].

La mera diferencia léxica entre ambos fragmentos, con el «explicar» como objetivo en el primer caso, y la insistencia en el «comprender» en el segundo (las cursivas de las citas son nuestras), expresa de manera elocuente la diferencia de planteamiento metodológico que existe entre las dos investigaciones.

2.2.1. Diseño de la investigación y recopilación de datos

La investigación de Jankowski es un ejemplo de «observación participante». Sin embargo, por contraste con la mayoría de investigaciones realizadas con observación participante, Jankowski decide desde el principio que su estudio debe ser un estudio comparado, a fin de entender los elementos comunes y específicos de las bandas. Por ello decide estudiar bandas de ciudades distintas, en concreto, áreas metropolitanas de Los Ángeles, Nueva York y Boston, con connotaciones étnicas diversas y tamaños diferentes. En total estudió 37 bandas durante diez años, en los que participó plenamente en la vida de las bandas, integrándose en sus actividades y su vida cotidiana, hasta el punto de que fue herido en enfrentamientos con bandas rivales y detenido por la policía en repetidas ocasiones.

Para poder estudiar un número tan elevado de bandas tuvo que planificar rigurosamente su participación (lo cual no es habitual en este tipo de estudios). Al principio, cuando una nueva banda aceptaba someterse al estudio, pasaba con ella un mes completo; más adelante, entre cinco y diez días; y en los últimos tres años de la investigación, pasaba entre tres y seis días con cada una de las bandas.

La recopilación de datos, como es habitual en estudios de observación participante, se realizó mediante el registro de apuntes en un cuaderno, al cual se añadían resúmenes y reflexiones diarias y semanales (trataremos con más detalle este tema en el capítulo 9). El autor también utilizó de manera esporádica un magnetófono para grabar conversaciones.

2.2.2. Hipótesis

Por contraste con la investigación de Sampson y Laub, en este caso no encontramos esa reflexión teórica sistemática que conduce a la elaboración de hipótesis y su comprobación empírica. En el primer capítulo de su estudio, Jankowski no pasa revista a la literatura, comparando diferentes tesis, ni formula hipótesis, sino que se limita a presentar su investigación y a exponer las conclusiones que ha sacado de su

experiencia. Este planteamiento no procede de una elección personal del autor, sino que se deriva de las características del enfoque interpretativo, que, como ya hemos visto, adopta un método fundamentalmente inductivo, evitando todo condicionamiento teórico inicial, según la máxima de que la teoría se debe «descubrir» a través de la investigación.

La originalidad del enfoque de Jankowski reside en que para él las bandas no constituyen una desviación patológica de las normas sociales, sino más bien una elección racional. «Hasta ahora la teoría ha considerado a las bandas como una forma de desorganización social: en ausencia de orden, control y valores comunes, la banda ofrece ese orden social y esa seguridad que necesitan los jóvenes sin medios de los suburbios». En cambio, según Jankowski, los barrios más desfavorecidos de las ciudades estadounidenses «en realidad, están organizados, pero lo están en torno a una fuerte competencia, que llega a generar conflictos, por acaparar los escasos recursos. Surge pues un orden social alternativo al tradicional [...] y la banda aparece como respuesta organizada para aumentar la competitividad de sus miembros» [*ibídem,* 22].

La reflexión de Jankowski sigue tres líneas: el individuo y su relación con la banda, la banda como organización, y la banda y la comunidad. Respecto al individuo, el autor crea el concepto de carácter «individualista y rebelde». Éste denotaría un fuerte espíritu competitivo, que a menudo se manifiesta en forma de agresión física y que está presente en cualquier comportamiento del sujeto, una desconfianza hacia los demás, de la que se derivaría individualismo, aislamiento social y autosuficiencia, y, por último, una idea de la existencia que el autor define como «darwinista», según la cual la vida es una lucha en la que sobreviven sólo los más fuertes y de la que se deriva un fuerte instinto de supervivencia. El lector habrá notado en esta breve descripción los rasgos del «tipo ideal» weberiano.

Lo que este individuo espera de la sociedad podría encontrarlo en la banda. El autor define la banda (otro tipo ideal) como un sistema social cuasi privado (no es accesible para todos los individuos) y cuasi secreto (sólo sus miembros conocen plenamente sus actividades), ordenado según una estructura jerárquica con roles definidos, donde la autoridad asociada a éstos debe ser legitimada, cuyas actividades buscan no sólo ofrecer servicios sociales y económicos a sus miembros, sino también asegurar su propia supervivencia como organización, que no se preocupa por la legitimidad de sus objetivos, y está exenta de burocracia (no tiene un personal administrativo específico). El individuo denominado «individualista rebelde» solicita formar parte de la banda porque cree que le conviene, que puede obtener de ella beneficios económicos, de estatus y de poder. La banda decidirá si lo acepta como miembro si el individuo satisface las necesidades de la organización (prestigio, eficacia, servicios ofrecidos).

El autor prosigue su análisis de la banda como organización: sus estrategias para retener y comprometer a los miembros, la estructura de la dirección y los mecanismos de legitimación de la misma, los incentivos y sanciones para procurar la obediencia de los miembros.

Por último, Jankowski aborda la relación de la banda con la comunidad de la que forma parte: la fuerte cohesión interna no es suficiente para garantizar su super-

vivencia: ésta sólo se puede conseguir mediante la integración de la banda en la comunidad local. La banda debe ser aceptada por los residentes como parte integrante del barrio, al cual debe proporcionar unos servicios. A cambio, obtendrá su apoyo en términos de protección de la policía y de los «piratas rivales» (las otras bandas). La capacidad de establecer estos vínculos será uno de los principales determinantes de la supervivencia de la banda a largo plazo.

2.2.3. Interpretación del material empírico

En los capítulos siguientes, cinco sobre las dinámicas internas de la banda y sus relaciones con la comunidad local, y tres sobre sus relaciones con el mundo exterior (instituciones públicas, sistema judicial y medios de comunicación), el autor interpreta los datos recopilados con ayuda de las categorías presentadas en el marco teórico. Para mostrar el procedimiento empleado, citaremos brevemente el primer capítulo, titulado *Gang Involvement,* en el que el autor intenta responder a la siguiente pregunta: ¿quién entra en una banda y por qué?

Jankowski rechaza las cuatro respuestas que la literatura especializada ha dado a esta pregunta: los adolescentes entran en una banda, *(a)* porque proceden de familias desestructuradas en las que el padre está ausente y buscan una identificación con otros varones y con figuras de autoridad masculinas; *(b)* porque en la banda encuentran un sustituto de la familia; *(c)* porque ha sido expulsados del sistema escolar y no tienen formación suficiente para aspirar a un trabajo y no encuentran nada mejor que hacer que formar parte de una banda; o, por último, *(d)* para seguir el ejemplo de chicos mayores que los convencen fácilmente. El autor afirma: «no he encontrado confirmaciones empíricas para estas hipótesis [...]. Lo que he percibido es que los individuos que viven en barrios desfavorecidos entran en las bandas por varias razones, pero todas se derivan de lo que consideran que es mejor para ellos en ese momento determinado» [*ibídem,* 40]. A continuación, Jankowski enumera las motivaciones que ha encontrado.

«Incentivos materiales»: se entra en la banda para conseguir dinero con más regularidad y correr menos riesgos que si actuaran por separado, para poder disponer de ingresos en situaciones de emergencia (la banda suele proporcionar una especie de asistencia mutua entre sus miembros), y por las expectativas de oportunidades económicas futuras, por ejemplo, participar en actividades de tráfico de drogas. El autor ilustra los diversos casos con extractos de entrevistas; a continuación, reproducimos uno relativo a la segunda motivación:

> ... el club (la banda) estuvo ahí cuando lo necesité. A veces yo no tenía para comer. A mi familia le iba muy mal, no llegaba a fin de mes, ¡a veces ni siquiera había nada en la mesa! Así que necesitaba dinero hasta que consiguiera reunir algo por mi cuenta o hasta que la situación de mi familia mejorara un poco. Ellos [la banda] estaban allí para ayudarme. Antes de entrar ya sabía que ayudarían, y por eso entré. Ellos están ahí cuando les necesitas, y seguirán estándolo.
>
> [Street Dog, puertorriqueño, 15 años, miembro de una banda de Nueva York durante dos años, *ibídem,* 42]

«Diversión»: la banda es también un lugar de diversión y entretenimiento; a menudo tiene una especie de club con bar, videojuegos, cartas, máquinas recreativas, organiza fiestas y ofrece ocasiones para conocer chicas:

> ... entré por todo lo que ofrecían los *Bats* [el nombre de la banda] [...]. Había chicas guapas que se relacionaban con el grupo y estaban con ellos. Además, sus fiestas eran magníficas. La música era estupenda y la hierba [marihuana] era la mejor que se podía encontrar [...]. Eran los amos del barrio y por eso entré.
>
> [Fox, 23 años, miembro de una banda de Nueva York durante siete años]

«Refugio y escondite»: la banda ofrece anonimato a quien lo necesite, en un entorno muy competitivo:

> ... decidí entrar en la banda porque la banda te ofrece cobertura. ¿Sabes lo que quiero decir? Cuando, por ejemplo, alguien está metido en algún trapicheo y forma parte de una banda, es difícil que lo identifiquen, porque lo que quieres que diga la gente es: ése es uno de los chicos de la banda. ¿Me entiendes? La banda es para mantenerme resguardado.
>
> [Junior J., 17 años, Nueva York, *ibídem,* 44]

El autor continúa ilustrando de forma análoga, mediante fragmentos de entrevistas, el resto de motivaciones para entrar en una banda. El resultado es totalmente distinto al de la investigación anterior: el objetivo no es construir modelos causales, en los que las variables están conectadas entre sí por relaciones causa-efecto, sino más bien clasificaciones y tipologías, a partir de la experiencia vivida, en una clara aplicación del paradigma interpretativo.

2.3. INVESTIGACIÓN CUANTITATIVA E INVESTIGACIÓN CUALITATIVA: COMPARACIÓN

Usando como ejemplo las dos investigaciones anteriores, haremos una comparación analítica de las respuestas que han dado las técnicas cuantitativas y cualitativas a los problemas más relevantes de la investigación social. Estudiaremos en qué aspectos y en qué medida se diferencian entre sí ambos enfoques, no tanto en los planteamientos filosóficos y epistemológicos, sino en la tarea concreta de la investigación. Para agilizar la comparación, en la tabla 2.1 presentamos un resumen que consta de cuatro partes: planteamiento, recopilación de datos, análisis y resultados, y que representa una ampliación de la tercera fila de la tabla 1.1 («metodología»).

2.3.1. Planteamiento de la investigación

Si tuviéramos que indicar un elemento único y general que marque la diferencia entre los dos tipos de investigación, probablemente se encontraría en la estructuración de las fases que abarcan desde el planteamiento de la pregunta inicial hasta las

conclusiones finales. La investigación de Sampson y Laub destaca por su geometría: exposición del marco teórico, formulación de un «modelo» que se pueda comprobar empíricamente, diseño de la investigación, obtención de datos, análisis de los datos, revisión de la teoría. Se trata de una geometría circular que parte de la teoría y vuelve a ella, que además se plantea como si hubiera varios círculos menores incluidos en un círculo mayor en cada uno de los capítulos, que se dividen en cuatro partes: marco teórico, observaciones empíricas, resultados del análisis, conclusiones teóricas. Este esquema se percibe fácilmente en el capítulo que hemos resumido antes. Nótese que no se trata de un orden meramente expositivo, sino de un orden conceptual que ha guiado el modo de proceder de los autores y que surge de un concepto de investigación entendida como proceso racional y lineal.

La forma de proceder de Jankowski es totalmente distinta. Su obra carece de un marco teórico que proponga una serie de hipótesis empíricamente comprobables. Su capítulo teórico es un combinación de teorías y conclusiones y no distingue claramente la teoría de los resultados empíricos. Él no parte de hipótesis claras, sino que las va construyendo sobre la marcha. Por ejemplo, cuando rechaza las razones psicosociales tradicionales para entrar en una banda, en favor de una serie de motivaciones racionales, nos explica que llegó a esta conclusión a partir de las entrevistas realizadas, y no a partir de un prejuicio teórico. Por tanto, podemos hablar de «teoría que surge de los datos».

En definitiva, la relación establecida entre *teoría e investigación* (primer punto de la tabla 2.1) es muy distinta en cada uno de los dos enfoques. En la investigación cuantitativa inspirada por el paradigma neopositivista, la relación está estructurada en fases que siguen una secuencia lógica, un planteamiento deductivo, es decir, la teoría precede a la observación, orientada a la comprobación empírica de la teoría formulada previamente. En este contexto adquiere una gran importancia el análisis sistemático de la *literatura* existente, de hecho, ésta es la que sugiere las hipótesis.

En la investigación cualitativa inspirada en el paradigma interpretativo, la relación entre teoría e investigación es abierta, interactiva. El investigador cualitativo suele rechazar deliberadamente la formulación de teorías antes de empezar a trabajar sobre el terreno, por considerar que podría inhibir su capacidad de «comprender» el punto de vista del sujeto estudiado, que podría cerrarle horizontes a priori. La elaboración de la teoría y la investigación empírica se producen, por tanto, de manera simultánea. Como es lógico, el enfoque cualitativo concede menos importancia a la reflexión sobre la literatura existente.

De ello se deriva también un uso distinto de los *conceptos*. Los conceptos son elementos constitutivos de la teoría, y al mismo tiempo permiten su comprobación empírica mediante la «operacionalización» de los mismos, es decir, su transformación en variables observables. En el enfoque positivista, la clarificación de los conceptos y su operacionalización en forma de variables se realizan incluso antes de que se inicie la investigación. Tomemos el concepto de «disgregación familiar» en la investigación de Sampson y Laub. Los autores lo operacionalizan del siguiente modo: el valor 1 (disgregación) se da cuando en una familia uno o los dos padres están ausentes por divorcio, separación, abandono, muerte; el valor 0 se da en el resto de los casos. Obviamente, esta definición operativa del concepto de disgregación fami-

Tabla 2.1. Comparación entre investigación cuantitativa y cualitativa

	Investigación cuantitativa	**Investigación cualitativa**
Planteamiento de la investigación		
Relación teoría-investigación	Estructurada, las fases siguen una secuencia lógica Método deductivo (la teoría precede a la observación)	Abierta, interactiva Método inductivo (la teoría surge de la observación)
Función de la literatura	Fundamental para la definición de la teoría y e hipótesis	Auxiliar
Conceptos	Operativos	Orientativos, abiertos, en construcción
Relación con el entorno	Manipulador	Naturalista
Interacción psicológica entre el estudioso y el objeto estudiado	Observación científica, distante, neutral	Identificación empática con el objeto estudiado
Interacción física estudioso-estudiado	Distancia, separación	Proximidad, contacto
Papel del sujeto estudiado	Pasivo	Activo

(sigue en la página 45)

Tabla 2.1. Comparación entre investigación cuantitativa y cualitativa *(continuación)*

	Investigación cuantitativa	Investigación cualitativa
Recopilación de datos		
Diseño de la investigación	Estructurado, cerrado, anterior a la investigación	Desestructurado, abierto, se construye en el curso de la investigación
Representatividad/inferencia	Muestra estadísticamente representativa	Casos individuales no representativos estadísticamente
Instrumento de investigación	Uniforme para todos los sujetos. Objetivo: matriz de datos	Varía según el interés de los sujetos. No se tiende a la estandarización
Naturaleza de los datos	*Hard,* objetivos y estandarizados	*Soft,* subjetivos y flexibles
Análisis de los datos		
Objeto del análisis	La variable (análisis por variables, impersonal)	El individuo (análisis por sujetos)
Objetivo del análisis	Explicar la variación de las variables	Comprender a los sujetos
Uso de técnicas matemáticas y estadísticas	Máximo	Ninguno
Resultados		
Presentación de los datos	Tablas (enfoque relacional)	Fragmentos de entrevistas, textos (enfoque narrativo)
Generalizaciones	Correlaciones. Modelos causales. Leyes. Lógica de la causalidad	Clasificaciones y tipologías. Tipos ideales. Lógica de la clasificación
Alcance de los resultados	Se buscan generalizaciones (inferencia)	Especificidad

liar permite la observación empírica del mismo, pero también tiene el inconveniente de reducirlo y empobrecerlo, con el riesgo de cosificar el indicador preseleccionado. En efecto, en el curso del análisis es probable que se pierdan de vista las limitaciones de dicha definición operativa [2].

Un investigador cualitativo nunca habría actuado de esta manera. No habría traducido el concepto en forma de variable desde el principio de la investigación, sino que habría utilizado el concepto de «disgregación familiar» como un *sensitizing concept,* tal y como lo define Blumer: un concepto «orientador», que predispone a la percepción, y que hay que agotar en el curso de la propia investigación, no sólo en términos operativos, sino también en términos teóricos:

> Mientras que los conceptos definitivos indican lo que se debe observar, los conceptos orientadores sirven como guía para acercarse a la realidad empírica [...], en una relación con esta realidad que permite comprobar, completar y enriquecer las propuestas sobre la misma gracias a los datos empíricos, [en un proceso que] va del concepto a la variedad concreta de la realidad, en lugar de tratar de enjaular a esta realidad en una definición abstracta del concepto.
>
> [Blumer 1969, 147-148]

Blumer añade que nuestros conceptos deben ser orientativos y no definitivos, no por inmadurez de la ciencia social o por falta de sofisticación científica, sino por la propia naturaleza del mundo empírico, «en el que cada objeto considerado por nosotros, ya sea una persona, un grupo, una institución, una conducta o cualquier otro, tiene un distintivo propio, una particularidad o característica única, y se sitúa en un contexto también específico» [*ibídem,* 148].

Tras esta serie de diferencias entre investigación cuantitativa y cualitativa, que conciernen a la relación entre teoría e investigación, hay un segundo grupo de diferencias relativas a la relación personal del investigador con la realidad estudiada. Primero consideraremos la *relación* general *con el entorno estudiado.* Uno de los problemas que debe afrontar el investigador social es el de la «reactividad» del objeto de estudio. El mero hecho de estudiar la conducta humana ya podría provocar cambios en la misma. Estudiar a los individuos no es como estudiar a las hormigas o los átomos: si alguien sabe que lo están observando, es probable que se comporte de forma poco natural.

Al enfoque neopositivista no parece preocuparle demasiado este problema. No es que el investigador social que se inspira en este paradigma no sea consciente de esta diferencia fundamental entre las ciencias naturales y las ciencias sociales. El investigador neopositivista no considera que la reactividad del objeto pueda representar un obstáculo serio, o al menos considera admisible un cierto margen de «manipulación controlada». La investigación cualitativa, en cambio, considera

---

[2] Un ejemplo muy debatido de este tipo de cosificación es el cociente intelectual; el instrumento utilizado para su medición, el CI, se ha convertido sin más en sinónimo de inteligencia, lo cual tiene graves consecuencias dado el marcado carácter reduccionista del instrumento.

que el *enfoque naturalista* es un requisito fundamental para la investigación empírica. Podemos describir el enfoque naturalista como una intervención sobre la realidad en la que el investigador se abstiene de toda manipulación, estimulación, interferencia o perturbación de esa realidad, que estudia en su curso natural.

Las dos formas de hacer investigación encuentran dos ilustraciones típicas y opuestas en las técnicas del experimento y la observación participante. En el experimento el investigador manipula la realidad social hasta el punto de construir una situación artificial, por ejemplo, poniendo a unos estudiantes en un laboratorio y sometiéndolos a un reportaje de propaganda política. Antes y después de esta exposición los somete a un test; además, al principio dividió a los estudiantes en dos grupos, siguiendo criterios artificiales, y aplicó el estímulo a uno de los grupos y al otro no. Se trata de una situación artificial en la que la intervención manipuladora del investigador es invasora. En el caso de la observación participante, el investigador se limita a observar lo que sucede en la realidad social, y a veces incluso evita esa intervención mínima consistente en realizar preguntas-encuestas a los sujetos observados.

Son técnicas de investigación muy diferentes. Entre ellas existen diversos matices. La propia observación participante rara vez es totalmente «naturalista», en el sentido de que la presencia de un observador extraño tiene que afectar de alguna manera a los sujetos, salvo en casos concretos como el de la observación de los niños. En general, todos los instrumentos del análisis cualitativo distintos de la pura observación, por ejemplo, las entrevistas detalladas, los relatos de experiencias vividas, etc., implican siempre una intervención sobre la realidad, al menos como estímulo para hacer o comunicarse a los sujetos. De manera simétrica, la investigación cuantitativa no siempre es tan manipuladora como en el caso del experimento. Una entrevista de preguntas abiertas, por ejemplo, respeta más el contexto natural del individuo que responde, que un cuestionario de preguntas cerradas. Existen, por otra parte, técnicas cuantitativas que, al no intervenir sobre los sujetos y actuar, por ejemplo, sobre fuentes estadísticas o bien sobre variables «no reactivas», no presentan este problema.

La segunda parte de la relación entre el estudioso y la realidad estudiada concierne a la relación del investigador con cada individuo (hasta ahora hemos hablado sobre todo de la relación del investigador con el contexto estudiado, considerado en su totalidad). Como hemos señalado antes, una diferencia fundamental entre los paradigmas neopositivista e interpretativo reside en la definición del objeto de investigación, en el primer caso la «comprobación empírica de las hipótesis», y en el segundo, el «descubrimiento del punto de vista del actor social». De esta doble perspectiva se derivan dos hechos, uno de carácter psicológico-cultural y otro de carácter físico-espacial.

El primero concierne a la interacción psicológica entre el estudioso y lo estudiado. El investigador cuantitativo se sitúa en un punto de observación externo al sujeto estudiado, propia del observador «científico» neutral y distante; por otro lado, estudia lo que, como investigador, le parece importante, o así lo considera la comunidad científica. El investigador cualitativo, en cambio, se sitúa lo más cerca posible del sujeto de análisis, con la intención de ver la realidad social «con los ojos

de los sujetos estudiados». Para conseguirlo, nunca se mantiene neutral o indiferente, sino que tiende a desarrollar una identificación empática con los sujetos. Jankowski escribe en el prólogo de su libro:

> ... los diez años y cinco meses que he dedicado a esta investigación han sido un viaje[3]. Un viaje no sólo en el tiempo, sino también en las vidas de los miembros de las bandas y de muchas otras personas que viven en las zonas pobres de Nueva York, Boston y Los Ángeles. Irónicamente, también ha sido un viaje de vuelta a mi juventud [...], en el cual he encontrado personas maravillosas que recordaré siempre con afecto, y otras no tan maravillosas, que tampoco olvidaré jamás.
>
> [Jankowski 1991, XI]

Esta inmersión psicológica, que no deja indiferente o neutral al investigador, pone de manifiesto el problema de la objetividad en la investigación cualitativa. Problema real que también existe en la investigación cuantitativa, ya que, sin duda, lo que el investigador percibe está influido por su punto de vista, sus experiencias, su cultura, sus valores. Es imposible cumplir el ideal absoluto de ciencia objetiva y libre de valores, al menos en el campo de las ciencias sociales. Pero también es cierto que el problema es mucho más evidente para el investigador cualitativo, que en su interacción empática con el sujeto estudiado se expone al riesgo de implicarse emocionalmente y, por tanto, de caer en interpretaciones unilaterales.

El segundo hecho, derivado del anterior, se refiere a la interacción física entre estudioso y estudiado, caracterizada por la distancia o la proximidad. La investigación cuantitativa no suele prever contacto físico alguno entre el estudioso y lo estudiado. Pensemos en una investigación realizada con un cuestionario sobre una muestra de la población, en la que las entrevistas se encarga a un instituto demoscópico. O en un experimento de laboratorio en el que el investigador se limita a observar la conducta de los sujetos. O en un análisis secundario como el de Sampson y Laub, en el que los investigadores no tienen la posibilidad de conocer físicamente a ninguno de los sujetos estudiados, puesto que la investigación se realizó más de treinta años antes.

En el caso de la investigación cualitativa, el encuentro y la interacción entre estudioso y estudiado es una condición previa para el estudio. Al describir su interacción con los sujetos estudiados, Jankowski escribe: «[a lo largo de los 10 años de estudio] he participado en casi todo lo que ellos hacían. He comido donde ellos comían, dormido donde ellos dormían, he estado con sus familias, he ido donde ellos iban, y en ciertas situaciones en las que no podía permanecer neutral, he luchado con ellos» [*ibídem,* 13]. El caso de la observación participante sirve una vez más como ejemplo del punto de vista de la investigación interpretativa. Pero la exigencia de proximidad física y espacial con el objeto estudiado está presente en todas las técnicas cualitativas (pensemos, por ejemplo, en las entrevistas en profundidad, en los relatos de experiencias vividas, en el análisis de las dinámicas de grupo, etc.).

Por todo lo dicho, resulta evidente la diferencia radical entre ambos enfoques respecto al *papel del objeto estudiado.* En el enfoque cuantitativo, el individuo estudiado se considera *pasivo*, y si esto no fuera posible, se intenta reducir al máxi-

mo su interacción con el investigador. La concepción de la investigación como «observación» y «estudio» implica que los individuos estudiados se consideran objetos (recordemos la idea positivista original de que los hechos sociales son «cosas»). En el enfoque cualitativo, en cambio, la concepción de la investigación como «interacción» implica un papel *activo* del sujeto estudiado. Su coparticipación directa y creativa en el proceso de investigación no se evita, sino que se busca, como se deduce claramente de los dos pasajes anteriores de Jankowski.

2.3.2. Recopilación de los datos

Una de las diferencias principales entre ambos enfoques está relacionada con el diseño de la investigación, es decir, dónde, cómo y cuándo se recopilan los datos, los instrumentos usados para la investigación, entrevistas u observación participante, cuestionarios o experimentos, etc., la localización, cuáles y cuántos sujetos se entrevistarán, cuáles y cuántas organizaciones se estudiarán, etc. La diferencia entre los dos enfoques reside en el grado de estructuración de los procedimientos. En el caso cuantitativo, el diseño de la investigación se construye antes de empezar a recopilar datos, y está estructurado y cerrado. En cambio, en la investigación cualitativa, el diseño no tiene una estructura fija, es abierto, de modo que permita captar lo imprevisto, y puede cambiar en el curso del proceso.

En la investigación de Sampson y Laub, una vez construida la muestra de 500 sujetos que cometieron delitos y 500 que no lo hicieron, ésta se mantuvo invariable. En la investigación de Jankowski, en cambio, una vez definidos algunos criterios para observar, como el número de bandas de Nueva York, Boston y Los Ángeles, los vínculos étnicos y la dimensión de las bandas, el investigador podía elegir las bandas más adecuadas para su objetivo. Tenía además libertad para entrevistar a quien quisiera, prolongar o acortar la observación según conviniera, etc.

Bajo este punto de vista, los dos ejemplos elegidos no son los más típicos. La investigación del matrimonio Glueck, de la que Sampson y Laub extrajeron sus datos, preveía diversas fuentes de información, por ejemplo, sobre las actividades delictivas de los sujetos, entre las que el investigador podía escoger libremente. Por su parte, Jankowski estableció unos objetivos de representación y comparación, por lo que podríamos decir que siguió un plan preestablecido para escoger las bandas. La investigación cuantitativa suele tener un diseño rígido, como en el caso de un estudio con un cuestionario de preguntas cerradas sobre una muestra probabilística, o un experimento, mientras que la investigación cualitativa es más flexible, es decir, el investigador decide sobre la marcha qué sujetos estudiar y qué instrumentos de investigación utilizar.

De este enfoque cerrado o abierto, establecido a priori o durante la investigación, se derivan otras dos características que distinguen a ambos enfoques. La primera se refiere a la *representatividad* de los sujetos estudiados. En los estudios cuantitativos el investigador está preocupado por la posibilidad de generalizar los

[3] El autor había señalado en una ocasión anterior que en el *Webster New American Dictionary*, entre los significados de la palabra «gang» (banda) se encuentra también el de 'viaje'.

resultados, y el uso de una muestra estadísticamente representativa es el indicador más claro de esta preocupación. Podríamos decir que el investigador cuantitativo está preocupado por la representatividad de la parte de la sociedad que estudia, más que por su capacidad de comprenderla.

El investigador cualitativo, en cambio, antepone la comprensión de los sujetos estudiados, aun a riesgo de perderse en el seguimiento de situaciones atípicas y mecanismos no generalizables. Al investigador cualitativo no le interesa en absoluto la representatividad estadística. Si acaso le interesará una especie de representatividad sustantiva, sociológica, que se decide no con fórmulas matemáticas, sino según el juicio del propio investigador. Éste elegirá los casos en que se debe profundizar, no por su difusión en la población, sino por el interés que según él presentan. Interés que, por otra parte, puede modificarse en el curso de la investigación: por ello, como hemos encontrado en un manual de investigación cualitativa, «el muestreo se realiza en función de la evolución de la relevancia teórica de los conceptos» [Strauss y Corbin, 1990, 179].

Tomemos como ejemplo la investigación sobre «comunistas y católicos» que el antropólogo estadounidense David Kertzer [1980] realizó en un barrio de Bolonia a finales de los años setenta. Siguiendo el planteamiento clásico de los «estudios de comunidad», Kertzer empleó la técnica de la observación participante, junto con una serie de entrevistas con «observadores privilegiados». Se utilizaron entrevistas en profundidad «no estructuradas» y entrevistas estructuradas, con las mismas preguntas para todos los sujetos. Sin embargo, los sujetos a entrevistar no fueron preseleccionados al inicio de la investigación.

En un principio el autor pensaba entrevistar a todos los agentes políticos y sociales del barrio, tanto de tendencia comunista como católica, miembros de la directiva de las cuatro secciones del Partido Comunista, activistas de la UDI (Unión de mujeres italianas), de la FGCI (Federación juvenil del PCI), sacerdotes activos de las parroquias, responsables laicos de las asociaciones católicas, etc. Sin embargo, en el curso de la investigación, Kertzer percibió que su objetivo era inalcanzable, por ejemplo, algunos activistas del PCI y sacerdotes se mostraron desconfiados ante las posibles preguntas de un estadounidense; en cambio, ciertos individuos que en un principio no se incluyeron entre los sujetos a entrevistar resultaron ser informadores muy valiosos, por ejemplo una camarera del barrio, que resultó ser una de las fuentes de información más ricas sobre las relaciones entre comunistas y católicos.

El segundo elemento diferenciador entre la investigación cuantitativa y la cualitativa es la uniformidad del *instrumento de investigación*. En investigación cuantitativa todos los sujetos reciben el mismo tratamiento. El instrumento de investigación utilizado es el mismo en todos los casos (por ejemplo, un cuestionario), o al menos tiene capacidad para uniformar (por ejemplo, una ficha codificada para uniformar unas preguntas abiertas o la información procedente de distintas fuentes, como sucedió en la investigación de Sampson y Laub). Esta uniformidad es necesaria porque el objetivo final de la recopilación de información es la «matriz de datos», una matriz de números ordenados en filas y columnas, donde se codifica la misma información para todos los casos.

La investigación cualitativa no persigue esta estandarización. Al contrario, la falta de homogeneidad de las informaciones es su principal elemento constitutivo, dado que el investigador asume informaciones distintas según el caso, con distinto nivel de profundización según la conveniencia. Este planteamiento distinto se atribuye a la diferencia del objetivo cognoscitivo, que en un caso es descubrir las uniformidades del mundo humano y en el otro es comprender las manifestaciones en su individualidad.

El último punto relativo a la «recopilación de datos» guarda relación con la *naturaleza de los datos*. En la investigación cuantitativa éstos son fiables, precisos, rigurosos e unívocos, o al menos, se espera que lo sean. El término inglés utilizado para designar estas características es *hard*, en español se expresaría con los términos «objetivo» y «estandarizado». Objetivo quiere decir que los datos no deben estar influidos por la subjetividad interpretativa del investigador, es decir, si cambiara el investigador, el resultado de la observación debería ser el mismo, ni por la subjetividad expositiva del sujeto estudiado, es decir, la misma situación aplicada a sujetos diferentes debería producir el mismo dato resultante de la observación. Estandarizado quiere decir que los datos recopilados sobre los distintos sujetos, y tal vez por investigadores distintos, deben ser comparables entre sí, para poder compararlos, se suelen expresar de forma numérica. Esto se puede conseguir fácilmente con algunas variables estructurales básicas, como el género, la edad, el nivel de estudios, o de conducta, como la práctica religiosa, el comportamiento electoral, pero resulta más difícil con conceptos más complejos, como la clase social o la inteligencia, y aun más con actitudes, como autoritarismo pedagógico, conservadurismo político, etc. La investigación cuantitativa intenta siempre que los datos tengan estas dos características: por ejemplo, en el caso de la actitud hacia un representante político, pide a los entrevistados que le asignen una calificación del 1 al 10 (como con las notas escolares), o una puntuación de 0 a 100 (midiendo el grado de simpatía con una especie de «termómetro de los sentimientos»).

La investigación cualitativa no se plantea el problema de la objetividad y la estandarización de los datos, y se preocupa en cambio de la riqueza y profundidad de los mismos. Los datos que produce la investigación cualitativa se describen con el término inglés *soft*, que en español se podría traducir como subjetivos, flexibles, relativos. Volviendo al ejemplo anterior, la popularidad de un político se valorará en función de las perspectivas, los puntos de vista y la cultura de los entrevistados, respetando las formulaciones de cada uno y su nivel de profundidad.

2.3.3. Análisis de los datos

El análisis de los datos es quizás la fase de una investigación social en la que resulta más visible la diferencia entre el enfoque cuantitativo y el cualitativo. El efecto de la instrumentación matemática y estadística utilizada en el enfoque cuantitativo, con sus tablas, gráficos, test estadísticos, etc., así como el complejo instrumental de ordenadores, bases de datos, programas de análisis estadístico, etc., contrastan con la sobriedad de un análisis cualitativo, que no utiliza herramientas estadísticas o matemáticas, y se limita a organizar el material empírico con la eventual ayuda de instrumentos infor-

máticos [4].

Pero más que en la instrumentación tecnológica utilizada para el análisis de los datos y en la presentación final de los resultados, la diferencia más importante se encuentra en la lógica que guía el propio análisis. Empecemos por el objeto del análisis y volvamos a la investigación de Sampson y Laub sobre la desviación juvenil. A partir de las correlaciones entre variables dependientes e independientes, los autores llegan a la conclusión de que «...los procesos de control social informal presentes en la familia y el colegio proporcionan la clave interpretativa para encontrar la explicación causal de la delincuencia en la infancia y la adolescencia» [Sampson y Laub, 1993, 246].

¿Cuál fue el proceso lógico que permitió llegar a estas conclusiones? Para realizar la investigación el equipo recogió evidencia empírica *para cada individuo,* como se hace en todas las investigaciones, cualitativas o cuantitativas. Es decir, de las 500 + 500 personas estudiadas se registraron todas las propiedades individuales que llamamos «variables» en la fase de análisis de los datos (actos de violencia cometidos, composición de la familia, profesión de los padres, entorno familiar, rendimiento escolar, etc.). Cada sujeto se describe de manera analítica a partir de estas propiedades. Podríamos decir que su unidad como individuo se fragmenta en tantos elementos como variables se utilicen para describirlo. A partir de este momento, el investigador ya no tratará al sujeto como persona. *El análisis de los datos se hará siempre a través de las variables,* de manera *impersonal.* Se hablará de medias de las *variables* (media de delitos cometidos, media de hijos por familia, renta media, etc.), de porcentajes de *variables* (porcentaje de sujetos que tienen antecedentes penales en su familia, un padre violento y autoritario, etc.), y de relaciones entre *variables* (correlación entre agresividad adolescente y violencia juvenil, etc.).

El *objetivo del análisis* será «explicar la variación» en los valores de las variables dependientes, es decir, encontrar las causas que provocan la variación de las variables dependientes en los sujetos, los factores que «explican» por qué algunos jóvenes se han convertido en delincuentes y otros no. Se estudiaría, por ejemplo, si las diferencias entre jóvenes delincuentes y no delincuentes se explican por el tipo de educación que han recibido de sus padres. Si *todos* los delincuentes proceden de familias con padre violento, y de los jóvenes cuyo padre no es violento *ninguno* es delincuente, habremos «explicado estadísticamente» la variación de la variable «delincuencia». Como veremos más adelante, este hecho también se puede interpretar diciendo que hemos encontrado la «causa» de la variable delincuencia, en este caso, la conducta del padre.

Este modo de proceder es el mismo que el se utiliza en ciencias naturales. La relación causal entre el consumo de tabaco y el cáncer de pulmón, por ejemplo, se ha deducido de la observación de una relación estadística entre la variación de dos variables en miles de sujetos, aislando la evolución de estas dos variables en la serie

[4] Es cierto que recientemente está aumentando el uso de programas informáticos para el análisis de datos cualitativos, por ejemplo, el programa Atlas-Ti o el programa Nudist.

de variables que varían con ellas.

El enfoque interpretativo critica este modo de proceder, acusándolo de asumir impropiamente el modelo científico de las ciencias naturales. Esta crítica ha surgido no sólo en las ciencias sociales, sino también en el campo de la medicina, donde se rechaza la conexión causa-efecto entre distintas variables en nombre de la unidad psíquica y somática del cuerpo humano. El argumento de dicha acusación es la imposibilidad de reducir el «hombre» a una serie de variables, y la necesidad de una perspectiva global u *holística* para el análisis de la conducta humana [5]. Este planteamiento subraya que el individuo es algo más que la suma de sus partes y no puede quedar reducido a la relación entre unas cuantas variables, y que la comparación de los sujetos mediante variables conlleva la desnaturalización de los mismos.

La investigación cualitativa adopta un planteamiento totalmente diferente para el análisis de los datos. *El objeto del análisis ya no está representado por la variable, sino por el individuo en su totalidad.* Si la investigación cuantitativa se basa en *variables,* la cualitativa se basa en *casos.* Para explicar esta diferencia volvemos de nuevo a la investigación de Jankowski. En el curso de la misma, el autor se pregunta cuáles son las causas de las conductas violencias en los individuos que pertenecen a las bandas. Jankowski parte de cuatro factores desencadenantes de la violencia: el miedo, la ambición, la frustración y la exhibición. Llegado a este punto, un investigador cuantitativo trataría de obtener la variable dependiente «violencia», por ejemplo, atribuyendo a los individuos una puntuación según el grado de violencia manifestado en su conducta, y del mismo modo trataría de obtener con indicadores adecuados las cuatro variables independientes mencionadas. Su objetivo sería por tanto «explicar la variación» de la variable dependiente «violencia» a partir de las variaciones de las variables independientes, es decir, encontrar la correlación entre las variables independientes y la variable dependiente, utilizando los instrumentos estadísticos adecuados.

Jankowski no fragmenta los sujetos en variables, sino que los clasifica por tipos. La clasificación es lo que conecta a los sujetos, igual que el modelo causal conecta las variables entre sí.

Jankowski distingue los episodios de violencia según sean expresiones de violencia individual u organizada (por la banda). Distingue también los contextos en los que tiene lugar esa violencia, identificando seis: violencia contra miembros de la misma banda, de otras bandas, personas de la comunidad, personas de fuera de la comunidad, bienes materiales de la comunidad y de fuera de la misma. La clasificación consta, por tanto, de doce situaciones, y dentro de cada una el autor distingue cuatro casos, en función de los cuatro factores emotivos que desencadenan la violencia: miedo, ambición, frustración y exhibición. De este modo, obtiene una tipología de 48 tipos para clasificar los episodios de violencia, en términos de los sujetos y sus conductas. Por ejemplo, para el tipo «violencia individual,

[5] Por «perspectiva holística» (del griego *hólos* = todo, entero) se entiende también un planteamiento en el que los sujetos sociales se estudian íntegramente como sistemas complejos, asumiendo que un sistema no se puede dividir en partes independientes.

contra miembros de la misma banda, causada por la ambición», el autor describe el caso de Shoes. Shoes es un chico de 16 años, miembro de una banda de Nueva York, que quiere convertirse en uno de los jefes, a pesar de ser demasiado joven para el cargo y llevar poco tiempo en la banda. Un día, durante una pelea con otra banda, Shoes percibe que dos de sus compañeros no se implican en la contienda, sino que se mueven con cautela en la retaguardia por temor a que los ataquen. En la siguiente reunión del grupo, Shoes ataca con violencia a uno de los dos, y cuando los otros los separan justifica su comportamiento acusando a los dos compañeros de cobardía y pidiendo su expulsión de la banda. Todo ello para conseguir buena reputación en el grupo.

Como puede verse, el *objetivo del análisis,* coherente con el paradigma interpretativo, es «comprender a las personas», interpretar el punto de vista del actor social, (en el ejemplo, comprender las motivaciones de la conducta violenta), igual que en el enfoque cuantitativo el objetivo era «explicar la variación de las variables».

La relación que tienen los enfoques cuantitativo y cualitativo con las *técnicas matemáticas y estadísticas* es muy diferente. En el primer caso, el lenguaje de las matemáticas se considera *el* lenguaje de la ciencia, por lo que se hace un gran esfuerzo por transformar los conceptos en términos matemáticamente tratables («medibles») y se utilizan al máximo las técnicas estadísticas para analizar los datos y hacer generalizaciones a partir de la muestra de población. En la investigación cualitativa, la formalización matemática se considera no sólo inútil, sino hasta nociva (restrictiva), por lo que se ignora completamente.

2.3.4. Resultados

De la diferencia de planteamiento, recopilación y análisis de los datos también se deriva una profunda discrepancia en los resultados que se logran con las dos formas de hacer investigación. Comencemos con el aspecto más llamativo, la *presentación de los datos.* Las dos formas clásicas, y también las más sencillas, de presentación de los datos en las tradiciones cuantitativa y cualitativa son, respectivamente, la «tabla» y la «narración».

Pongamos como ejemplo una investigación realizada en Italia a mediados de los años sesenta sobre los militantes de base de la Democracia Cristiana y del Partido Comunista [Alberoni *et al,* 1967]. Se estudió a 108 activistas (54 democristianos y 54 comunistas) mediante entrevistas dirigidas, con un diseño común y una duración media de 6 a 7 horas, subdivididas en varias sesiones (entre tres y seis). Las entrevistas se grabaron, se transcribieron y se codificaron, clasificando las respuestas en categorías y atribuyendo a cada categoría un valor numérico (hablaremos con más detalle de este procedimiento en el capítulo 10). Una de las «variables» obtenida de este modo se definió como «visión dicotómica de la realidad», es decir, la tendencia (o no) a interpretar las fuerzas en la sociedad en términos de «frentes opuestos [...] dos partes, una de las cuales es la parte positiva con la que nos identificamos, y la otra es el enemigo al que enfrentarse y vencer y destruir» [*ibídem,* 381]. Esta tendencia a tener una visión dicotómica de la realidad varía según los sujetos, que se

clasificaron en las cuatro categorías de la tabla 2.2 , que abarcan desde «totalmente dicotómica» a «en absoluto dicotómica» Analicemos el contenido de esta tabla. En primer lugar, la columna de los entrevistados (DC + PCI) nos sugiere que los militantes estudiados están repartidos de forma equilibrada entre «dicotómicos» y «no dicotómicos» (casi el 50% en cada grupo, si sumamos las dos primeras y las dos últimas categorías de la variable). Ahora bien, si consideramos a los militantes de DC y del PCI por separado, podemos sacar las siguientes conclusiones: entre los primeros los «dicotómicos» están en minoría (menos de un tercio), en cambio, entre los segundos predomina la visión dicotómica (más de dos tercios). La información de carácter cuantitativo que ofrece la tabla es concisa, escueta y compacta: con pocos números, la tabla ilustra una característica importante de los activistas, la relación entre afiliación política y orientación ideológica. Por eso, en la tabla 2.1 hablábamos de *enfoque relacional.*

Sin embargo, estos datos presentan dos limitaciones. En primer lugar, la «visión dicotómica de la realidad» es una categoría conceptual de los investigadores, que se sitúa entre quien lee la tabla y las categorías mentales efectivas de los entrevistados. Por otra parte, tras este concepto básico hay una riqueza que no se puede percibir si nos limitamos a leer la tabla. Los extractos de entrevistas están menos afectados por estas dos limitaciones. En primer lugar, las palabras del entrevistado permiten «ver la realidad con los ojos de los sujetos estudiados»[6] de un modo más efectivo. En segundo lugar, la cita confiere al dato una dimensión inexistente en la tabla, y

Tabla 2.2. Visión de la realidad social por parte de los activistas de partido (en porcentajes sobre el total)

	DC	PCI	DC + PCI
Totalmente dicotómica	5,6	31,5	18,5
Predominantemente dicotómica	25,9	38,9	32,4
Concibe la realidad como lucha entre posiciones, pero es más sensible a las alternativas	51,8	25,9	38,9
En absoluto dicotómica	16,7	3,7	10,2
Total	100	100	100
(N)	(54)	(54)	(108)

FUENTE: Alberoni *et al.* [1967].

[6] No queremos decir que con el mero uso de la cita estamos transmitiendo al lector la idea que tienen los individuos estudiados de la realidad. Lo que se transmite siempre es la interpretación del investigador: él decide a quién citar entre los sujetos estudiados, la importancia que concede a un fragmento concreto de conversación, o las conexiones entre las distintas citas presentadas. Al margen de estas decisiones, la respuesta, o alguna otra forma de comunicación, del entrevistado se cita en su forma original, mientras que en la tabla hay un elemento de mediación posterior, que es la codificación de la respuesta dentro de las categorías establecidas por el investigador.

permite al lector «visualizar» a los entrevistados, como una fotografía permite un conocimiento distinto y más completo que la mera descripción de rasgos externos. Es lo que en la tabla 2.1 hemos denominado *perspectiva narrativa.*

Reproducimos un fragmento de entrevista referente a la «visión dicotómica de la realidad», que muestra que se trata de un concepto muy sintético compuesto por varias dimensiones específicas. Por ejemplo, implica un concepto claro de la política internacional, que en el caso de un activista comunista se expresa en la idealización dogmática de la Unión Soviética:

> ... desde que era joven siempre me gustó la Unión Soviética: ahora la conozco aun mejor. Estoy convencido de que la política de Rusia es la correcta. Estoy convencido de que Rusia hace todo lo posible por evitar la guerra. Rusia no quiere la guerra; su guerra es la propaganda que hace en el mundo y que logra simpatías en los otros estados y atrae al mercado comercial desde todos los estados del mundo. Este sistema ha salido ganando con ello, y hoy o mañana el socialismo tendrá que estar en todo el mundo y a esos estados los pondrá de rodillas sin guerras porque provocará su desaparición [...]. En Rusia no tienen ninguna necesidad de hacer huelga; allí trabajan para el pueblo [...], trabajan para sí mismos [...]. Allí está el socialismo, y con el socialismo no hay necesidad de ir a la huelga [...]. Allí están muy por delante [...]. No sé si el año pasado o hace algunos años, el pan allí era gratis para todo el pueblo [...]. Cuando un pueblo llega a esto, significa que no necesitan hacer huelga.
>
> [Alberoni *et al.*, 1967, 479]

Si bien tabla y narración son dos formas típicas de presentar los resultados en la investigación cuantitativa y cualitativa respectivamente, no son instrumentos unilaterales utilizables sólo por uno o el otro tipo de investigación. Son muy comunes las investigaciones de planteamiento neopositivista en las que la narración se emplea con fines ilustrativos, con la misma función que cumple una fotografía en un artículo de periódico. En este caso el análisis de los datos se hace con instrumentos cuantitativos, sobre variables numéricas, mediante tablas y análisis multivariados. La narración sirve para ejemplificar los resultados, para que el lector pueda comprender mejor el universo real que subyace al dato numérico. Por ejemplo, en la investigación citada antes, tras distinguir las variables que caracterizan mejor a los distintos tipos de activista de partido, en nuestro caso, la edad, la educación, el grado de compromiso con el activismo, la actitud dicotómica, la búsqueda de beneficios personales, los investigadores reconocieron en algunos entrevistados este conjunto de características específico y citaron extractos tomados de sus entrevistas.

Mucho más raro es el caso contrario. El investigador que sigue el enfoque interpretativo raras veces utiliza la tabla para mostrar sus variables relevantes [7], y es reacio a aplicar categorizaciones propias a las respuestas y actitudes de los entrevistados. Su objetivo es presentar la visión de la realidad de los sujetos estudiados, más

[7] Naturalmente podrá utilizarlas sin problemas para variables secundarias y descriptivas de la muestra, como las variables sociodemográficas básicas.

que localizar unos rasgos generalizables.

A continuación hablaremos de las *generalizaciones*. La tabla y la narración son dos formas de presentación básicas y, en cierto sentido, fragmentadas de los datos. La conclusión de una investigación debería ir más allá de la simple exposición de distribuciones de variables o ilustraciones de casos, y establecer relaciones entre las variables o conexiones entre los casos, ya que su objetivo no debería ser sólo describir la realidad, sino también ofrecer sistematizaciones y síntesis de orden superior, ya sean explicaciones o interpretaciones. Éste es, en efecto, el único modo de conectar la investigación con la teoría, que representa una forma de abstracción racional y sintética de la realidad.

En investigación cuantitativa el camino para llegar a estas síntesis está claro, y consiste en el proceso que lleva al planteamiento de relaciones causales entre las variables, mediante el estudio de las relaciones entre ellas. Tras fragmentar al individuo en variables, el análisis cuantitativo aporta una primera síntesis expresada en la correlación entre variables, con indicadores numéricos como el coeficiente de correlación, una red de conexiones causales entre variables, y, en los casos más afortunados, expresiones sintéticas de forma similar a las «leyes» de las ciencias naturales.

Considérese, por ejemplo, una conocida investigación realizada en Estados Unidos en los años sesenta, sobre la estratificación social [Blau y Duncan, 1967]. Los autores abordaron, entre otros, la tarea de valorar los efectos sobre la posición social de los individuos del estatus adscrito, en este caso, la posición social del padre, y el estatus adquirido, en este caso, el nivel de formación. El objetivo del estudio era saber si para conseguir el éxito social es más importante tener un padre rico o un buen currículum. Todo ello a sabiendas de la dificultad para distinguir la acción de los dos factores, ya que las dos principales variables causales están relacionadas entre sí. Es decir, la posición social del padre influye también en la formación del hijo.

En la figura 2.1 hemos representado los nexos causales, en forma de flechas, que se supone que existen entre las variables: influencia directa de la «escolarización del hijo» sobre la «posición social del hijo» ($Z \rightarrow Y$); influencia directa de la «posición social del padre» sobre «posición social del hijo» ($X \rightarrow Y$); y, por último,

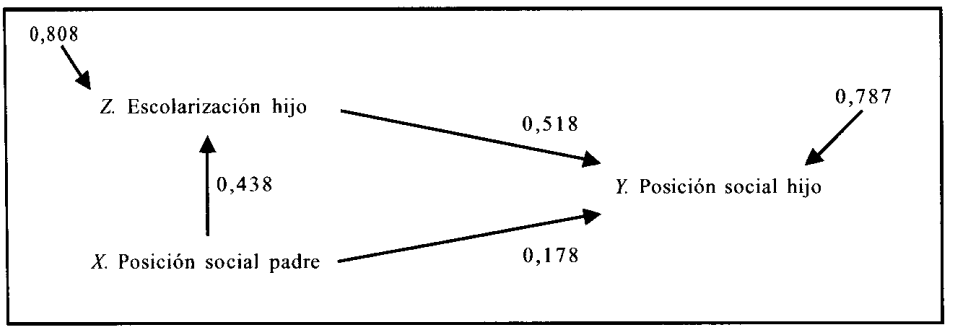


Figura 2.1. Modelo causal relativo a los vínculos entre posición social del padre, escolarización del hijo y posición social del hijo [Blau y Duncan, 1967].

influencia indirecta de la «posición social del padre» sobre la «posición social del hijo», pasando por la «escolarización» ($X \rightarrow Z \rightarrow Y$: la buena posición social del padre permite estudiar al hijo, y la escolarización de éste influye de manera positiva en su posición social).

Mediante la técnica estadística del *path analysis* es posible cuantificar el peso de estas componentes causales, expresado en el coeficiente que aparece por encima de las flechas del gráfico. En concreto se puede constatar, en vista de los valores referidos en el gráfico, que de las dos componentes directas, el efecto de la escolarización es mayor que el de la posición social del padre. Incluso considerando el efecto total de la posición social del padre, la suma de los efectos directo e indirecto de este factor no llega a igualar el efecto directo de la escolarización (ver tabla 2.3). De este modo la investigación puede cuantificar los efectos de las variables adscritas y adquiridas, y demostrar que las segundas prevalecen sobre las primeras.

Esta forma de proceder es bastante común entre los investigadores cuantitativos. En el enfoque cualitativo es más difícil encontrar modos de sintetizar la información que estén lo bastante extendidos entre los investigadores. Sin embargo, muchos autores [entre otros, Lofland, 1971; Spradley, 1980; Hammersley y Atkinson, 1983] consideran que la identificación de «tipos» es la vía para alcanzar estos objetivos de síntesis. Y ésta es también, por otra parte, la solución que la inspiración original weberiana sugiere al enfoque interpretativo con el concepto de «tipo ideal».

Como indicamos en el capítulo 1, el tipo ideal es una categoría conceptual que no tiene correlato exacto y específico en la realidad. Es una construcción que, aunque procede de la observación de casos reales, elimina los detalles y accidentes de la realidad para extraer sus características fundamentales, dándoles un nivel de abstracción superior, para utilizar el «modelo» obtenido como concepto-límite con el que interpretar esa realidad.

Como ejemplo de utilización del tipo ideal para interpretar la realidad podemos citar la investigación de Jenkins, *Lads, Citizens and Ordinary Kids,* realizada a finales de los años setenta, con entrevistas y observación directa de un grupo de adolescentes (53 chicos y 33 chicas) de la clase obrera de Belfast. Los «rebeldes», varones que muestran en sus conductas y valores de referencia ciertos rasgos machistas y

Tabla 2.3. Datos relativos a la figura 2.1

Relación original (r)	Efecto causal		Parte no causal	
	Directo	Indirecto	Total	(Espuria)
XY 0,405	0,178	0,226 *	0,405	—
XU 0,438	0,438	—	0,438	—
UY 0,596	0,518	—	0,518	0,078

* Según las reglas del *path analysis,* el coeficiente del efecto indirecto se obtiene por la multiplicación de los coeficientes de los efectos directos que lo componen: 0,438 × 0,518 = 0,226.

antiburgueses, asociados tradicionalmente con la clase obrera, los «ciudadanos», que encarnan los respetables valores burgueses de la austeridad, la diligencia, la independencia y el deseo de ascenso social, y los «niños corrientes», son tres categorías conceptuales sintéticas que el autor perfila tras su inmersión en la realidad estudiada, y que utiliza para releer, situar e interpretar analíticamente el material empírico disponible.

Con estos tres tipos ideales, el investigador interpreta los estilos de vida (consumo, forma de vestir, ocio, etc.), el modo de interactuar con el sexo opuesto, la relación con la iglesia y el deporte, la opinión sobre la familia y el matrimonio, el desarrollo escolar y las primeras experiencias laborales, etc. La categorización en tres tipos, sirve, por ejemplo, para interpretar las distintas ideas sobre el matrimonio y la familia. Para los rebeldes y los niños corrientes, el lugar de la mujer es la casa, y, como normal general, piensan que después del matrimonio la mujer debe dejar el trabajo y dedicarse a la familia, tener hijos, cuidarlos, etc. Los ciudadanos, en cambio, piensan que después del matrimonio la mujer debe seguir trabajando para ahorrar y comprar la casa, y posponer el nacimiento de los hijos. Los ciudadanos también opinan que se debe llegar al matrimonio siguiendo los pasos del cortejo, el noviazgo y la boda religiosa; aceptan las relaciones prematrimoniales, pero en el contexto de una relación sólida orientada al matrimonio. Esto no se aplica a los otros dos tipos. La tipología se utiliza de manera análoga para interpretar las diferencias en la muestra en relación con los otros temas estudiados.

Hay que subrayar que en todos estos casos no sólo se describe la realidad, sino que también se interpreta, se lee, se analiza y por último se *recompone y sintetiza* a partir de las categorías clasificatorias o de los tipos ideales reconocidos. Jenkins reencuadra los casos, relee sus características, reinterpreta los datos a la luz de los tres tipos propuestos. Del mismo modo, Jankowski utilizaba los cuatro mecanismos emotivos que desencadenaban la violencia, cuatro tipos ideales: miedo, ambición, frustración y exhibición, para interpretar los diversos episodios de violencia que presenció.

Concluimos este apartado con un resumen sobre la diferencia entre el enfoque cuantitativo y el cualitativo en relación con los dos mecanismos de la «explicación» y de la «interpretación». En las investigaciones cualitativas aquí descritas no encontramos ninguna pregunta sobre los mecanismos causales que han provocado diferencias de actitudes, conductas y estilos de vida entre los sujetos. Jenkins no se pregunta por qué los rebeldes tienen una idea del matrimonio distinta de la que tienen los ciudadanos, más bien intenta describir las diferencias entre ambas a la luz de las características generales de los dos tipos ideales. En otras palabras, mientras la investigación cuantitativa se pregunta por los *porqués,* la cualitativa se pregunta por los *cómo.* Denzin, un claro defensor del enfoque interpretativo, escribe: «En mi estudio sobre alcohólicos anónimos no he preguntado a las personas por qué se convirtieron en alcohólicos, sino cómo llegaron a ello. Esta forma de plantear las cuestiones lleva al meollo del proceso social, y no a las preocupaciones sobre variables causales previas [...]. Yo prefiero profundizar en la forma en que se produce un hecho o un proceso, más que preguntarme sólo por qué se ha producido, o cuál ha sido su causa» [Denzin 1989, 26].

El objetivo de la investigación cuantitativa es, en cambio, la identificación del mecanismo causal. Aunque no siempre es posible formular un «modelo causal», con variables independientes y dependientes conectadas claramente por nexos causales, como en el ejemplo citado de Blau y Duncan, la mente del investigador cuantitativo estará siempre guiada por la lógica del mecanismo causa-efecto. Hemos dado un ejemplo de ello en la investigación de Sampson y Laub, donde las variables se pueden agrupar en «causas», «efectos» y «condiciones», y el investigador se guía siempre por la pregunta de «qué causa qué» y bajo qué condiciones[8].

Por último, podemos añadir que el modelo causal enlaza las *variables* entre sí (lógica de la «causalidad»), mientras que la tipología representa el esquema teórico que relaciona a los *sujetos* (lógica de la «clasificación»).

Una última cuestión, sin duda importante por abarcar muchos de los temas tratados, es el *alcance de los resultados*. Ya hemos hablado de esta problemática a propósito del muestreo y la representatividad de los casos estudiados. Dadas sus exigencias de profundidad e identificación con el objeto estudiado, la investigación cualitativa no puede abordar un número de casos muy elevado. El ejemplo de la investigación de Jankowski, que participó en la vida de nada menos que 37 bandas, es prácticamente único, si bien es verdad que su observación participante duró diez años, hecho muy poco común. Normalmente los estudios cualitativos abarcan pocos casos. Con frecuencia se presenta el denominado «estudio de caso», una investigación sobre una situación concreta (una banda, un barrio, una fábrica, una organización, un acontecimiento, etc.). ¿Cómo es posible inferir, de situaciones tan específicas, aunque elegidas de modo que sean lo más representativas posible, observaciones y conclusiones de alcance general?

Sin duda, en una investigación sobre pocos casos se podrá profundizar más, pero se reducirán también las posibilidades de generalizar sus resultados. Como escribe Michael Patton:

> ... se puede estudiar a un único individuo durante un periodo de tiempo largo, por ejemplo, estudiar en profundidad una semana en la vida de un niño. Eso significa recopilar información detallada sobre todos los episodios que se produzcan y las interacciones que le afecten cada día. Si limitamos el campo de las preguntas de la investigación, podríamos estudiar a varios niños durante un periodo más breve. Con preguntas aun más dirigidas o una entrevista de media hora podríamos estudiar a un mayor número de niños sobre un número de preguntas reducido. El caso extremo consistiría en plantear una sola pregunta al máximo número de niños que permitieran los recursos económicos y de tiempo.
>
> [Patton, 1990, 166]

[8] Para mayor claridad, en la comparación entre métodos cuantitativos y cualitativos hemos contrapuesto «explicación» a «interpretación», «búsqueda de porqués» a «búsqueda de cómos», «causalidad» a «clasificación», «análisis por variables» a «análisis por casos», pero en la realidad los límites nunca son tan claros. Las tipologías también se usan en la investigación cuantitativa, y, lo que es más importante, el mecanismo causal está muy presente en muchos enfoques interpretativos, por ejemplo, el de Weber, que admite abiertamente la existencia de «leyes», aunque sean sólo instrumentos para comprender la conducta individual, y no constituyan el objetivo de la investigación social [cfr. Kaplan, 1964, 115].

Profundidad y amplitud están pues en relación inversa. Profundizar tiene un coste, y el precio a pagar depende del número de casos estudiados. Pero el número de casos va ligado a la posibilidad de generalizar los resultados. Cuanto más amplia sea la investigación, es decir, cuanto mayor sea el número de sujetos que abarca, más representativa de la situación real será la muestra y, por tanto, sus resultados podrán aplicarse con más legitimidad a toda la población, siempre que no se produzcan distorsiones sistemáticas [9]. En conclusión, los resultados de la investigación cuantitativa tienen sin duda más posibilidades de generalización que los resultados de la investigación cualitativa.

2.4. DOS FORMAS DIFERENTES DE CONOCER LA REALIDAD SOCIAL

Para finalizar este capítulo nos planteamos la siguiente pregunta: ¿Es mejor, más correcto científicamente, más adecuado cognitivamente, hacer investigación social utilizando la perspectiva cuantitativa o la cualitativa? ¿Se puede afirmar que uno de los dos enfoques es superior al otro desde el punto de vista «científico»? Distinguimos tres posturas al respecto.

La primera sostiene que el enfoque cuantitativo y el enfoque cualitativo, el paradigma neopositivista y el paradigma interpretativo, representan dos puntos de vista incompatibles, porque se caracterizan por planteamientos filosóficos divergentes. Los defensores de cada una de las posturas afirman que la suya es la adecuada, y la otra es errónea. Los partidarios del criterio cuantitativo argumentan que el enfoque cualitativo no es ciencia; los segundos, en cambio, aseguran que los primeros, con su cosificación de las ciencias naturales, no pueden captar la verdadera esencia de la realidad social.

La segunda postura, muy difundida entre los científicos sociales de orientación cuantitativa, es la de aquellos que, pese a haberse decantado por el paradigma neopositivista, reconocen que las técnicas cualitativas pueden aportar una contribución válida a la producción y el análisis de evidencia empírica. No obstante, se considera que son adecuadas sólo para la fase exploratoria de la investigación, pues cumplen una función de estimulación intelectual preliminar, pero quedan fuera de la fase científica propiamente dicha. El siguiente pasaje de Blalock ilustra este punto de vista sobre la investigación cualitativa:

> ... por lo general las técnicas de observación participante son muy útiles para proporcionar impresiones e ideas iniciales que después pueden conducir a una

[9] Tampoco vale la hipótesis de la «elección cuidada», es decir, la argumentación de que si se eligen con un criterio adecuado y la suficiente diversificación, unos cuantos casos pueden representar el conjunto de las variaciones presentes en la población. ¿Cómo podemos garantizar que los casos elegidos son los adecuados con relación al universo de posibilidades, si la investigación se realiza precisamente para conocer la realidad estudiada? Por otra parte, a veces los casos desviados son precisamente los más esclarecedores.

> formulación precisa del problema e hipótesis explícitas, pero se prestan a la crítica porque sus resultados se basan en experiencias subjetivas y difícilmente repetibles. Por este motivo muchos científicos sociales prefieren considerar la observación participante como un instrumento útil en la etapa inicial del proceso de investigación, pero no un planteamiento del que se puedan derivar resultados definitivos de la investigación.
>
> [Blalock, 1970; 45-46].

La tercera postura sostiene la legitimidad, utilidad y dignidad de ambos métodos, y anuncia el desarrollo de una investigación social que, según las circunstancias y las posibilidades, opte por un enfoque o el otro (o por ambos). Esta postura ha ganado adeptos en los últimos años, y procede, más que de una reflexión filosófica y epistemológica nueva, de la constatación pragmática de que la sociología y la investigación social han recibido contribuciones valiosas y fundamentales de investigaciones tanto cuantitativas como cualitativas. Albert Bryman afirma de manera explícita que las diferencias entre las dos maneras de hacer investigación no son epistemológicas, sino puramente técnicas:

> La distinción entre investigación cuantitativa e investigación cualitativa es una cuestión meramente técnica, en la que la elección entre una o la otra sólo tiene que ver con su capacidad para responder al problema particular planteado por la investigación [...], igual que otras decisiones técnicas [...], como la decisión sobre la conveniencia de utilizar el cuestionario postal o sobre las modalidades de construcción de una muestra estratificada.
>
> [Bryman, 1988, 109]

Bryman llega a sostener la utilidad de un enfoque que utiliza métodos cualitativos y métodos cuantitativos en la misma investigación. Esta postura también la adopta otro manual de metodología de investigación cualitativa, en el que los autores defienden, bajo el título de «dos estilos de investigación, una lógica de inferencia», que el marco de referencia subyacente está representado por la misma lógica [...]; las diferencias entre las tradiciones cuantitativas y cualitativas son sólo estilísticas, y no tienen importancia metodológica ni sustantiva» [King, Keohane y Verba, 1994].

De las tres posturas presentadas suscribo la tercera, pero con una matización importante. No creo que los métodos cuantitativos y cualitativos representen dos variaciones puramente técnicas de un mismo modo de entender el mundo social y la finalidad de la investigación. En mi opinión, las dos formas de hacer investigación, en contra de lo que afirma Bryman, no difieren entre sí por meras cuestiones técnicas, sino que son la expresión directa y lógica de dos perspectivas epistemológicas distintas, dos paradigmas diferentes que implican modos alternativos de comprender la realidad social, los objetivos de la investigación, el papel del investigador y la instrumentación técnica.

Pero ¿por qué la aceptación de diferencias entre ambos enfoques debe implicar también que uno sea correcto y el otro erróneo?, ¿no pueden dos visiones distintas de la misma realidad aportar dos contribuciones válidas para su conoci-

miento, como contribuyen a la ilustración de una ciudad tanto su fotografía panorámica como la de uno de sus callejones más característicos?

Hemos estudiado la investigación de Sampson y Laub, que acaba con la comprobación empírica de un modelo teórico sobre el proceso que lleva al individuo a cometer un delito y adoptar conductas ilícitas, analiza el proceso en términos de variables dependientes e independientes usando las categorías de causa y efecto y extrayendo datos cuantitativos sobre una muestra de 1000 sujetos. También hemos estudiado la investigación de Jankowski, que trata de comprender las motivaciones y razones por las que un joven entra en una banda y se deja llevar por la misma cometiendo actos de violencia, para lo cual se introduce en la vida de algunas bandas metropolitanas y participa en su vida cotidiana. ¿Podemos decir de estas dos perspectivas que una es científica o correcta y la otra es errónea o pseudo-científica, que una ha enriquecido nuestros conocimientos sobre la desviación juvenil y la otra nos ha dado una visión deformada y engañosa de la misma? En mi opinión, ambas investigaciones han contribuido de manera significativa al conocimiento del fenómeno social de la delincuencia juvenil.

Ahora bien, considero difícil, si no imposible, utilizar los dos planteamientos en un mismo diseño de investigación. Los procedimientos y los instrumentos utilizados difieren demasiado. En realidad, los ejemplos de estudios que utilizan ambos enfoques son investigaciones basadas en una de las dos perspectivas, que se sirven de técnicas tomadas del otro enfoque, pero como instrumentación auxiliar y complementaria. Me parece difícil que un mismo investigador pueda llevar a cabo una investigación que utilice de manera simultánea instrumentos de producción y análisis de datos cualitativos y cuantitativos. Su formación, e incluso la estructura de su personalidad científica, hacen muy difícil mostrar la flexibilidad necesaria para combinar ambos enfoques.

Para concluir, el enfoque neopositivista y el interpretativo, las técnicas cuantitativas y cualitativas conducen a conocimientos diferentes. Pero esto no es una limitación, sino una riqueza, ya que para poder conocer completamente la realidad social se precisa una aproximación polifacética y múltiple. La investigación social es como un dibujo de la realidad social. Para hacerlo se escoge una perspectiva determinada, pero puede haber infinidad de ellas, no sólo en términos del ángulo visual (dibujar a un personaje a lomos de un caballo en lugar de retratar su rostro o elegir el perfil de una figura en lugar de su vista frontal), sino también de la fidelidad a la apariencia formal (se puede favorecer la expresión de rasgos psicológicos con líneas y colores deformadores o poner al personaje en un contexto surrealista[10]). No existe un retrato absoluto, como no existe una representación absoluta y «verdadera» de la realidad.

[10] Hace ya varios siglos se discutía sobre si los pintores, con sus imágenes, construían una realidad «distinta» o bien representaban la «verdadera». Un ejemplo de ello es el debate entre las corrientes de lo sublime («constructivista») y de lo pictórico («realista»), en el siglo XVIII.

2.5. SÍNTESIS DEL CAPÍTULO 2

- El análisis secundario de Sampson y Laub sobre los datos de una encuesta a una muestra sobre delincuencia juvenil es un ejemplo de investigación cuantitativa que se inspira en el paradigma neopositivista. Dicho análisis ilustra un procedimiento muy sistemático, en el que cada capítulo sigue un esquema de cuatro fases: enfoque teórico, observación empírica, análisis de los resultados, vuelta a la teoría. El análisis de los datos se realiza sobre las variables utilizando técnicas estadísticas para producir modelos causales en los que las variables están conectadas por relaciones causa-efecto.
- El estudio de Jankowski es un ejemplo de observación participante, una de las técnicas clásicas de investigación cualitativa basada en el paradigma interpretativo. Aunque las preguntas eran similares a las de la investigación de Sampson y Laub, el método de trabajo es muy distinto. El autor participó personalmente en la vida de las bandas objeto de su estudio, registrando sus observaciones en cuadernos en el curso de la observación. Su objetivo no es identificar relaciones causa-efecto entre variables, sino comprender las motivaciones que subyacen a la conducta de los miembros de las bandas y construir clasificaciones y tipologías.
- Las diferencias entre investigación cuantitativa y cualitativa, consideradas no ya en función de sus premisas filosóficas y epistemológicas, sino en términos de sus aplicaciones empíricas concretas, se pueden comprender mejor si se examinan en las cuatro fases de la investigación empírica: diseño de la misma, obtención de información, análisis de los datos y producción de resultados.

  3.1. *Diseño de la investigación.* La diferencia entre investigación cuantitativa y cualitativa se encuentra en el diseño estructurado y predefinido, con hipótesis deducidas de la teoría, en el primer caso, frente al plan de trabajo abierto e interactivo de la investigación cualitativa, cuyas modalidades de desarrollo surgen en el curso de la propia investigación. Por otra parte, en la investigación cuantitativa la actitud del investigador hacia los sujetos estudiados es neutral y distante, mientras que en la investigación cualitativa se caracteriza por la empatía y la identificación.

  3.2. *Obtención de información.* La investigación cuantitativa suele trabajar con una muestra representativa del universo estudiado, con el objetivo de producir una «matriz de datos», es decir, recoger el material empírico de forma estandarizada, según un modelo idéntico para todos los casos. La investigación cualitativa no se plantea problemas de estandarización ni de representatividad, y prefiere tratar los casos que son objeto del estudio de forma heterogénea, según la relevancia de cada situación.

  3.3. *Análisis de los datos.* Mientras que en la investigación cuantitativa el análisis de los datos gira alrededor de las variables, es decir, las características de los casos estudiados, que se analizan con herramientas matemáticas y estadísticas, la investigación cualitativa analiza y estudia en profundidad a los sujetos, con el fin de comprenderlos.

3.4. *Producción de resultados.* El objetivo de la investigación cuantitativa es hacer generalizaciones, síntesis de orden superior tanto en el aspecto de la abstracción conceptual como en el aspecto del campo de aplicación. La investigación cualitativa, en líneas generales, se preocupa mucho menos por la generalización de sus resultados y dedica más atención a defender la especificidad de las distintas situaciones sociales que a identificar los rasgos comunes a todas ellas.

4. El enfoque neopositivista y el enfoque interpretativo, la investigación cuantitativa y la investigación cualitativa, llevan a conocimientos diferentes. Esto no es una limitación, sino una riqueza, dado que para poder conocer la realidad social se precisa un enfoque polifacético y múltiple.

2.6. LECTURAS COMPLEMENTARIAS

Un texto muy útil sobre la comparación entre investigación cuantitativa y cualitativa es el de A. Bryman, *Quantity and Quality in Social Research* (Routledge, 1988, 198 pp.). Siempre como introducción, sobre todo para ayudar al principiante a elegir entre ambos enfoques, puede consultarse la obra de J. W. Creswell *Qualitative and Quantitative Approaches* (Sage, 1994, 227 pp.).

Para una reconciliación entre los enfoques cuantitativo y cualitativo, y una evaluación de sus diferencias para estudiar mejor la realidad social, sugerimos dos libros: I. Newman y C. R. Benz, *Qualitative-Quantitative Research Methodology: Exploring the Interactive Continnum* (Southern Illinois U.P., 1998, 218 pp.) y A. Tashakkori y C. Teddlie, *Mixed methodology: Combining Qualitative and Quantitative Approaches* (Sage, 1998, 185 pp.).

Sobre la investigación cualitativa, el texto de referencia puede ser, por la amplísima y cuidada selección de material y reflexiones, N. K. Denzin y Y. Lincoln, *Handbook of Qualitative Research* (Sage, 2000, 1.065 pp); presenta un gran número de ensayos (41 para ser exactos), que reconstruyen la historia de los métodos cualitativos (señalamos el de A. J. Vidich y M. L. Stanford, *Qualitative Methods: Their History in Sociology and Anthropology*), los paradigmas de referencia, las distintas estrategias de investigación y los diferentes métodos para recopilar, analizar e interpretar el material empírico.

En cuanto al enfoque cuantitativo, sugerimos dos libros. El primero, P. S. Maxim, *Quantitative Research Methods in the Social Sciences* (Cambridge U.P., 1999, 405 pp.), ofrece un planteamiento más metodológico-epistemológico, se enfrenta a las bases filosóficas de la investigación empírica, la deducción estadística, la medición, el *scaling,* el diseño de la investigación, o el muestreo. La segunda obra, T. R. Black, *Doing Quantitative Research in The Social Sciences: An Integrated Approach to Research Design, Measurement and Statistics* (Sage, 1999, 751 pp.) es de corte más técnico y dedica más de la mitad de su contenido a los procedimientos estadísticos para la transformación de información en datos y el análisis de los mismos.

LA PRODUCCIÓN DE LOS DATOS: TÉCNICAS CUANTITATIVAS

En los seis capítulos siguientes abandonamos el discurso metodológico para adentrarnos en las técnicas de investigación específicas del enfoque cuantitativo. En realidad, el primero de estos capítulos, dedicado a la traducción empírica de la teoría, sigue tratando temas de carácter muy general, situados en un ámbito que podríamos definir como pretécnico y que no está limitado a la investigación cuantitativa. Aun así, hemos preferido incluir este tema dentro del enfoque cuantitativo, por dos motivos.

En la investigación cualitativa no existe el equivalente de la *operacionalización* de conceptos, que en el campo cuantitativo representa el núcleo central de la transformación de las hipótesis teóricas en proposiciones comprobables por métodos empíricos. Ya hemos mencionado que los conceptos de la investigación cualitativa son, como afirma Blumer, «conceptos orientadores», marcos de referencia abiertos, cuyo objetivo es orientar al investigador y hacerlo más sensible a temas y problemas concretos en el curso de la entrevista. Están, por tanto, exentos de ese carácter de explicación y definición derivado de su *operacionalización*, propia del enfoque cuantitativo.

Hay que añadir, además, que el propio concepto de «traducción empírica de la teoría» lleva implícita la tesis de que la teoría precede a la investigación, tesis que rechaza gran parte de la investigación de inspiración cualitativa. Es conocido el planteamiento que sostiene que los conceptos y las hipótesis no deben preceder a la observación, porque la teoría está «arraigada» a la realidad observada y la tarea del investigador es descubrirla. Según los autores que defienden este punto de vista, el hecho de extraer de la realidad los conceptos y las hipótesis en lugar de imponerle esquemas teóricos preconcebidos, permite formular teorías «que se adaptan mejor a los datos y funcionan mejor, puesto que las categorías se descubren mediante el examen de los propios datos», y que resultan más fáciles de comprender para el hombre común, al deducirse con su misma forma de pensar. Ello permite también evitar «el uso oportunista de teorías de dudoso valor y de dudosa capacidad explicativa», adoptadas a menudo por los investigadores sólo por pereza intelectual y adhesión a las modas dominantes [Glaser y Strauss, 1967, 3-4].

Es evidente que estos planteamientos son incompatibles con la secuencia «teo-

ría-hipótesis-recopilación de datos-análisis-vuelta a la teoría» que representa, como veremos, el modelo de estructura de la investigación cuantitativa clásica.

Por estas razones hemos preferido situar dentro del marco de la investigación cuantitativa un tema tan general como el del paso de la teoría a la realidad empírica. Hay que decir, en todo caso, que la parte inicial de este capítulo, con las nociones de teoría, hipótesis y conceptos, supera los límites del enfoque cuantitativo y abarca todo el contexto de la teoría y la investigación social.

CAPÍTULO 3

LA TRADUCCIÓN EMPÍRICA DE LA TEORÍA

3.1. ESTRUCTURA «TIPO» DE LA INVESTIGACIÓN CUANTITATIVA 68

3.1.1. Las cinco fases del proceso de investigación 69

3.2. DE LA TEORÍA A LAS HIPÓTESIS 71

3.2.1. Teoría 71
3.2.2. Hipótesis 72
3.2.3. La transformación de la teoría en hipótesis 73

3.3. DE LOS CONCEPTOS A LAS VARIABLES 76

3.4. UNIDAD DE ANÁLISIS 79

3.4.1. Los distintos tipos de unidad de análisis 80

3.5. VARIABLES 81

3.6. VARIABLES NOMINALES, ORDINALES Y CARDINALES 84

3.6.1. Variables nominales 85
3.6.2. Variables ordinales 87
3.6.3. Variables cardinales 88

3.7. CONCEPTOS, INDICADORES E ÍNDICES 91

3.8. ERROR TOTAL 95

3.9. FIABILIDAD Y VALIDEZ 99

3.9.1. Fiabilidad 99
3.9.2. Validez 101

3.10. Síntesis del capítulo 3 102

3.11. LECTURAS COMPLEMENTARIAS 104

En este capítulo hablaremos de la operacionalización, es decir, la transformación de hipótesis teóricas en operaciones de una investigación empírica. En primer lugar explicaremos qué se entiende por teoría, hipótesis y conceptos. Después presentaremos la noción de «variable», que representa el concepto clave de la investigación social empírica, y analizaremos los diferentes tipos con los que se representa. Por último, abordaremos el problema general de los errores que se pueden cometer en el proceso de traducción de la teoría a la investigación empírica.

3.1. ESTRUCTURA «TIPO» DE LA INVESTIGACIÓN CUANTITATIVA

Para empezar conviene decir que *la investigación científica es un proceso de descubrimiento creativo que sigue un itinerario prefijado y unos procedimientos preestablecidos y consolidados dentro de la comunidad científica.* Este paralelismo entre la «creatividad» y los sucesivos «procedimientos preestablecidos» no debe parecer contradictorio. El atributo «creativo» es importante en esta definición, evoca las capacidades personales del investigador, su perspicacia e inteligencia intuitiva. Pero esto es sólo un aspecto del proceso de investigación científica.

Reichenbach introdujo una distinción entre la concepción de una idea nueva y la presentación, justificación, defensa y comprobación de la misma, momentos que denominó, respectivamente, *contexto del descubrimiento* y *contexto de la justificación*. Según Reichenbach, en el primer estadio no es posible establecer reglas ni procedimientos. «El acto del descubrimiento escapa al análisis lógico; no hay reglas lógicas con las que construir una "máquina descubridora" que desempeñe la función creativa del genio» [Reichenbach, 1951; 231]. Pero la labor del científico no consiste únicamente en producir teorías, sino también en comprobarlas. Y esta fase, tanto si aborda el examen lógico-formal de la teoría, como su congruencia con la realidad, debe seguir unas reglas muy precisas.

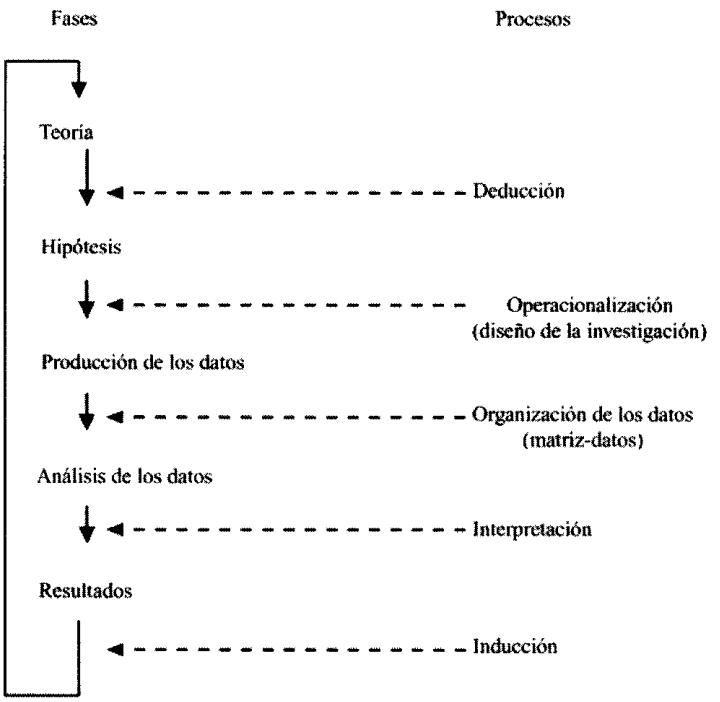
La primera regla fundamental de la investigación empírica es que debe desarrollarse dentro de un marco aceptado por la comunidad científica. La investigación científica, por contraste con el arte, es un proceso colectivo. «La ciencia es pública y no privada», escribió Merton. Este aspecto colectivo-público de la ciencia tiene una doble connotación. Por una parte, implica el *control:* «Los conceptos y procedimientos empleados, incluso por el sociólogo más intuitivo, deben estar estandarizados, y los resultados de sus intuiciones deben poder ser comprobados por otros investigadores» [Merton, 1968]. Por otra parte, implica la *acumulación:* «Si conseguí llegar más lejos que otros hombres, es porque me subí a hombros de gigantes», afirmó Isaac Newton, frase que Merton utilizó para construir un brillante ensayo sobre sociología de la ciencia [Merton, 1965]. Recordemos también que una de las definiciones más comunes de la ciencia es «acumulación sistemática de conocimientos».

El marco de referencia colectivo que debe tener presente el científico social al realizar su investigación, sobre todo en el enfoque cuantitativo, consta de dos elementos: la estructura lógica del proceso de investigación y la instrumentación técnica a utilizar. En el siguiente apartado trataremos el primero de estos elementos.

3.1.1. Las cinco fases del proceso de investigación

El *proceso* que sigue el investigador social para la realización de una investigación consiste en un recorrido cíclico que parte de la teoría, pasa por las fases de recopilación y análisis de los datos, y vuelve a la teoría. Este proceso se describe de manera más o menos similar en todos los manuales de investigación empírica, y para explicarlo podemos utilizar el esquema de la figura 3.1, donde distinguimos cinco fases y cinco procesos que conectan estas fases.

La primera fase es la de la *teoría.* La segunda es la de las *hipótesis,* y el paso entre las dos se produce mediante el proceso de *deducción.* La hipótesis representa una formulación parcial de la teoría, y se sitúa a un nivel inferior en cuanto a generalidad [1]. La teoría es «general» y la hipótesis es «específica». Por ejemplo, de la proposición teórica general que afirma que «existe una correlación positiva entre la participación política y el prestigio social» pueden deducirse, entre otras, tres



FUENTE: Adaptado de Bryman [1988, 20].

Figura 3.1. **Estructura de la investigación cuantitativa.**

[1] Y algunas veces también en cuanto al ámbito de aplicación, es decir, el área geográfica o lapso de tiempo a los que se refiere la hipótesis.

hipótesis más específicas: que la participación electoral es superior en el caso de los varones, los adultos (en comparación con los jóvenes y los ancianos), y las personas con más éxito profesional. Al hacer esto, hemos limitado el concepto de participación política al de participación electoral, y el de prestigio social a los tres aspectos mencionados.

La tercera fase es la de *recopilación de datos,* y llegamos a ella mediante el proceso de *operacionalización,* que consiste en la transformación de las hipótesis en afirmaciones observables empíricamente. Este proceso es muy complejo y se puede dividir en dos partes. La primera consiste en la *operacionalización de los conceptos,* o sea, la transformación de los conceptos en variables, o entidades medibles. Por ejemplo, el concepto de éxito profesional se podrá operacionalizar mediante los ingresos, o la obtención de un empleo mejor que el de nuestros padres, etc. La segunda consiste en la *elección del instrumento* y los procedimientos para medir los conceptos. En el ejemplo de la participación política se podría trabajar con datos agregados, utilizando las estadísticas oficiales de voto y averiguando si la participación electoral varía con el género y la edad; también se podría recurrir a datos de encuestas de opinión pública, o incluso servirse de entrevistas en profundidad sobre una muestra de ciudadanos. La decisión sobre estos aspectos nos lleva a la definición del *diseño de la investigación*, un «plan de trabajo» que establece las diversas fases de la observación empírica (por ejemplo: se entrevistará a tantas personas, ubicadas en las siguientes áreas geográficas..., elegidas con el siguiente criterio..., utilizando el siguiente cuestionario..., etc.).

Después de la recopilación de material empírico, llegamos a la fase del *análisis de los datos,* que será precedida por la *organización y sistematización de los datos obtenidos.* Por lo general se da el nombre de *información* al material empírico bruto que aún no ha sido sistematizado, y el de *datos,* a ese mismo material organizado de tal forma que se pueda proceder a su análisis. Pensemos, por ejemplo, en un paquete de cuestionarios y varias cintas de entrevistas grabadas (información), y esos mismos cuestionarios o entrevistas transformados en secuencias de números, comparables entre sí y obtenidas atribuyendo un número a cada respuesta (datos).

El proceso de organización de los datos suele consistir en transformar la información en una matriz de números, la *matriz de datos,* llamada también matriz «casos por variables» (C × V): ésta se obtendrá, por ejemplo, codificando de forma sistemática las respuestas a los cuestionarios, estandarizando las respuestas de las entrevistas abiertas, uniformando las fuentes estadísticas, aplicando una plantilla de análisis del contenido a los textos, etc. La matriz de datos obtenida representa el punto de partida para el análisis de los datos, que por lo general consiste en operaciones matemáticas realizadas con un programa de análisis estadístico.

La quinta fase consiste en la presentación de los *resultados,* mediante un proceso de *interpretación* del análisis estadístico realizado en la fase anterior. Por último, el investigador vuelve al punto inicial, es decir, a la *teoría,* mediante un proceso de *inducción,* que consiste en comparar los resultados empíricos con las hipótesis teóricas y el conjunto de la teoría, para confirmarla o reformularla.

Como puede verse, el esquema que hemos expuesto empieza con la teoría y

vuelve a ella, en un proceso cíclico que podríamos visualizar como una *espiral,* donde teoría e investigación están ligadas en el proceso de acumulación de conocimientos. Nótese que la sucesión descrita representa el proceso ideal para la investigación social de tipo cuantitativo, y que el esquema básico puede sufrir variaciones significativas en cada aplicación concreta del mismo. Por ejemplo, en relación con la teoría, el equilibrio entre deducción e inducción puede variar. Aunque el esquema básico prevé que la investigación no comience antes de que se haya realizado un estudio exhaustivo de la teoría y se hayan formulado hipótesis muy precisas, hay investigaciones que parten de un embrión de teoría, y ésta no se desarrolla ni toma cuerpo hasta la fase de análisis del material empírico. Otras veces se obvia toda la fase de la recopilación de datos, incluidos los procesos de operacionalización y organización de los mismos, porque se usan datos ya existentes para la investigación. Pero en cualquier caso el esquema de referencia sigue siendo el que aparece en la figura 3.1.

En este capítulo trataremos los procesos de deducción y operacionalización. En el apartado siguiente nos ocuparemos del primero. El resto del capítulo lo dedicaremos íntegramente al proceso de operacionalización de conceptos: la atención dedicada a este tema queda justificada por su complejidad y su relevancia en el proceso de investigación.

3.2. DE LA TEORÍA A LAS HIPÓTESIS

3.2.1. Teoría

Podemos definir una *teoría* como un *conjunto de proposiciones relacionadas de manera orgánica, que se formulan con un alto grado de abstracción y generalización respecto a la realidad, se deducen de hechos empíricos que se producen con regularidad y de las que pueden inferirse predicciones empíricas.* A continuación explicamos con más detalle los cuatro elementos de esta definición.

Conjunto de proposiciones: sistema coherente de afirmaciones que con frecuencia tiene estructura causal. *Abstracción y generalización:* la teoría va más allá de las expresiones empíricas concretas, en el aspecto conceptual y en el campo de aplicación. *Se deduce de hechos empíricos que se producen con regularidad:* la teoría surge de la constatación de repeticiones en la realidad observada, y se confirma con ellas. *Inferencia de predicciones empíricas:* de una teoría que surge de la observación de hechos empíricos que se repiten se pueden deducir hechos en otros contextos diferentes.

Ilustramos lo apuntado hasta aquí con un ejemplo clásico: la teoría sobre el suicidio de Durkheim. Como es sabido, Durkheim identifica tres «tipos ideales» de suicidio: el suicidio egoísta, el altruista y el anómico. En su teoría encontramos la siguiente proposición causal que representa el fundamento teórico del suicidio egoísta: «Cuanto más elevada sea la tasa de individualismo en un determinado grupo social, mayor será la tasa de suicidios en ese grupo». Por individualismo entendemos una situación social y cultural en la que la persona es plenamente libre y está sola para tomar decisiones en la vida; en el extremo

opuesto se encuentra la situación de cohesión social, en la que la actividad de una persona está sometida a controles sociales de manera permanente, y sus decisiones vienen determinadas en gran parte por normas colectivas.

La proposición teórica enunciada, que vincula el individualismo con el suicidio, es una *proposición causal,* que presenta las características siguientes: *abstracción,* puesto que los conceptos mencionados de individualismo, cohesión social y suicidio tienen significados abstractos respecto a la realidad observada; *generalización,* dado que la conexión que Durkheim propone se considera válida para una gran parte de la sociedad, cuando no para todas las sociedades; *se deduce de hechos empíricos que se repiten,* ya que Durkheim, como describe en su investigación, dedujo y comprobó empíricamente su teoría mediante el análisis de las fuentes estadísticas de su época; genera *predicciones empíricas,* porque permite prever, por ejemplo, que Irlanda tendría una tasa de suicidios inferior a la de Inglaterra.

3.2.2. Hipótesis

Una proposición teórica debe poder expresarse en hipótesis específicas. Una *hipótesis* es una *proposición que implica una relación entre dos o más conceptos, situada en un nivel inferior de abstracción y generalidad con relación a la teoría y que permite una traducción de ésta en términos que se puedan someter a prueba empírica.* Las dos características distintivas de la hipótesis son, por una parte, su menor abstracción (o mayor concreción) en términos conceptuales, y su menor generalidad (o mayor especificidad), en términos de extensión, con relación a la teoría. Y por otra parte, su provisionalidad: la hipótesis es una afirmación provisoria que se debe comprobar, derivada de la teoría pero que precisa una comprobación empírica para poder confirmarse.

De la teoría general de Durkheim pueden deducirse una serie de hipótesis específicas. Por ejemplo, es de esperar una tasa de suicidio más alta en las sociedades protestantes que en las católicas, por el hecho de que la religión protestante confía las decisiones a la libertad individual, mientras que la católica es impositiva y normativa; una tasa de suicidio más baja en las personas casadas y con hijos, por el conjunto de exigencias, vínculos, controles y deberes que conlleva la situación familiar; una tasa de suicidio más baja cuando una sociedad atraviesa un periodo de crisis política o está en guerra, por la consolidación de los lazos sociales entre ciudadanos expuestos al mismo peligro; y así sucesivamente.

Estas hipótesis se pueden *comprobar de manera empírica* mediante definiciones operativas adecuadas: por ejemplo, se definirá el concepto de «integración familiar» considerando que las personas casadas y con hijos tienen el grado máximo de integración, y las personas que viven solas el grado mínimo, con los grados intermedios necesarios, y se verá si en los datos existe una asociación entre esta variable y el suicidio. También se podrá constatar, como hizo Durkheim, que en igualdad de condiciones en distintas sociedades, existe una relación entre la tasa de suicidio y el tipo de religión predominante.

La validez de una teoría depende de su capacidad para transformarla en hipóte-

sis que se pueden comprobar de manera empírica. *El criterio de la verificación empírica es el criterio del carácter científico.* Si una teoría es imprecisa y confusa, difícilmente se podrá transformar en hipótesis comprobables, y sin una confirmación empírica, no superará el ámbito preteórico de las conjeturas. En las ciencias sociales el riesgo de que se formulen teorías vagas, demasiado confusas como para operacionalizarlas, es especialmente elevado. En este aspecto compartimos la máxima de que «es mejor una teoría errónea que una teoría imprecisa».

3.2.3. La transformación de la teoría en hipótesis

A continuación presentamos algunos ejemplos de transformación de una teoría en hipótesis concretas. El primer ejemplo se refiere a la teoría formulada por Inglehart a mediados de los años setenta sobre el cambio de valores en los países occidentales. Según esta teoría, los valores de las generaciones que crecieron en la posguerra serían distintos de los valores de las generaciones anteriores, y el cambio consistiría en el desplazamiento de la atención prestada a aspectos como la seguridad física y económica hacia otras cuestiones como la calidad de vida, la autorrealización, y la satisfacción intelectual y estética, lo que supone pasar de una orientación predominantemente materialista a otra que se ha definido como «postmaterialista». Las causas de este cambio serían varias. Aquí comentaremos sólo un aspecto, quizás el más importante: los cambios de las condiciones de vida durante los años de formación de los ciudadanos.

Las generaciones de la posguerra no han sufrido las crisis económicas que afectaron a las generaciones que crecieron entre las dos guerras mundiales y durante la Gran Depresión; además, en la posguerra las sociedades occidentales conocieron una prosperidad sin precedentes que permitió satisfacer las necesidades de subsistencia a prácticamente toda la población. Por otra parte, las nuevas generaciones no han vivido la experiencia de la guerra, que no obstante, sí afectó directamente a sus padres y abuelos. Según la teoría de las necesidades, el hombre empieza a poner en primer plano objetivos no materiales, sólo después de conseguir la seguridad física y económica.

De esta teoría general pueden derivarse hipótesis específicas. Primero tendremos que encontrar en todos los países occidentales una diferencia notable en las escalas de valores de jóvenes y ancianos, como consecuencia del hecho de que la orientación de valores que se constituye en los años de formación tiende a mantenerse durante toda la vida aunque las condiciones sociales y ambientales cambien.

La segunda hipótesis se refiere a la distinta magnitud de esta diferencia generacional en los distintos países. Si la teoría es válida, la diferencia debería ser más marcada en países donde ha sido mayor la variación de la calidad de vida entre los que crecieron antes y después de la Segunda Guerra Mundial. Por ejemplo, en Alemania esta diferencia debería ser máxima: durante la Primera Guerra Mundial los alemanes sufrieron escasez de recursos, una fuerte inflación, depresión económica, graves conflictos internos y la tragedia, en términos de destrucción del país y pérdidas de vidas humanas, de la Segunda Guerra Mundial. En cambio, en la pos-

guerra vivieron un periodo de bienestar económico y paz social notables. El caso de Inglaterra debería ser opuesto: antes de la Segunda Guerra Mundial era el país europeo más rico, el único de los países europeos que no vivió la guerra en su territorio ni sufrió la invasión extranjera. Sin embargo, en la posguerra experimentó un estancamiento económico persistente que provocó niveles de bienestar inferiores a los de muchos otros países europeos y sin duda por debajo de los que tenía en la primera mitad del siglo. Por tanto, en este caso se puede esperar una diferencia menor en la escala de valores entre generaciones.

La tercera hipótesis concierne también a la diferencia entre países, pero se refiere no tanto al distanciamiento generacional como a la cifra absoluta de quienes abrazan los valores postmaterialistas. Éstos deberían ser más numerosos cuanto más rico sea el país, y, por tanto estarían presentes sobre todo en Bélgica, Holanda y Suiza, y, en menor medida en España, Grecia e Italia [cfr. Inglehart, 1977].

Todas las hipótesis que hemos mencionado se pueden operacionalizar fácilmente y, por tanto, se pueden someter a su comprobación empírica.

En este punto debemos hacer una aclaración sobre la conexión entre teoría e hipótesis. En la práctica, la investigación no siempre se realiza en el orden descrito, primero la elaboración de la teoría y después la formulación de hipótesis que sirven de guía para el «diseño de la investigación». Con frecuencia sucede que las hipótesis se formulan *después* de la recopilación de datos y se comparan con estos datos a posteriori.

Tomemos como ejemplo una investigación sobre decisiones escolares individuales. El autor analiza la decisión de seguir o no estudiando después de la enseñanza secundaria y primaria, comparando tres concepciones distintas de la acción individual. En primer lugar, el punto de vista «estructuralista» sostiene «que los actos humanos dependen de vínculos externos que no dejan espacio para tomar decisiones realmente importantes; en segundo lugar, el punto de vista de la «no intencionalidad», que considera que el individuo está condicionado, aunque no tanto por factores estructurales externos como por factores psicológicos internos, la mayoría inconscientes; y, finalmente, el punto de vista de la «intencionalidad», que «supone que los individuos actúan [...] de acuerdo con ciertos propósitos, persiguiendo sus objetivos lo mejor que pueden» [Gambetta, 1987, 8, 16].

Como puede verse, se trata de una formulación clásica de la teoría de la acción individual. Lo que nos interesa es que, para dar respuestas empíricas a estas preguntas, los autores de los dos ejemplos utilizan los datos *ya existentes* de dos encuestas por muestreo realizadas con fines más bien generales (descripción de las orientaciones políticas y culturales de los jóvenes). Las encuestas contenían también preguntas sobre la intención de seguir estudiando, que se utilizaron para responder a las cuestiones antes mencionadas. El autor parte de tres hipótesis concretas de conducta, y analiza sus datos para ver cuál de ellas se cumple efectivamente en el caso empírico analizado. En este ejemplo, la elaboración teórica, aunque es posterior a la recopilación de datos, precede al análisis de los mismos.

Otras veces se recurre a la teoría *después* de analizar los datos para explicar un hecho anómalo o un resultado inesperado. Por ejemplo, en una investigación sobre abstencionismo electoral en Italia realizada en los años ochenta [Corbetta y Parisi,

1987], los autores plantearon dos hipótesis: que el abstencionismo electoral es un fenómeno de *apatía* política debido a la poca capacidad de los partidos para movilizar al electorado con sus aparatos organizativos, o bien un fenómeno de *protesta* por parte de los electores para mostrar a los partidos descontento y resentimiento.

En la primera hipótesis los autores pronosticaron una mayor abstención en las capas más periféricas de la sociedad, menos accesibles para la capacidad de movilización y comunicación de los partidos (el caso extremo serían las mujeres ancianas de las zonas rurales del sur de Italia, donde las cuatro variables de género, edad, ruralismo-urbanización y norte-sur tienen los valores más marginales). En la segunda hipótesis los autores pronosticaron un mayor abstencionismo entre los jóvenes varones de las grandes ciudades del norte, clases sociales centrales expuestas a los flujos de nuevas ideas, lo que mostraría un desacuerdo entre la parte más moderna del país y la política de épocas anteriores.

Sin embargo, los datos arrojaron unos resultados sorprendentes, contrarios a las dos hipótesis. Las categorías más abstencionistas resultaron ser las mujeres de las capitales de provincia. Un análisis detallado del perfil de abstencionistas reveló que se trataba en gran parte de mujeres que habían votado anteriormente a la Democracia Cristiana. Para explicar este resultado imprevisto, los autores recurrieron a la teoría de las «presiones cruzadas», según la cual «cuando el elector es sometido de manera simultánea a estímulos de igual intensidad pero opuestos, es probable que eluda el conflicto de la elección [...] y opte por la abstención» [*ibidem*, 59]. La debilidad de la religión en una sociedad secularizada y la distancia deliberada de la Iglesia del compromiso político, junto con la existencia de un nuevo sistema de valores canalizado sobre todo en los medios de comunicación de masas, creaban en el electorado femenino, hasta entonces sensible al mensaje de la Iglesia, una situación de punto muerto entre los valores antiguos y nuevos que generaba incertidumbre y, por tanto, abstención. Este proceso se iniciaba en las grandes ciudades, lugares de cambio por antonomasia.

En el caso citado los autores recurrieron a una teoría diferente, utilizada en otras ocasiones para interpretar la conducta electoral. Otras veces puede descubrirse una nueva teoría durante la fase empírica. Un ejemplo muy conocido es el de los experimentos realizados en los años veinte en las fábricas de la Western Electric de Hawthorne (cerca de Chicago). Los experimentos pretendían estudiar un problema muy simple: conocer cómo afectaba la variación de la iluminación al rendimiento de los trabajadores de un departamento concreto. En un principio se constató que el aumento de la iluminación hacía aumentar la productividad. Pero después se vio que, contra todo pronóstico, al volver a los niveles de iluminación iniciales, la productividad seguía siendo elevada.

Tampoco se confirmaron las expectativas de los investigadores cuando se dividió a las trabajadoras en un grupo experimental (para el que se hacía variar la iluminación) y un grupo de control (para el que la iluminación se mantenía constante): ambos grupos tenían un rendimiento superior a la media, sin variaciones perceptibles vinculadas al nivel de iluminación. Entonces un equipo de sociólogos de la organización dirigido por Elton Mayo estudió el caso y se fijó en otros factores no ligados estrictamente a la comodidad ambiental, empleando otras técnicas

como la observación participante y la entrevista. De este modo se descubrió que la variable determinante del rendimiento laboral no era la iluminación (ni otras variables de carácter técnico u organizativo, como la duración del descanso para el café, etc.), sino el hecho de que el grupo de trabajadoras se sintiera centro de atención y objeto de un experimento.

A partir de los resultados de estos experimentos, se empezó a perfilar una nueva corriente teórica dentro de la sociología laboral, la de las denominadas «relaciones humanas», que incluía entre las condiciones para un rendimiento laboral óptimo variables psicológicas como la cohesión del grupo de trabajo, la percepción por parte del individuo de la importancia que tiene en su grupo de trabajo y la utilidad de la labor que desempeña, etc.

Sobre la alteración de la secuencia teoría-hipótesis-observación hay que decir que en ocasiones la observación viene antes que las hipótesis por causas de fuerza mayor. Es el caso del «análisis secundario», que consiste en aplicar un segundo análisis a datos recopilados previamente por otros investigadores (la investigación de Gambetta es un ejemplo de ello, así como *Crime in The Making,* de Sampson y Laub, que comentamos en el capítulo 2).

Conviene señalar que, a menudo, la teoría no está tan bien definida como para admitir hipótesis claras y precisas. A veces el tema es nuevo, desconocido, y por tanto la investigación se mueve en un plano eminentemente descriptivo. Otras veces la observación no parte de una teoría específica, sino que pretende abarcar una amplia gama de problemas, para permitir sucesivos análisis diversificados (es el caso de encuestas periódicas como los barómetros del CIS en España, la encuesta del Eurobarómetro, la encuesta europea de valores, etc.).

3.3. DE LOS CONCEPTOS A LAS VARIABLES

Con el término «concepto» nos referimos al *contenido semántico* (es decir, al significado) *de los signos lingüísticos y las imágenes mentales.* Su origen etimológico (del latín *cum capio* = tomar de manera conjunta), expresa la acción de ordenar varios elementos bajo un único acto de pensamiento, y el acto de extraer de la inmediatez de las impresiones sensibles y las representaciones concretas una abstracción con significado universal. Es el instrumento que permite al hombre conocer y pensar, y es el fundamento de toda disciplina científica, que consiste en conocer mediante conceptos universales.

De esta definición se desprende que «el término tiene un significado muy general, y puede incluir todo tipo de signos o procedimientos semánticos, sea cual sea el objeto al que se refiere, abstracto o concreto, cercano o lejano, universal o particular. Se puede tener un concepto de una mesa o del número tres, del hombre o de Dios, del género y de la especie [...] de un periodo o una institución históricos (el Renacimiento o el Feudalismo)» [Abbagnano, 1971, 146]. Los conceptos, además, pueden hacer referencia a construcciones mentales abstractas que no se pueden observar de manera directa, como el poder, la felicidad o la clase social, o bien referirse a entidades concretas y observables, como la flor o el obrero.

En las páginas anteriores hemos definido la *hipótesis* como una *conexión entre conceptos* o una *proposición que implica una relación entre dos o más conceptos*. La hipótesis de que la tasa de suicidios es más elevada entre las personas más instruidas implica un nexo, una relación, entre el concepto de suicidio y los conceptos de formación o religión. La hipótesis de que los valores postmaterialistas están más difundidos entre los jóvenes de las clases burguesas que entre los de la clase obrera establece una relación entre el concepto de valor y el de clase social.

Podemos decir, por tanto, que los conceptos son los «ladrillos de la teoría», y que mediante la operacionalización de los conceptos realizamos la traducción empírica de una teoría. El concepto tiende un puente que conecta la orilla teórica con la orilla del mundo empírico observable. Como escribió Blumer, «la teoría tiene sentido para la ciencia sólo en la medida en que logra conectarse con el mundo empírico de manera provechosa; los conceptos son los únicos instrumentos que permiten establecer tal conexión» [Blumer, 1969, 143].

Si la teoría es una red de conexiones entre entidades abstractas representadas por los conceptos, cuando estas entidades se hagan concretas, toda la red teórica se concretará también, y será posible establecer las mismas conexiones entre los conceptos transformados en entidades observables. Si la hipótesis teórica es que los valores postmaterialistas están más difundidos en las sociedades ricas, cuando estemos en condiciones de medir empíricamente la presencia de los valores postmaterialistas y la riqueza en distintas sociedades, podremos también verificar la validez de la teoría, observando si los datos de los dos conceptos operacionalizados muestran una correlación positiva.

¿Cómo se operacionaliza un concepto? Pensemos en los conceptos típicos de las ciencias sociales: poder, clase social, autoritarismo, participación electoral, desviación, subdesarrollo, etc. ¿Cómo podemos transformar estos conceptos en entidades observables?

El primer paso del proceso de traducción empírica de un concepto consiste en aplicarlo a objetos concretos, es decir, *convertirlo en atributo o propiedad de objetos concretos,* que llamaremos *unidades de análisis*. El concepto de poder, por ejemplo, puede ser una propiedad de unidades de análisis como cargos empresariales, cargos políticos (consejero municipal, funcionario de un partido, diputado, etc.). Así mismo, los conceptos de clase social y autoritarismo pueden ser propiedades de individuos; los conceptos de participación electoral y desviación pueden caracterizar a individuos o unidades territoriales (en este caso en forma de porcentajes de votantes y tasas de homicidios); el concepto de subdesarrollo puede ser una propiedad de los países. Nótese que estos conceptos del análisis social no tienen por qué ser necesariamente complejos: el género y la edad, el lugar de residencia y el tiempo empleado para ir de casa al trabajo también son ejemplos de conceptos más fáciles de medir.

Estas propiedades se presentan en distintos grados en los diferentes objetos, es decir, varían en las distintas unidades de análisis. Por ejemplo, unos cargos empresariales tienen más poder que otros, la clase social varía entre los distintos sujetos, la participación electoral y la criminalidad varían en los diferentes municipios.

El segundo paso para operacionalizar el concepto-propiedad consiste en formular una *definición operativa* de mismo, es decir, *establecer las reglas para su traducción en operaciones empíricas*. Por ejemplo, se puede definir operativamente el poder de un cargo empresarial por el número de personas que tiene a su cargo; el autoritarismo se puede operacionalizar con las respuestas a preguntas sobre actitudes, con puntuaciones entre 0 y 5; la participación electoral y la criminalidad de un municipio se pueden operacionalizar con el porcentaje de votantes sobre el número de electores en unas elecciones concretas y el tanto por mil de delitos sobre el total de la población.

El tercer paso consiste en *aplicar las* mencionadas *reglas a los casos concretos estudiados:* ésta es la fase de operacionalización en su sentido más estricto. La definición operativa se hace en el plano teórico, con la operacionalización se pasa a la práctica. Como escribió Marradi, la definición operativa es un «texto», mientras que la operacionalización es un «hecho».

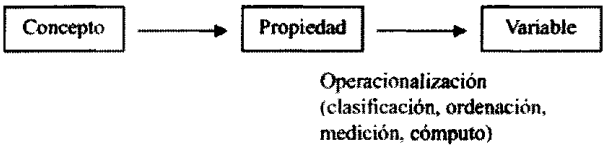
La propiedad operacionalizada se denomina *variable*, y las *modalidades* son los «estados» operacionalizados de la propiedad. A cada modalidad se le asigna un *valor* simbólico distinto, que suele ser un número. Por ejemplo, el concepto de «nivel cultural» se puede representar con la propiedad «nivel de estudios» referida a los individuos, que adoptará diferentes estados en los distintos sujetos estudiados; estos estados se clasifican en alguna de las cinco modalidades siguientes: «sin estudios», «graduado escolar», «enseñanza secundaria», «diplomatura» y «licenciatura»; a cada una de estas modalidades les asignamos los *valores* 0, 1, 2, 3, 4.

Llegados a este punto conviene hacer una aclaración sobre el término «operacionalizar». En el lenguaje corriente habríamos dicho «medir», puesto que la medición es el proceso de asignación de valores numéricos a los estados de una propiedad. De hecho, la definición de medición que se suele usar en todos los manuales de metodología procede de las formulaciones de Stevens en los años cuarenta, según las cuales «la medición consiste en la atribución de números a objetos o acontecimientos siguiendo determinadas reglas» [Stevens, 1946].

En realidad, como ha señalado Marradi [1981, 602 sq.], cuando no se dispone de unidades de medida es inadecuado hablar de medición: en estos casos el paso de la propiedad a la variable es una operación distinta a la *medición* (por ejemplo, medimos la edad con la unidad de medida del año solar), que, según los casos, es una *clasificación* (por ejemplo, clasificamos la nacionalidad según la nación), una *ordenación* (por ejemplo, ordenamos las profesiones en función de su prestigio social), o un *cómputo* (por ejemplo, contamos el número de hijos de una persona).

Sin embargo, en el lenguaje común no existe un término único para designar esa operación general de medición-ordenación-cómputo-clasificación. Las ciencias naturales, donde las unidades de medida casi siempre son definibles, han impuesto el término «medir» para referirse a la operación general. Pero para la clasificación o la ordenación el término «medir» resulta inadecuado, por ello hemos empleado el término «operacionalizar». Este término tiene un sentido amplio y se emplea con el significado general de «traducir del lenguaje teórico al lenguaje empírico», pero en sentido estricto se refiere al *paso de las propiedades a las variables,* que, en el proceso más amplio que conecta la teoría y la investigación, representa el puente

crucial entre las dos vertientes, y que puede esquematizarse del siguiente modo:



3.4. UNIDAD DE ANÁLISIS

La «unidad de análisis», como ya se ha indicado, representa el objeto social al que se refieren las propiedades estudiadas en la investigación empírica. Una reflexión teórica no necesita definir unidades de análisis exactas. Pensemos, por ejemplo, en la interpretación de la rebelión social y el radicalismo político en términos de «privación relativa». Según esta interpretación, «la insatisfacción con el sistema de privilegios y recompensas de una determinada sociedad no [es proporcional] al grado de desigualdad que afecta a sus miembros» [Runciman, 1966; 3], sino que depende más bien del sentido de privación que tengan dichos miembros, que a su vez deriva de los «grupos de referencia» y comunidades con los que los individuos se comparan, y las expectativas que tienen. Por ello, «la manifestación de descontento es poco frecuente en periodos de privaciones constantes, mientras que tiende a aumentar [...] cuando se entrevé una posibilidad de mejora [...] [y] es probable que las revoluciones se produzcan en periodos de bienestar creciente» [*ibidem*, 21]. Podría desarrollarse una teoría similar sobre la rebelión en el plano de la reflexión filosófica, histórica, social y psicológica, aplicándola a distintos contextos, problemáticas y hechos históricos. Pero en el momento en que queramos comprobar esta teoría mediante una investigación cuantitativa concreta, al definir el «diseño de la investigación», debemos definir ante todo la unidad de análisis. Ésta podría estar representada por el *episodio* de rebelión social, como en el estudio de Gurr y sus colaboradores [Gurr, 1972, 92-98], que recopilaron de manera sistemática datos sobre actos de violencia política y protesta social en el periodo 1961-65 en 38 países, en concreto, una serie de informaciones publicadas en la prensa. Por ejemplo, el número de participantes y víctimas, la duración de la agitación, el tipo de personas implicadas, los motivos del descontento, las modalidades de la protesta, las características del contexto social (tipo de régimen político, variables económicas, libertades civiles, etc.), llegando a registrar más de 1000 episodios. Todo ello con el objetivo de comprobar si los casos de conflicto y violencia estaban ligados a situaciones de cambio social que habían originado expectativas no satisfechas, y en general a situaciones de privación relativa más que de privación absoluta.

La unidad de análisis de una investigación orientada a comprobar la teoría mencionada también podría estar representada por un *núcleo territorial*. Aceptando una sugerencia de Tocqueville, para quien la revolución francesa se desarrolló en las zonas más ricas del país, puesto que «las partes de Francia que habían de ser el foco principal de esa revolución son precisamente aquellas donde el progreso es mayor» [Tocqueville, 1856], se podría realizar una investigación tomando como

unidad de análisis la región. Se trataría de encontrar un indicador de bienestar económico y fuerza del impulso revolucionario para las regiones de Francia en la época de la revolución, y relacionar ambas variables.

También se podría tomar como unidad de análisis al *individuo*. En 1962, Runciman entrevistó a unos 1400 ciudadanos de Inglaterra y Gales, utilizando preguntas abiertas y cerradas, para identificar los grupos de referencia de cada sujeto, la clase a la que consideraba que pertenecía, el grado de satisfacción con su posición social, las aspiraciones insatisfechas, es decir, las componentes de la privación relativa, junto con las orientaciones políticas, para ver si existía una correlación entre las dos variables.

Hemos dicho que la traducción empírica de un concepto (que es abstracto por definición) se hace mediante la asignación de este concepto a un objeto social concreto («unidad de análisis»), como si fuera una propiedad de dicho objeto. En una primera aproximación, podemos identificar los siguientes tipos de unidades de análisis: el *individuo,* el *conjunto de individuos,* el *grupo-organización-institución,* el *acontecimiento* y el *producto cultural.*

3.4.1. Los distintos tipos de unidad de análisis

La unidad de análisis más frecuente en las investigaciones sociales está representada por el individuo. Por otra parte, como escribe Galtung, «la sociología se define a menudo como la ciencia de la interacción social, de lo que se deriva que la unidad de los análisis sociológicos debería ser el actor social [...], y en la mayoría de los análisis sociológicos la elección recae sobre el actor social por excelencia, el individuo humano» [Galtung, 1967, 37].

También es bastante frecuente que la unidad de análisis esté representada por un *colectivo.* Este colectivo puede ser un conjunto de individuos o un grupo-organización-institución. El ejemplo más común de *conjunto de individuos* son las fuentes estadísticas oficiales, basadas en agrupaciones de individuos por territorios (municipios, provincias, comarcas, regiones...). En este caso las variables proceden de operaciones matemáticas realizadas sobre variables registradas de forma individual (por ejemplo, renta media, porcentaje de desempleados, etc.). Si hacemos una distinción, como a veces se hace, entre *unidad de análisis* y *unidad de registro,* en este caso la unidad de registro se sitúa en un nivel inferior respecto a la unidad de análisis, y está representada por el individuo.

En un *grupo-organización-institución* la mayoría de las variables está representada por lo que llamaremos «variables colectivas estructurales» o «globales», y la unidad de registro es el propio colectivo. Pensemos en un estudio sobre centros escolares, en el que las variables se refieren al tipo de gestión (privada/pública), la realización (o no) de programas experimentales, el tipo de perfil social al que se dirige la escuela, el número de clases, el número de alumnos, la presencia de mujeres en el profesorado, el índice de aprobados, etc. Aunque algunas variables (como las dos últimas citadas) hacen referencia a ámbitos individuales subyacentes, los datos se recogen de forma colectiva (por ejemplo, mediante una entrevista con el director y una consulta de los archivos del instituto), por lo que la unidad de análisis

y la unidad de registro coinciden.

En investigación social estas unidades de análisis son bastante frecuentes. Pensemos en *grupos* como familias, asociaciones, sectas religiosas, grupos étnicos, bandas juveniles, pandillas, y *organizaciones-instituciones* como los sindicatos, partidos, organizaciones laborales, unidades administrativas locales, instituciones públicas, hasta sociedades enteras y los propios países.

Una cuarta unidad de análisis es el *acontecimiento*. Ya hemos citado la investigación de Ted Gurr sobre los episodios de rebelión política. Otro acontecimiento podría ser la elección política: podemos citar como ejemplo la investigación de Bartolini y Mair [1990], que estudiaron las elecciones de las democracias europeas celebradas entre 1885 y 1995, en relación con la inestabilidad electoral, el voto de clase, la participación electoral, etc.; cada una de estas variables, obviamente, representa una propiedad de la unidad de análisis «acontecimiento electoral». Otros acontecimientos podrían ser huelgas, guerras, golpes de estado, procesos judiciales, ceremonias religiosas, campañas electorales, manifestaciones de protesta, etc.

El último tipo de unidad de análisis es lo que podríamos llamar *representación simbólica-producto cultural.* Pensemos en toda la corriente del análisis de contenidos, donde la unidad de análisis está representada casi siempre por mensajes de los medios de comunicación de masas, escritos, orales o audiovisuales: artículos de periódicos, textos literarios, programas electorales, discursos políticos, fotografías, programas de televisión, espectáculos de teatro, reportajes.

Finalmente queremos precisar que llamamos *observaciones* a los ejemplares de una determinada unidad de análisis incluidos en una investigación. La «unidad de análisis» es una definición abstracta, que designa el tipo de objeto social al que se refieren las propiedades (como el elector o el episodio de protesta). Esta unidad se localiza en el tiempo y en el espacio, definiendo «la población de referencia» de la investigación (por ejemplo, los electores españoles en las elecciones generales de 2000; los episodios de protesta política registrados en España entre 2002 y 2003). Esta población de referencia puede estudiarse de manera integral, por ejemplo, la investigación sobre los gobiernos que se han sucedido en España desde la llegada de la democracia en 1977 hasta el año 2004, o de manera parcial. A menudo se intentará extraer una muestra casual, otras veces se podrán usar distintos criterios de selección. Las observaciones son los ejemplares de esa unidad de análisis que se estudian, sobre los cuales se recopilan datos. Mientras que la unidad de análisis es singular y abstracta, las observaciones son múltiples y concretas, y constituyen el objeto específico de la investigación empírica. En definitiva, en una matriz de datos las observaciones son todos y cada uno de los números incluidos en las casillas.

3.5. VARIABLES

Como ya hemos dicho, una *variable* es *un concepto operacionalizado*. Más en concreto, consiste en la *propiedad operacionalizada* de un objeto, ya que para poder operacionalizar el concepto, debemos aplicarlo a un objeto, y de este modo se convierte en propiedad. Entre concepto, propiedad y variable existe la misma diferen-

cia que entre el peso (concepto), el peso de un objeto (propiedad) y el peso concreto de un objeto medido con una balanza (variable). La figura 3.2 ilustra este paralelismo. No existe una correspondencia biunívoca entre «concepto» y «variable», ya que un concepto se puede operacionalizar de distintas maneras. Se puede asociar, como propiedad, a diferentes unidades de análisis; por ejemplo, el concepto de poder puede asociarse a un individuo, un cargo empresarial, un cargo político, una institución, etc., y, como propiedad, puede dar lugar a diversas variables. Por ejemplo, la propiedad «nivel cultural» de un individuo puede definirse operacionalmente mediante *a*) su título universitario, *b*) el número de libros que ha leído o lee al año, *c*) las costumbres culturales cotidianas (lectura del periódico, cine, teatro), *d*) un test de cultura general, u otros.

Una variable puede adoptar distintos valores, que se corresponden con los diferentes estados de la propiedad correspondiente. El género, por ejemplo, es variable, puesto que puede adoptar los estados de macho o hembra. Nótese que, aunque una propiedad puede variar, también puede mantenerse invariable para el subconjunto concreto de los objetos estudiados. Por ejemplo, la nacionalidad es una propiedad que puede variar entre los individuos, pero si el estudio se realiza sobre la población española, será invariable (salvo para la población inmigrante). En este caso, esa propiedad, en su forma operacionalizada, ya no se llama variable, sino *constante*.

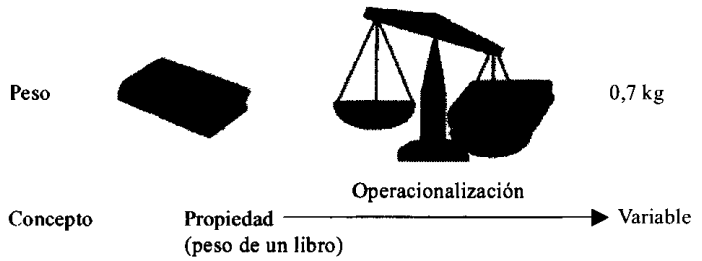


Figura 3.2. De los conceptos a las variables.

En los ejemplos citados hasta ahora, hemos hecho referencia a variables que varían entre los distintos objetos estudiados (por ejemplo, el género, que varía entre los individuos, o el poder, que varía entre los distintos cargos empresariales). Pero una variable puede variar de dos modos: *en el tiempo,* sobre un mismo caso, o *entre casos,* en un mismo momento. Pongamos dos ejemplos extraídos del ámbito de la medicina. Supongamos que queremos estudiar la reacción de un paciente a un determinado tratamiento médico, por ejemplo la relación entre la administración de quimioterapia y la variación en la concentración de glóbulos blancos. Para realizarlo haremos variar la dosis del medicamento, observando cómo varía la concentración de glóbulos blancos en la sangre. En este caso las dos variables *varían en el tiempo,* sobre el mismo sujeto. En cambio, para comprobar empíricamente la hipó-

tesis de que el cáncer de pulmón está relacionado con el consumo de tabaco, hay que proceder de forma distinta. Se podría comparar, por ejemplo, a un grupo de pacientes sanos con uno de enfermos de cáncer, estudiando si los distintos niveles de enfermedad/salud se corresponden con distintos grados de consumo de tabaco. En este caso las variables consideradas, enfermedad y consumo de tabaco, *varían entre los sujetos*. En el primer caso hablamos de *estudio longitudinal* (o *diacrónico*), y en el segundo caso, de estudio transversal (o *sincrónico*).

En las ciencias naturales el primer caso es muy frecuente. Además, el experimento, el principal procedimiento de las ciencias naturales para el estudio de las relaciones causales, se funda en variaciones de las variables en el tiempo («manipulo los valores de *x* y observo cómo varía *y*»). En las ciencias sociales, en cambio, es mucho más frecuente la segunda situación, donde las variables varían simultáneamente en todas las unidades de análisis. Ello se debe a que la mayoría de las variables sociales no se pueden manipular: por ejemplo, no podemos hacer variar la edad o la afiliación religiosa de un individuo y observar si con la variación de estas variables varía también su tendencia a abrazar valores postmaterialistas.

La «variable» tiene una importancia enorme en la investigación social empírica. Así como hemos dicho que los conceptos son los ladrillos de la teoría, podemos afirmar que las variables son el elemento fundamental del análisis empírico, el «vocabulario» de las ciencias sociales. Como escriben Lazarsfeld y Rosenberg (1955: 6) «la formación de las variables, el estudio de sus interrelaciones y el análisis de su variación en el tiempo constituyen la espina dorsal de toda investigación social». Y no es casual que Lazarsfeld definiera el lenguaje de la investigación social como «el lenguaje de las variables» (aunque sabemos que esta definición sólo es válida para la investigación de tipo cuantitativo).

Para visualizar mejor ese proceso crucial en la investigación empírica, el paso de los conceptos a las variables, usaremos un ejemplo expuesto previamente. En la investigación de Inglehart se planteaba el problema de operacionalizar el concepto de escala de valores materialista o postmaterialista. Ese concepto fue operacionalizado mediante una serie de preguntas. El autor partió de la jerarquía de las necesidades de Maslow, identificando dos grupos de necesidades materialistas (necesidad de subsistencia y necesidad de seguridad) y dos grupos de necesidades postmaterialistas (necesidad de pertenencia y aprecio, y necesidad de autorrealización estética e intelectual).

Para cada uno de los cuatro grupos de necesidades se formuló una afirmación en términos de objetivo de la política nacional (por ejemplo, mantener una alta tasa de crecimiento económico, garantizar una defensa militar fuerte, aumentar el poder de decisión de los individuos, etc.), y se pidió al entrevistado que marcara el objetivo que le parecía más importante de los cuatro. La pregunta se repetía tres veces, con tres series de objetivos diferentes. Al final las respuestas se combinaban, dando lugar a una puntuación que iba de 0 (materialismo máximo) a 5 (postmaterialismo máximo), teniendo en cuenta tanto las elecciones como las jerarquías.

Con esta explicación, el lector habrá percibido la gran arbitrariedad de las definiciones operativas. La forma de operacionalizar el concepto de escala de valores escogida por el autor es totalmente discutible. En este sentido hay que decir que no

existe una definición operacional correcta o errónea. La decisión sobre el modo de operacionalizar un concepto depende por completo del investigador, al cual sólo se le puede exigir que: *a*) explicite, y *b*) argumente sus decisiones. En todo caso, conviene señalar que siempre existirá un desfase entre variable y concepto. En este sentido puede afirmarse que una definición operativa nunca es perfectamente adecuada. Es un paso necesario pero raras veces suficiente para captar toda la complejidad de un concepto teórico.

La definición operativa implica por tanto una limitación y un empobrecimiento del concepto, pero el peligro que entraña no reside en esta ineludible reducción, sino en una posible «cosificación». Por ejemplo, el hecho de etiquetar como orientación postmaterialista un conjunto de respuestas a una serie de preguntas, y seguir usando el nombre del concepto cuando hablamos de relaciones que conciernen a su forma operacionalizada, la variable, puede hacernos olvidar que no estamos hablando de escala de valores en sentido estricto, sino de una interpretación y operacionalización de ella muy particular y totalmente arbitraria. Como ejemplo, recordemos el daño causado durante años por la identificación de la inteligencia con una operacionalización del concepto particular y con connotaciones culturales, el IQ (el test para determinar el cociente intelectual).

La definición operativa es, por tanto, arbitraria y subjetiva, pero paradójicamente, en ella se fundamenta el carácter científico y objetivo de la investigación social. Podemos decir que representa el *criterio* para establecer si una disciplina puede considerarse *científica:* la necesidad de definir operacionalmente las propiedades a estudiar es el discriminante más seguro entre la investigación científica y otros tipos de actividades, como por ejemplo la especulación filosófica. Y conviene que añadir que, si bien es cierto que la definición operativa es arbitraria y subjetiva, a medida que una ciencia avanza y madura y sus técnicas se perfeccionan, su arbitrariedad disminuye, puesto que dentro de la comunidad científica surgen convergencias que permiten construir definiciones convencionales de uso más generalizado.

La definición operacional representa también el criterio de objetividad de la investigación científica, ya que establece las directrices para que otros investigadores puedan repetir la misma investigación. De esta forma se reduce el grado de subjetividad de las afirmaciones del investigador. Ya no son opiniones, sino afirmaciones dotadas de apoyo empírico. Por ejemplo, afirmar que Fulano es un individuo autoritario es una opinión, si no va acompañada de hechos que la fundamenten. Pero afirmar que es autoritario porque ha contestado de un modo concreto a una determinada serie de preguntas, es algo que se puede verificar.

3.6. VARIABLES NOMINALES, ORDINALES Y CARDINALES

Una clasificación muy importante de las variables concierne al tipo de operaciones que pueden efectuarse con ellas. Las *características lógico-matemáticas* de una variable hacen referencia a las operaciones lógicas o matemáticas que se pueden

realizar con sus valores.

Se trata de una clasificación de gran importancia para las variables, puesto que determina las operaciones estadísticas que se les puede aplicar. Y dado que el investigador recurre a la variable para poder tratar una propiedad con instrumentos matemáticos y estadísticos, puede entenderse lo importante que es esta característica, pues condiciona todas las operaciones para las que ha sido construida dicha variable. Hasta las operaciones estadísticas más comunes, como el cálculo de la media o un índice de asociación, dependen de las características lógico-matemáticas de la variable (por ejemplo, se puede calcular la edad media de una muestra de individuos, pero no la nacionalidad media).

De acuerdo con ello, clasificaremos las variables en tres tipos (variables nominales, ordinales y cardinales, ver la tabla 3.1). Esta distinción, que toma como referencia el *análisis de los datos,* está muy relacionada con la naturaleza de las operaciones empíricas realizadas para el *registro de datos,* es decir, para operacionalizar los estados de la propiedad al transformarla en variable. Por tanto, volveremos también sobre los procedimientos de operacionalización, que, como hemos visto, son cuatro (clasificación, ordenación, cómputo, medición)[2].

3.6.1. Variables nominales

Cuando la propiedad que vamos a registrar adopta estados discretos que no se pueden ordenar hablamos de una variable nominal. Los estados «discretos» implican que la propiedad puede adoptar solamente una serie de estados finitos; en el lenguaje común podríamos decir que pasa de un estado a otro con un salto, y no puede adoptar estados intermedios (la religión puede ser católica o musulmana, pero no puede haber una vía intermedia entre ambas; el género puede adoptar sólo los estados masculino o femenino).

Los estados «no ordenables» son aquellos entre los que no es posible establecer un orden o jerarquía. Podemos decir que una persona es de nacionalidad española, sueca o china, pero no podemos ordenar tales estados en una secuencia jerárquica, igual que no podemos establecer un orden entre los estados masculino y femenino, o entre los estados de católico, protestante, musulmán, judío, ateo, etc. Las únicas relaciones que podemos establecer entre las modalidades de una variable nominal son las relaciones de «igual» y «desigual» (en lo que concierne a la religión, un católico es igual que otro católico y distinto de un protestante).

Llamamos *modalidades* a los estados de la variable operacionalizados y *valores* a los símbolos asignados a las modalidades (éstos suelen ser números, aunque no siempre es así). La operación que permite pasar de la propiedad a la variable es, en este caso, la *clasificación*. Como hemos indicado, estamos en los niveles inferiores de formalización matemática, y, en efecto, la clasificación es la operación primera y la más sencilla que se puede realizar en cualquier ciencia.

---

[2] En nuestra exposición seguiremos la propuesta metodológica planteada por Alberto Marradi [en particular, Marradi, 1980 y 1993].

Las categorías en las que clasificamos los estados de la propiedad deben cumplir dos requisitos: *a*) *exhaustividad*, lo que exige que cada caso que examinamos se debe poder clasificar al menos en una de las categorías, y *b*) *exclusividad mutua,* lo que implica que un caso no se puede clasificar en más de una categoría. De este segundo requisito se deriva también la *unicidad del criterio de clasificación,* lo que significa, por ejemplo, que no podemos clasificar la nacionalidad utilizando categorías como italiana, francesa, china, protestante, nórdica, inglesa, europea, asiática, etc.

A las modalidades se les asocia un símbolo que, como hemos dicho, llamamos valor, y que, sin embargo, no tiene ningún significado, aparte de servir para identificar a la categoría. Suele ser un número, pero éste carece de significado numérico. Para la variable «religión» es indiferente asignarle el valor 1 a católico, el 2 a protestante, el 3 a judío, el 4 a musulmán, el 5 a otras religiones y el 6 a ateo. El orden podría ser cualquier otro, y también podríamos utilizar otros números (el único requisito es que sean diferentes entre sí), o letras, u otras etiquetas de cualquier tipo. La operación de subdivisión de una propiedad en categorías no ordenadas consiste, de hecho, en la simple asociación de un nombre para cada categoría: por este motivo este tipo de variable se denominada «nominal».

Existe un tipo de variables nominales que sólo tiene dos modalidades: hombre y mujer, empleado y desempleado, casado y no-casado, etc. Estas variables se llaman *dicotómicas* y tienen la peculiaridad de que se les puede aplicar operaciones estadísticas que no suelen ser aplicables a las variables nominales, siempre y cuando estén codificadas con los valores 0 y 1.

Tabla 3.1. Tipos de variables

Estados de la propiedad	Procedimiento de operacionalización	Tipo de variable	Características de los valores	Operaciones posibles
Discretos no ordenables	Clasificación	Nominal	Nombres	$= \neq$
Discretos ordenables	Ordenación	Ordinal	Números con características exclusivamente ordinales	$= \neq$ $> <$
Discretos enumerables	Cómputo	Cardinal	Números con características cardinales	$= \neq$ $> <$ $+ - \times :$
Continuos	Medición	Cardinal	Números con características cardinales	$= \neq$ $> <$ $+ - \times :$

3.6.2. Variables ordinales

En el caso de las variables ordinales, la propiedad a registrar adopta *estados discretos ordenables*. Como ejemplos podemos citar el nivel de estudios (enseñanza primaria, secundaria, diplomatura, licenciatura), la jerarquía militar (teniente, capitán, coronel, general), o la clase social (nobleza, burguesía y proletariado). También podría servir como ejemplo un cuestionario con varias respuestas a elegir en forma de escala (mucho, bastante, poco, en absoluto). Lo que distingue a este tipo de variables del anterior es la ordenación, que permite establecer relaciones de igualdad y desigualdad entre las distintas modalidades, y relaciones de orden, «mayor que» y «menor que». Conviene señalar que en las variables ordinales no se conoce la distancia existente entre las distintas modalidades. Sabemos que, en la jerarquía laboral, un obrero industrial ocupa un nivel superior al de un trabajador agrícola temporal, pero no conocemos la distancia entre ambos.

El procedimiento de operacionalización de las propiedades es, en este caso, la «asignación de modalidades ordenadas», que llamamos *ordenación*. Por lo tanto, la asignación de los valores a cada una de las modalidades ya no podrá ser casual, sino que tendrá que aplicar un criterio que preserve el orden entre los estados. Por ello casi siempre se utiliza la serie de los números naturales, que gozan de las propiedades ordinales de los números. Si atribuimos el valor 1 a la ausencia de título académico, 2 al título de enseñanza primaria, 3 al título de enseñanza secundaria, 4 a la diplomatura y 5 a la licenciatura, estos números deben considerarse en el contexto de la secuencia que representan, pero sin atribuirles el valor del número correspondiente. Por ejemplo, no podemos decir que la distancia entre ausencia de título y título de enseñanza primaria sea la misma que la distancia entre diplomatura y licenciatura (aunque en ambos casos haya una diferencia de una unidad en la secuencia numérica propuesta), o bien que un diplomado universitario (al que hemos asignado el valor 4) tenga el doble de estudios que una persona con un título de enseñanza primaria (valor 2).

Dado que los números tienen un significado puramente ordinal, al asignarlos a las distintas modalidades, sólo indican una secuencia y nada más. De ello se deriva que, por ejemplo, la secuencia propuesta para el nivel de estudios (1, 2, 3, 4, 5) podría sustituirse por cualquier otra secuencia numérica creciente, como 12, 25, 32, 43, 55. Sin embargo, es conveniente utilizar el criterio de la serie de los números naturales en secuencia simple (1, 2, 3...), o un criterio basado en una estimación, aunque sea aproximada y subjetiva, de las distancias entre las categorías. Ya que la secuencia simple es tan arbitraria como cualquier otra serie ordenada de números, es razonable utilizar una secuencia que refleje las distancias entre las distintas categorías [3].

[3] En el caso del nivel de estudios, se podría afirmar que las personas sin formación se encuentran una situación de marginalidad que los separa del resto, y que su distancia de quienes tienen estudios de primaria es mayor que la distancia entre éstos y quienes tienen estudios de secundaria. También podría sostenerse que la distancia entre diplomados y licenciados universitarios es pequeña. Por ello, en este caso podría proponerse la secuencia numérica 1, 4, 6, 8, 9.

Las variables pueden ser ordinales por dos motivos: porque derivan de propiedades que en un principio estaban constituidas por estados discretos, como en los ejemplos citados (nivel de estudios, clase social), o porque derivan de propiedades continuas que han sido registradas en una secuencia ordinal a falta de instrumentos de medición adecuados. Éste es el caso de los citados cuestionarios con varias respuestas a elegir en forma de escala (mucho, bastante, poco, en absoluto).

3.6.3. Variables cardinales

Son variables para las cuales los números asignados a las modalidades (los «valores» de la variable) tienen significado numérico pleno, en el sentido de que no sólo presentan las propiedades ordinales de los números, sino también las cardinales. Ejemplos de variables de este tipo son la edad, la renta o el número de hijos. Dado el carácter cardinal de los valores, entre las modalidades de una variable de este tipo podrán establecerse relaciones de igualdad y desigualdad (la edad de 20 años es distinta a la edad de 22), relaciones de orden (20 años es menor que 22), y también podrán realizarse operaciones de suma y resta.

En este tipo de variables se conoce la *distancia* existente entre dos valores (existe la misma diferencia de edad entre 20 y 22 años que entre 48 y 50 años), y entre estas distancias también pueden aplicarse las operaciones de multiplicación y división (la diferencia de edad entre 16 y 20 años es el doble que entre 20 y 22). Es posible determinar estas distancias o intervalos, puesto que existe una unidad de referencia (que, como veremos, puede ser una unidad de medida o una unidad de cálculo). Esto permite aplicar a las distancias las cuatro operaciones aritméticas, y por tanto permite realizar operaciones matemáticas y estadísticas más avanzadas sobre las variables [4].

Las variables cardinales se pueden obtener con dos procesos de operacionalización de la propiedad: la medición y el cómputo. La *medición* se produce cuando

---

[4] La medición puede producir variables que incluyan el cero convencional o el cero absoluto (un cero «físico» que significa ausencia total de la propiedad). Pensemos en la temperatura en grados Celsius y en grados Kelvin. El cómputo produce variables con cero absoluto (Stevens ha denominado escalas «a intervalos» a las variables con cero convencional, y escalas «de razón» a las variables con cero absoluto). En las variables con cero convencional, es posible sumar y restar las modalidades de las variables (se puede calcular la diferencia entre 20 y 22 grados o entre 40 y 44). A las *diferencias* calculadas entre los valores se pueden aplicar las cuatro operaciones aritméticas. Por ejemplo, puede decirse que la diferencia entre 10 y 30 grados es el doble de la diferencia entre 0 y 10 grados. Y esto puede hacerse independientemente de la unidad de medida seleccionada: puesto que 0 ° Celsius se corresponde con 32° Fahrenheit, 10° C = 50° F, 20° C = 68° F, 30° C = 86° F, puede verse que, igual que con grados Celsius, la diferencia correspondiente en grados Fahrenheit 86° F – 50° F = 36° F también es el doble de la diferencia 50° F – 32° F = 18° F. El hecho de que puedan realizarse las cuatro operaciones aritméticas con las diferencias entre los valores (intervalos), hace posible aplicar a estas variables casi todos los cálculos estadísticos. En el caso de variables con cero absoluto, las cuatro operaciones aritméticas son aplicables a los valores de la variable, lo que amplía las posibilidades de utilización de los cálculos estadísticos. Para simplificar, en este manual no distinguiremos entre variables con cero convencional y cero absoluto, dado que los cálculos estadísticos más comunes se aplican indistintamente a los dos tipos de variables.

concurren las dos condiciones siguientes: *a*) la propiedad a medir es *continua*, es decir puede adoptar estados intermedios infinitos en cualquier intervalo entre dos estados, y *b*) disponemos de una *unidad de medida* preestablecida que nos permite comparar la magnitud a medir con una de referencia.

El ejemplo de medición más inmediato es la longitud medida con la unidad de medida convencional del metro, pero naturalmente podemos imaginar muchos otros. El proceso de medición conlleva el redondeo del número real correspondiente al estado medido en un número de cifras compatible con la aproximación que hemos establecido para nuestra medición (por ejemplo, una estatura se podrá medir en centímetros, como 167, donde el último dígito se obtiene por redondeo).

Hablamos de *cómputo* cuando: *a*) la propiedad a registrar es *discreta*, es decir adopta estados finitos, no fraccionables, y *b*) existe una *unidad de cálculo*, es decir una unidad elemental que está contenida un número finito de veces en la propiedad del objeto. En este caso, la operacionalización consiste en «contar» cuántas unidades de cálculo incluye la cantidad de propiedad que posee el objeto. Así, contamos el número de hijos de una persona, el número de habitaciones de un apartamento, el número de empleados de una empresa, el número de veces que una persona va a la iglesia al mes, el número de periódicos que lee a la semana. La unidad de cálculo es «natural», a diferencia de la unidad de medida, que es «convencional». En los ejemplos mencionados, la unidad de cálculo es, según el caso, el hijo, la habitación, el empleado, la visita a la iglesia, el periódico. Al tratarse de propiedades discretas (la unidad de cálculo es indivisible), en el cómputo no puede haber decimales (una persona puede tener 2 hijos, pero no 2,3), ni pueden realizarse redondeos. Por otra parte, al realizarse ciertas operaciones matemáticas, la variable podrá dar lugar a decimales: por ejemplo, podrá obtenerse un número medio de 0,7 habitaciones por persona, o una tasa de fecundidad de un país igual a 1,2 hijos.

En física se distingue entre variables *fundamentales* y variables *derivadas*. La longitud, la masa, el tiempo y la temperatura serían variables fundamentales; la densidad, y la velocidad serían variables derivadas. Las segundas no son otra cosa que operaciones matemáticas con las variables fundamentales. En las ciencias sociales también hay muchas variables cardinales que se derivan de operaciones realizadas con otras variables cardinales. En una familia, los metros cuadrados de vivienda disponibles por persona se derivan de la división entre dos variables cardinales: una que se refiere a una propiedad continua (la superficie de la habitación), y otra a una propiedad discreta (el número de componentes de la familia). Las variables derivadas son especialmente importantes cuando la unidad de análisis está representada por una agrupación territorial de individuos. El porcentaje de votos obtenido por un partido en relación con el total de votos válidos en una determinada provincia, el porcentaje de desempleados en relación con el total de la población activa, el porcentaje de licenciados, etc. son propiedades derivadas de relaciones entre variables cardinales contadas sobre estados discretos.

Nótese que las características de los tres tipos de variables son acumulables. Así, entre los valores de las variables nominales sólo se pueden establecer relaciones de igualdad y diferencia; entre los valores de las variables ordinales, *además* de las relaciones de igualdad y diferencia, se pueden establecer relaciones de ordena-

ción; y entre los valores de las variables cardinales, *además* de las relaciones mencionadas, se pueden establecer relaciones derivadas del conocimiento de la distancia entre los valores. Esto implica que es posible pasar de un nivel de variables superior a otro inferior. Por ejemplo, los valores de la variable cardinal «edad» se pueden recodificar en las tres categorías ordenadas de jóvenes, adultos y ancianos (variable ordinal); también podemos sumar los distintos grados de «religiosidad», ordenados en una escala con diferentes modalidades, desde «va a la iglesia todos los días» hasta «no va a la iglesia nunca», en una variable nominal dicotómica que nos permita distinguir entre «practicantes» y «no practicantes».

Dentro de las variables cardinales podemos distinguir un subgrupo representado por las *variables cuasicardinales*. Hemos hablado de variables cardinales derivadas de propiedades continuas, y variables cardinales derivadas de propiedades discretas. Las del primer tipo son poco frecuentes en las ciencias sociales. Ya hemos mencionado la edad, podríamos recordar otras variables basadas también en el tiempo (por ejemplo, el tiempo empleado para realizar una actividad cualquiera) y podríamos citar también la renta, la distancia, pero pocas más. Además, todas estas variables derivan de propiedades típicas de las ciencias naturales. No obstante, las propiedades más características de las ciencias sociales, desde la religiosidad hasta el sistema político, pasando por el autoritarismo, la cohesión social, las escalas de valores, etc. pueden considerarse propiedades continuas, que varían de forma gradual entre los individuos. Sin embargo, esas propiedades no logran pasar de la condición de propiedad continua a la de variable cardinal por un defecto en la fase de operacionalización, en concreto, por la dificultad de aplicar una unidad de medida a las actitudes humanas.

Hasta la fecha ha habido numerosos intentos de superar ese límite, en particular, la «técnica de las escalas» *(scaling),* que desde los años veinte se propuso medir opiniones, actitudes y valores, y más en general propiedades continuas relacionadas con la estructura psicológica y los valores del individuo. Las «escalas de autoposicionamiento», como el «termómetro de los sentimientos», o las «escalas de posicionamiento izquierda-derecha» para la orientación política, también intentan superar dicho límite. El objetivo es conseguir algo cercano a una «medición», es decir, variables que permitan conocer la distancia entre dos valores (lo que equivale a decir que se tiene una unidad de medida y, por tanto, se tiene una variable cardinal). Parece totalmente lícito tratar las variables obtenidas con estas técnicas utilizando los instrumentos matemáticos-estadísticos previstos para las variables cardinales. Para subrayar esta aceptación, y al mismo tiempo «la diferencia con respecto a las variables cardinales», aceptamos la sugerencia de Marradi, que propone llamarlas variables cuasicardinales [Marradi, 1993, 98].

En cuanto a las técnicas de cálculo estadístico aplicables a las variables, se pueden dividir en dos grandes grupos: las técnicas para las variables cardinales y las técnicas para las variables nominales. En muy pocos casos disponemos de técnicas concebidas específicamente para las variables ordinales[5]. En rigor, éstas

[5] Los modelos de probabilidad discreta son una excepción, ya que están diseñados precisamente para la estimación de variables dependientes cualitativas. Sobre esta técnica, véase Long, 1997.

deberían tratarse exclusivamente con las técnicas de cálculo estadístico utilizadas para las variables nominales, sobre todo si la variable ordinal deriva de una propiedad con estados discretos ordenados. Hay que decir, sin embargo, que esta cuestión ha suscitado una amplia polémica entre los expertos en metodología de las ciencias sociales, y a menudo se analizan las variables ordinales con técnicas estadísticas que han sido desarrolladas para variables cardinales.

3.7. CONCEPTOS, INDICADORES E ÍNDICES

En este apartado vamos a profundizar sobre el paso de los conceptos a las variables. Ya hemos dicho que en el proceso de traducción empírica, un concepto se conecta con un objeto (unidad de análisis), se convierte en propiedad, y se operacionaliza, es decir, se registra en forma de variable. Por ejemplo, el concepto de práctica religiosa se define como propiedad de los sujetos humanos, y se puede operacionalizar mediante el número de veces que una persona va a la iglesia al mes (variable). Pero si el concepto que queremos operacionalizar es un concepto más general, como la religiosidad, tendremos más dificultades para darle una definición operativa. La práctica religiosa puede ser un aspecto del concepto, pero sin duda no es el único. En otras palabras, hay conceptos difícilmente transformables en propiedades de las unidades de análisis.

Podemos clasificar los conceptos dentro de una serie continua que viene dada por el distinto grado de generalidad-especificidad («escala de generalidad»). Por ejemplo, con respecto a la religiosidad, podemos imaginar cinco formulaciones conceptuales en orden de generalidad decreciente (o especificidad creciente): 1) creer en la existencia de una divinidad; 2) creer en el dios cristiano; 3) pertenecer a la iglesia católica; 4) actuar conforme a lo que dicta la iglesia, y 5) ir a misa todos los domingos. Esta escala de generalidad está correlacionada con una «escala de abstracción» (cuyos polos son abstracto-concreto), que contempla la medida en que un concepto es observable. Esta correlación se deriva del hecho de que un concepto específico también es, en general, más fácilmente observable que un concepto general. Por ejemplo, la fe en dios no es observable, pero la práctica religiosa sí.

Ahora bien, la mayoría de los conceptos sociológicos tienen un alto grado de generalidad. Por ejemplo, conceptos como alienación, socialización, poder, o conflicto, son representaciones teóricas alejadas de la experiencia. Sin embargo, dado que nuestro objetivo es comprobar empíricamente la teoría, debemos definir empíricamente esos conceptos. Pero, ¿cómo podemos formular una definición operacional de un concepto tan abstracto como el de «alienación»?

Aquí es donde entran en juego los *indicadores*. Se trata de conceptos más sencillos, «específicos», traducibles a la realidad, que están ligados a los conceptos «generales», por lo que se define como una *relación de indicación,* o representación semántica (o sea, de significado). Se trata de pasar, en la escala de generalidad que acabamos de definir, de conceptos generales a conceptos específicos, ligados entre sí por afinidad de significado. Volviendo al ejemplo anterior, si no se puede

observar empíricamente la religiosidad, se podrá, sin embargo, observar una forma «específica» de religiosidad, que es la práctica religiosa.

Es evidente que la relación entre concepto e indicador es sólo parcial. Por una parte, un concepto (general) admite más de un indicador (específico), y por otra, un indicador puede servir para concretar más de un concepto.

Analicemos la primera parte de esta afirmación. Los indicadores, precisamente por el hecho de ser específicos, pueden captar sólo un aspecto de la complejidad de un concepto general. Por ejemplo, la práctica religiosa podrá ser un indicador del componente ritual de la religiosidad (junto a otros indicadores como la oración, lectura de textos sagrados, etc.). Pero el concepto de religiosidad incluye otros componentes. Glock menciona, además de los ritos, las experiencias religiosas, la ideología y el comportamiento inspirado por las convicciones religiosas [Glock, 1959]. Como ejemplos de indicadores de este último componente podemos citar los actos de caridad o una moral sexual acorde con las convicciones religiosas.

Por ello es *necesario recurrir a más de un indicador* para registrar operacionalmente un concepto. Lazarsfeld sugiere la expresión «universo de los indicadores», acuñada por Guttman [Lazarsfeld, 1967, 190]. El mismo Lazarsfeld propone llamar «dimensiones» a los distintos aspectos con los que se especifica un concepto, para llegar a la identificación de los indicadores (en el ejemplo anterior, los cuatro aspectos mencionados, ritual, experiencias, ideología y conducta, serían cuatro dimensiones distintas del mismo concepto: la religiosidad).

Hemos dicho que un indicador puede servir para concretar varios conceptos, y éstos pueden tener contenidos semánticos radicalmente distintos. Por ejemplo, en las sociedades dominadas por las instituciones eclesiásticas, la participación en los ritos religiosos puede ser un indicador de conformismo social más que de religiosidad. Según Marradi [1980, 36] un indicador se compone de dos partes: la *parte de indicación* y la *parte ajena*. La primera es la parte de contenido semántico que tienen en común el indicador y el concepto que especifica, y la segunda es la parte del indicador ajena al mismo. Obviamente, siempre daremos prioridad a los indicadores con la máxima parte de indicación y la mínima parte ajena.

Otro aspecto relevante es la arbitrariedad de la elección de un indicador. Hemos dicho que el indicador representa sólo una parte de un concepto y que puede haber varios indicadores para un mismo concepto. El investigador elegirá entre esta variedad de indicadores según su propio criterio, y la única obligación que tiene con la comunidad científica es argumentar y justificar su elección, pero no «demostrar» que es correcta, ya que eso es imposible.

Resumiendo y esquematizando, podemos decir que la observación de un concepto que no es directamente observable pasa por cuatro fases [6]: 1) la división del concepto en dimensiones; 2) la elección de los indicadores; 3) la operacionalización, y 4) la formación de los índices.

La primera fase consiste en una reflexión teórica en la que se analizan las com-

[6] He utilizado parte del esquema de Lazarsfeld [1965], unificando sus dos primeras fases y añadiendo la fase de operacionalización.

ponentes principales de significado del concepto (las «dimensiones»). En el ejemplo de la religiosidad, en esta fase se identificarían las cuatro componentes mencionadas (rituales, experiencias, etc.). En el caso de la inteligencia, se identificarían las distintas habilidades y capacidades mentales. Lazarsfeld [1965, 43] cita el concepto de «gestión de una fábrica» mencionando como dimensiones (positivas) «falta de desacuerdos, buena comunicación entre los distintos niveles jerárquicos, flexibilidad de la autoridad, gestión racional, importancia relativa de los cargos, etc.».

La segunda fase consiste en la identificación de los indicadores correspondientes a cada dimensión. Nótese que, aunque aún estamos en una fase conceptual (los indicadores todavía son conceptos), ya se empieza a plantear el problema de la observación. Por tanto, se elegirán como indicadores conceptos específicos que sean observables, «cercanos a la experiencia», para emplear la expresión de Geertz [1977]. Dado el carácter de parcialidad del indicador, el investigador suele identificar más de uno para la misma dimensión conceptual. Por ejemplo, para la dimensión ritual de la religiosidad hemos mencionado la participación en los ritos colectivos, la oración individual y la lectura de textos sagrados.

En la tercera fase se procede a la operacionalización de los indicadores, es decir, su transformación en variables. Podemos operacionalizar la práctica religiosa con el número de veces que una persona va a misa al mes, o cuantificar el tiempo dedicado a la oración y la lectura de textos sagrados. Para las capacidades intelectuales podríamos elaborar un test que las operacionalice en un cómputo.

La cuarta y última fase es la de la construcción de los *índices*. Ésta operación no se había mencionado anteriormente porque interviene sólo en conceptos complejos que requieren más de un indicador, y no se plantea para conceptos sencillos traducibles en una sola variable (como la edad o el nivel de estudios). Cuando un concepto se descompone en dimensiones y se registra mediante varios indicadores, suele surgir la necesidad de sintetizar en un solo índice las múltiples variables. Así, después de registrar la religiosidad mediante variables que operacionalizan los rituales, las experiencias, la ideología y el comportamiento religiosos, podemos desear recomponer la unidad original en lo que podríamos llamar *índice* (global) de religiosidad. Podrá ser un índice unidimensional que «ordena» la religiosidad según una puntuación de las características ordinales, o la «mide», dando una puntuación cuasicardinal; también puede ser un índice tipológico que, por ejemplo, «clasifica» la religiosidad en tipos diferentes entre sí y no graduables, como devotos, practicantes, conformistas, y no religiosos.

El proceso que acabamos de describir está representado en la figura 3.3 [7]. Ya hemos dado un ejemplo con el caso de la religiosidad. La traducción en términos operacionales del concepto de escala de valores materialista y postmaterialista también podría servir para ilustrar el proceso. En ese caso el concepto se dividió en las cuatro dimensiones de los valores ligados a las necesidades de subsistencia, seguridad, pertenencia-aprecio y autorrealización. De estas dimensiones se extrajeron los

[7] Para simplificar no hemos incluido en el esquema la distinción entre «concepto» y «propiedad», ya que ambas nociones prácticamente se solapan.

indicadores relativos a opiniones sobre «combatir la criminalidad», «libertad de expresión», «defensa del medioambiente», etc. Inglehart consideraba, por ejemplo, que el hecho de dar prioridad a la defensa del medioambiente respecto a la lucha contra la criminalidad era un indicador de orientación postmaterialista. Estos indicadores se operacionalizaron mediante preguntas en un cuestionario, y, por último, el autor sintetizó las respuestas en un índice de materialismo-postmaterialismo (con una puntuación que iba de 0 a 5).

El proceso de traducción de un concepto en operaciones empíricas, que pasa por las fases de separación de dimensiones, definición de los indicadores y operacionalización, no siempre se realiza en la secuencia indicada aquí. En particular, cuando el investigador no crea directamente la evidencia y utiliza datos ya existentes, se encuentra ante la imposibilidad de planificar la definición de los indicadores según su análisis conceptual, puesto que sólo puede servirse de la evidencia ya existente, recopilada anteriormente por otros investigadores.

El problema es especialmente relevante cuando se trabaja con fuentes estadísticas, en las que la unidad de análisis está formada por núcleos territoriales (municipios, provincias, países). Por ejemplo, en una investigación sobre religiosidad realizada en Italia usando estadísticas provinciales, el concepto se operacionalizó mediante los siguientes indicadores: importancia del matrimonio civil y difusión del semanario católico *Familia Cristiana* (con variables como «tasa de matrimonios civiles sobre el total de matrimonios»; «número de copias de la revista difundidas por cada 1000 habitantes») [Cartocci, 1993]. La combinación de estos dos indicadores permite al autor obtener un índice tipológico de religiosidad, clasifi-

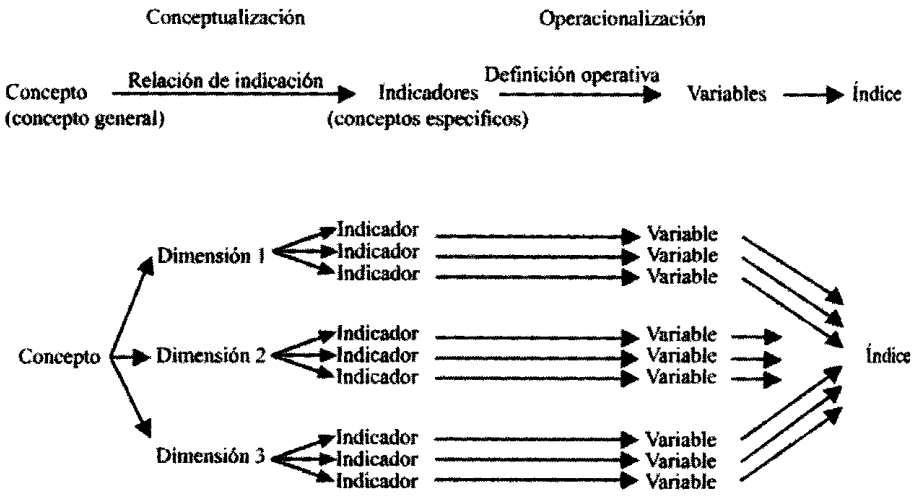


Figura 3.3. Representación esquemática del proceso de traducción empírica de un concepto complejo.

cando las provincias en cuatro tipo (practicantes, conformistas, en transición y secularizadas). Además de los dos utilizados, el autor indica otros indicadores de religiosidad, tomando siempre como unidad de análisis la provincia italiana: tasa de natalidad, número de hijos nacidos fuera del matrimonio, número de divorcios, separaciones y abortos, asignación a la Iglesia del ocho por mil de la contribución de la renta. Todo ello derivado no de una reflexión autónoma sobre las dimensiones del concepto de religiosidad, sino de la limitación que imponían los datos.

Hay que añadir que en estas situaciones es especialmente elevado el riesgo de indicadores con una «parte ajena» importante, es decir, indicadores que están ligados semánticamente a conceptos totalmente ajenos al que se pretende registrar. La asignación de la cuota del ocho por mil de la renta a la Iglesia puede indicar desconfianza en las instituciones estatales, más que un alto grado de religiosidad [*ibidem*, 135]. Por tanto hay que tener mucho cuidado a la hora de elegir indicadores para operacionalizar conceptos complejos cuando, por el motivo que sea, esta elección no puede orientar la recopilación de datos. Lamentablemente, ante conceptos complejos para los que no disponen de datos adecuados, muchos investigadores intentan arreglarse con lo que tienen, y fuerzan la correspondencia semántica entre indicador y concepto, y puede llegar incluso a utilizar indicadores con una parte de indicación demasiado baja. Todo ello afecta a la calidad de los resultados de sus análisis y, lo que es más importante, a sus conclusiones.

3.8. ERROR TOTAL

En la reflexión sobre el proceso que permite al científico social pasar del mundo de los conceptos y teorías al mundo de la experiencia y los sentidos, tiene un papel fundamental lo que los psicométricos llaman «error de medición», que nosotros llamaremos «error total», ampliando su sentido. Este error representa el desfase entre el concepto (teórico) y la variable (empírica).

Este error tiene dos componentes: el error sistemático y el error aleatorio. Por lo tanto, puede decirse que el *valor observado,* es decir, el valor de la variable empírica tal y como lo registra el investigador, es el resultado de la suma de tres partes: el *valor verdadero* no observado ni observable del concepto que la variable pretende registrar, y los dos componentes del error. Expresado formalmente[8]:

$$\underset{\text{(variable)}}{\text{Valor observado}} = \underset{\text{(concepto)}}{\text{Valor verdadero}} + \text{Error sistemático} + \text{Error aleatorio}$$

Lo que podemos expresar también de la siguiente forma equivalente:

[8] Llamamos «valor verdadero» y «valor observado» a lo que en la terminología psicométrica en inglés se denomina *true score* y *observed score.* La razón por la que evitamos el término «puntuación» es que, si bien en el ámbito de los tests psicométricos se utilizan efectivamente las puntuaciones, éste no suele ser el caso en ciencias sociales, donde a menudo la variable es nominal u ordinal.

Error = Valor observado – Valor verdadero = Error sistemático + Error aleatorio

El *error sistemático*, o «distorsión», es un error *constante*, porque aparece en todas las investigaciones, ya sean estudios sobre individuos diferentes o réplicas de investigaciones sobre los mismos individuos. Su valor medio sobre el total de los casos observados no es igual a cero, sino que toma un valor positivo o negativo, en el sentido de que el «valor observado» tiende sistemáticamente a ser inferior o superior al «valor verdadero». Por ejemplo, si estudiamos la participación electoral de una muestra de ciudadanos mediante una encuesta en la que les preguntamos si votaron en las últimas elecciones, sabemos que existe una tendencia generalizada a no declarar la abstención, por lo que en todas las encuestas se obtendrá una tasa media de participación electoral «observada» que será sistemáticamente superior a la «verdadera».

El *error aleatorio* es un error *variable*, porque varía para cada registro. Varía en hipotéticas réplicas del mismo registro en el mismo individuo, y varía al pasar de una muestra de individuos a otra. Se trata de oscilaciones que tienden a un valor medio (valor esperado) igual a cero, sobre todas las réplicas posibles del registro y sobre todas las muestras.

Podemos decir que si el error sistemático es la parte de error común a todas las aplicaciones de una investigación, el error aleatorio es la parte de error específica de cada una de las investigaciones.

¿En qué fases del proceso de transformación de los conceptos en variables surgen estos errores? Hemos visto que dicha transformación tiene dos fases (ver la figura 3.4): una fase *teórica*, que podríamos llamar de indicación, que consiste en

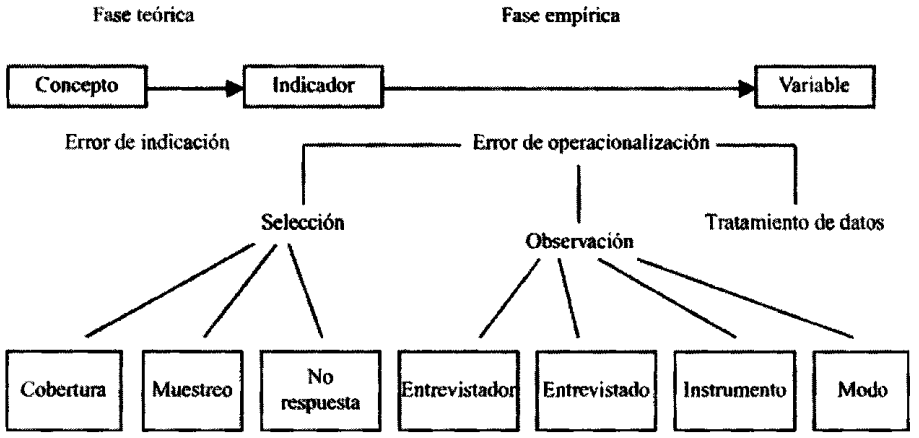


Figura 3.4. Tipos de error y su posición en el proceso de transformación de un concepto en variable.

la elección de los indicadores, y una fase *empírica*, que hemos denominado de *operacionalización* y que consiste en la obtención de los propios indicadores. Se pueden cometer errores en las dos fases.

• El error en la fase de *indicación*, es decir, en la elección de los indicadores adecuados para medir un determinado concepto, es un error de tipo *sistemático*. Por ejemplo, tomar como indicador de participación política la afiliación a un sindicato puede ser una elección afectada de error sistemático; de hecho, la afiliación a un sindicato puede expresar un cierto conformismo social o una preocupación por proteger los intereses propios, o un comportamiento que se repite cada año por inercia, más que ser indicio de una participación política real. En este caso el indicador no «cubre» el concepto de manera adecuada o completa y surge un *defecto en la relación de indicación*.

Recordemos lo dicho sobre la «parte de indicación » y la «parte ajena» de un indicador con relación al concepto al que representa: la primera es la parte de superposición semántica entre concepto e indicador. Un indicador será más *válido* cuanto mayor sea su parte de indicación y menor su parte ajena. Si en el indicador representado por la afiliación a un sindicato la parte ajena supera a la parte de indicación en relación con el concepto de participación política, el indicador adolece de un error sistemático.

• El error cometido en la fase de *operacionalización* puede ser *sistemático* y *aleatorio*. Conviene reflexionar sobre esta cuestión. Por ejemplo, en una investigación basada en encuestas realizadas a una muestra de sujetos, el proceso de operacionalización tiene lugar en varias fases, y en cada una de ellas se podrían cometer errores. En particular, podemos distinguir tres fases[9]: *a*) la fase de *selección* de las unidades estudiadas; *b*) la fase de recopilación de los datos, que llamaremos de *observación*, y c) la fase de *tratamiento de los datos*. Podemos cometer errores en cada una de estas fases.

Errores de selección. Se trata de los errores debidos al hecho de que en una investigación concreta no se analice evidencia sobre el total de la población, sino sobre una muestra de sujetos. Podemos identificar tres tipos de errores de selección. El *error de cobertura,* debido a que la lista de la población de la que extraemos los casos de la muestra no es completa, y excluye a una parte de los sujetos. Esto sucede, por ejemplo, cuando realizamos una entrevista telefónica para estudiar las intenciones de voto del electorado: todas las personas que no tienen teléfono están excluidas a priori, y esto conlleva un error. Otro tipo de error de selección es el *error de muestreo:* hacer la investigación sobre una muestra, en lugar de hacerla sobre la población total, conlleva un error que, con una muestra distinta, también habría sido distinto. El tercer tipo de error de selección es el *error de no respuesta*:

[9] El modelo de investigación empírica al que nos referimos en las páginas siguientes es la encuesta por muestreo, que conlleva, entre otras cosas, el uso de una muestra de la población estudiada, la consideración del individuo como unidad de análisis, el registro de datos mediante preguntas planteadas a los sujetos (por ejemplo, con un cuestionario). Pero las reflexiones tienen validez general y se pueden aplicar también a otros modelos de investigación empírica, con las adaptaciones necesarias.

algunos sujetos incluidos en la muestra que se va a estudiar, pueden no estar disponibles para el entrevistador o negarse a responder.

Errores de observación. Estos errores pueden proceder de cuatro fuentes: el *entrevistador*, que puede condicionar al entrevistado, o equivocarse al registrar las respuestas, etc.; el *entrevistado*, que podría comprender erróneamente la pregunta, dar respuestas deliberadamente falsas («para quedar bien», por ejemplo), no acordarse de la respuesta en preguntas relativas al pasado, etc.; el *instrumento*, que podría contener preguntas mal formuladas que inducen la respuesta, preguntas tendenciosas, etc.; el *modo de aplicación*, la entrevista telefónica, por ejemplo, implica una conversación apresurada que podría provocar respuestas más superficiales que una entrevista cara a cara, mientras que el cuestionario postal puede generar otro tipo de errores, como la falta de motivación por parte del encuestado.

Errores en el tratamiento de los datos. Se producen después de la recogida de datos y consisten en errores de codificación, de transcripción, de memorización en el soporte informático, de elaboración, etc.

Todos estos errores pueden ser sistemáticos o aleatorios. Pensemos, por ejemplo, en los errores debidos al instrumento: una pregunta sobre la renta conllevará siempre una disminución de la renta verdadera (error sistemático); del mismo modo, una pregunta puede implicar malentendidos que pueden variar entre los distintos sujetos (error aleatorio).

Esta breve reflexión sobre el error, esquematizada en la figura 3.4, permite percibir la variedad en la naturaleza y la fuente del error. Esta forma de considerar el error también se denomina enfoque del «error global» ([cfr. Groves, 1989, 14-15]). El error global de una investigación es imposible de evaluar porque hay demasiados componentes que escapan a nuestro control; y aunque las réplicas de una investigación permitirían evaluar el error aleatorio, no ocurre lo mismo con el error sistemático, que por definición se presenta en todas las posibles réplicas de cada investigación y escapa por lo tanto a cualquier forma de control.

Lo que sí podemos medir es un componente de este error global, el error de muestreo. Existe una parte muy consolidada de la disciplina estadística que ha desarrollado técnicas para calcular el error derivado del muestreo. Pero éste es sólo un componente del error global, y probablemente ni siquiera el más relevante en la mayoría de las investigaciones.

Sin embargo, dado que el error de muestreo es el único cuantificable, muchos investigadores (y de forma poco rigurosa) consideran este error como el error total de la investigación. De hecho, tanto en la publicación en un periódico de los resultados de un sondeo, como en un informe de investigación especializada pueden encontrarse afirmaciones del tipo: «las estimaciones presentadas están afectadas de un error de N (por ejemplo $\pm$ 3) puntos porcentuales». Esta afirmación es incorrecta, puesto que da a entender que ése es el error global de la estimación, cuando en realidad es sólo la parte de error debida al muestreo. A ésta tendríamos que añadirle un número de puntos porcentuales para obtener el error global que afecta a los datos y a los resultados.

3.9. FIABILIDAD Y VALIDEZ

Con respecto a la fase que hemos llamado de «observación», que otros llaman medición, por lo que hablan de «error de medición», que excluye, por tanto, los problemas de selección de los sujetos y tratamiento de datos, los psicométricos han elaborado dos nociones que trataremos a continuación: la fiabilidad y la validez.

La *fiabilidad* tiene que ver con la posibilidad de «reproducir» el resultado, e indica el *grado en que un procedimiento concreto de traducción de un concepto en variable produce los mismos resultados en pruebas repetidas con el mismo instrumento de investigación (estabilidad), o con instrumentos equivalentes (equivalencia).* Cuando pesamos un mismo objeto varias veces y el resultado de la medición es casi siempre el mismo, llegamos a la conclusión de que la báscula es fiable. Si al aplicar un test de actitud a un individuo, un día obtenemos una puntuación de 110 y al día siguiente, con el mismo test u otro similar (para evitar efectos de memorización o aprendizaje), la puntuación obtenida es 80, podemos afirmar que ese test no es fiable.

La *validez,* en cambio, hace referencia a el *grado en que un determinado procedimiento de traducción de un concepto en variable registra efectivamente el concepto en cuestión.* Si el CI (cociente intelectual) permite registrar efectivamente la inteligencia, si el PIB (producto interior bruto) permite registrar efectivamente la riqueza de un país, son preguntas que tienen que ver con la validez de estas definiciones operativas. En este caso no es la repetición de la prueba lo que permite controlar este tipo de error. Un procedimiento operativo concreto puede ser perfectamente estable en aplicaciones sucesivas (es decir, fiable), pero ello no implica necesariamente que sea válido. Si, por ejemplo, medimos la inteligencia de las personas por el número de pie que calzan, obtendremos un dato fiable, pero evidentemente no tendrá validez.

Por lo general la fiabilidad se suele asociar al error *aleatorio* y la validez al error *sistemático.* Por ello, la fiabilidad se puede controlar más fácilmente que la validez, puesto que el error aleatorio puede localizarse repitiendo la observación sobre el mismo sujeto (las variaciones entre las repeticiones se deben al error aleatorio), mientras que la validez es más difícil de identificar, puesto que el error sistemático está presente en cada una de las observaciones, y por tanto hace que no se pueda conocer el estado efectivo de la propiedad estudiada.

3.9.1. Fiabilidad

La primera forma de *fiabilidad* estudiada es la *estabilidad* en el momento de la observación. Ésta puede medirse con la técnica de *test-retest,* consistente en repetir la observación (sobre los mismos sujetos) y calcular la correlación entre los dos resultados. Esta solución, sin embargo, es difícil de aplicar en las ciencias sociales por dos motivos: la reactividad del sujeto humano y el cambio que puede introducir en el sujeto entre la primera y la segunda exploración. Es decir, la primera exploración no deja inalterada la propiedad que queremos registrar. Por ejemplo, el recuerdo del primer test o lo aprendido después del mismo pueden alterar el resultado del

segundo. Además, la repetición en el tiempo permite controlar las variaciones entre dos momentos de aplicación del test, pero no los otros dos tipos de variación accidental mencionados, atribuibles al instrumento y a los métodos de investigación.

Se ha propuesto una segunda definición de fiabilidad en términos de *equivalencia,* según la cual la fiabilidad se mide por la correlación entre dos procedimientos distintos, aunque muy similares entre sí. La primera técnica diseñada según esta definición es la denominada *split-half,* en la que la fiabilidad se obtiene por la correlación entre dos mitades del mismo test (las preguntas de un test se subdividen en dos grupos, por ejemplo, preguntas pares y preguntas impares, las puntuaciones se calculan por separado y se relacionan entre sí).

Un procedimiento similar es el de *parallel forms.* Dos test se denominan «paralelos» cuando se considera que miden el mismo «valor verdadero» subyacente, y difieren sólo en el error aleatorio. Generalmente se trata de tests de tamaños iguales, con preguntas formuladas de la misma forma, sobre el misma tema, etc. (pensemos en dos tests de inteligencia en los que las capacidades matemáticas están representadas por la solución de dos bloques de problemas muy similares entre sí y que requieren el mismo tipo de conocimientos y operaciones lógicas, o en los que las capacidades verbales están representadas por el reconocimiento de cuarenta palabras que cambian entre los dos tests, etc.).

También se puede medir la fiabilidad partiendo del supuesto de que los errores aleatorios varían no sólo entre los distintos tests, sino también entre las preguntas de un mismo test. Así, se han propuesto medidas basadas en la *coherencia interna,* en las que la fiabilidad del test se mide por la correlación entre las respuestas para cada pregunta y las respuestas para el resto de las preguntas.

Conviene señalar que todas las técnicas que hemos presentado proceden de la psicometría, donde el único procedimiento de operacionalización es el test psicológico, que consiste en una serie de preguntas sobre el mismo tema (como en un test de inteligencia en el que el sujeto realiza diversas pruebas y la puntuación final es una combinación de las puntuaciones obtenidas en cada una de las pruebas). En otras situaciones resulta mucho más difícil evaluar la fiabilidad de una investigación. Por ejemplo, cuando, manteniendo al individuo como unidad de análisis, se quieren registrar conductas específicas (como el voto, la práctica religiosa, etc.), o bien cuando la unidad de análisis no es el individuo (podría ser un núcleo territorial, un municipio o una región).

En vista de la dificultad (cuando no imposibilidad) de la repetición, se mantiene el planteamiento de base según el cual la fiabilidad de una investigación se puede controlar mediante investigaciones múltiples con instrumentos distintos. Por ejemplo, la práctica religiosa se puede investigar con una pregunta directa o con una pregunta indirecta para saber qué ha hecho el sujeto la mañana del domingo; la lectura de periódicos, mediante una pregunta directa y una pregunta sobre un hecho actual publicado por todos los periódicos, etc.

3.9.2. Validez

Mucho más laboriosa y difícil de controlar es la *validez,* a pesar de su importancia. Si tras un error de validez se esconde un error sistemático, es muy difícil advertirlo, ya que éste se mantiene en todas las investigaciones. Por lo general, el error de validez se sitúa en el paso del concepto al indicador, y surge del «error de indicación», comentado anteriormente (figura 3.4). La validez de un determinado indicador es, en efecto, mucho más difícil de reconocer, y muy difícil de medir.

Los psicométricos han propuesto diversas formulaciones del concepto de validez y los correspondientes procedimientos de validación de un indicador. Pero, en nuestra opinión, el concepto de validez se puede formular en torno a dos aspectos y con dos procedimientos de validación correspondientes: la validez de contenido y la validez por criterio.

La *validez de contenido* se sitúa en un plano teórico y denota que el indicador o los indicadores seleccionados para un concepto cubren efectivamente el significado del concepto. Una validación de este tipo sólo puede darse en un plano puramente lógico (de hecho, algunos autores hablan precisamente de «validación lógica»), y consiste en la descomposición analítica del concepto estudiado, que se corresponde con la fase de división del concepto en dimensiones, para cerciorarse de que todas las dimensiones quedan cubiertas por los indicadores seleccionados.

En la *validez por criterio* la validación se basa en la correspondencia entre el indicador y un criterio externo que se considera correlacionado con el concepto. Este criterio puede estar representado: *a*) por otro indicador aceptado como válido, o *b*) por un hecho objetivo, generalmente de carácter conductual. Dada la posibilidad de cuantificar esta correspondencia (por ejemplo, mediante un coeficiente de correlación), este tipo de validez también se llama «validez empírica», por contraste con la validez de contenido, que se podría denominar «validez teórica» [Lord y Novick, 1968, 261].

Se distinguen varios tipos de validez por criterio: la *validez predictiva,* que consiste en relacionar el dato del indicador con un hecho sucesivo vinculado al mismo. Por ejemplo, el resultado de un test de acceso a la universidad puede correlacionarse con las calificaciones de los exámenes para ver en qué medida puede servir para pronosticar el rendimiento en los estudios. Un test de actitudes sobre una determinada profesión puede validarse mediante la comparación entre el resultado obtenido en el test y el rendimiento sucesivo en el trabajo.

Se habla de *validez simultánea* cuando el indicador está correlacionado con otro registrado en el mismo momento. Por ejemplo, un indicador de conservadurismo político puede correlacionarse con una pregunta sobre el partido al que se ha votado. Un tipo de validez simultánea importante es la *validez por grupos conocidos,* en la que el indicador se aplica a sujetos cuya posición sobre la propiedad en cuestión es conocida. Por ejemplo, un indicador de religiosidad puede aplicarse a personas pertenecientes a grupos religiosos que asisten regularmente a la iglesia, para controlar si efectivamente, como se espera, el indicador registra valores de religiosidad especialmente elevados para estas personas. De forma similar, una escala sobre planteamientos democráticos/antidemocráticos podría validarse aplicán-

dola a sujetos miembros de asociaciones antidemocráticas (por ejemplo, grupos filofascistas).

Los manuales de psicometría incluyen un tercer tipo de validez denominada *validez de constructo*. En mi opinión, ésta es una combinación de los dos tipos de validez anteriores. La validez de constructo consiste en la concordancia de un indicador con las expectativas teóricas sobre las relaciones con otras variables. Por ejemplo, sabemos por diversas investigaciones realizadas, que existe una correlación inversa entre el nivel de estudios y los prejuicios raciales. Si establecemos un nuevo indicador del prejuicio, su validez de constructo se puede evaluar según su concordancia con estas expectativas: si las expectativas no se confirman (por ejemplo, se demuestra que las personas con valores altos en este indicador son también las que tienen un nivel de formación más alto), podemos afirmar que probablemente el indicador considerado no es un indicador válido del prejuicio. El procedimiento se basa en consideraciones teóricas y en la referencia a otros indicadores ya aceptados como válidos, de ahí que pueda considerarse una combinación de los dos criterios de validación anteriores [10].

Para finalizar esta reflexión sobre el error total, recordemos su importancia y la necesidad de que el investigador ponga especial cuidado en el mismo. Los errores cometidos en la transformación de los conceptos en variables son muy graves pues, al producirse al principio de la fase empírica, afectan a todo el desarrollo posterior de la investigación. Los procedimientos de muestreo más minuciosos y los cálculos estadísticos más sofisticados podrían resultar totalmente inútiles si se basan en variables que no reflejan correctamente los conceptos. Además, son errores especialmente delicados, ya que pueden escapar con facilidad al analista menos atento.

3.10. SÍNTESIS DEL CAPÍTULO 3

1. El proceso habitual en investigación social es un proceso cíclico que empieza y acaba en la teoría. Las fases que componen este proceso son: formulación

[10] En otros campos de las ciencias sociales, como la psicología experimental o la ciencia política, se usan también la noción de *validez interna* y *validez externa*. Se trata, sin embargo, de un uso del concepto de validez muy diferente del presentado aquí, ya que no se refiere específicamente a la transformación del concepto en variable, sino a la cuestón más general del significado que se debe atribuir al resultado de una investigación, y en particular a la correspondencia entre la relación hallada entre dos variables según los datos, y la relación que existe efectivamente entre ellas en el mundo real. Para poder hablar de validez interna es necesario que la relación hallada entre X e Y esté presente en los datos, y no sea falsa o ilusoria. Para que haya validez externa esta relación debe existir también fuera del contexto específico (en términos de sujetos estudiados, condiciones experimentales, etc.) en el que se ha realizado la investigación. Por tanto, la validez interna está relacionada con la adecuación de la investigación y el análisis realizado dentro de ella, y la validez externa se refiere al grado de generalización de los resultados obtenidos en situaciones diferentes a las estudiadas. Como puede percibirse, se trata de una aplicación del concepto de validez al margen de la problemática específica de la relación entre concepto y variable, que no utilizaremos aquí por ser poco útil y dar lugar a confusión.

de la *teoría;* deducción de *hipótesis* a partir de la teoría; *recopilación de datos*, a la que se llega mediante la *operacionalización,* que es la transformación de hipótesis en afirmaciones que se pueden registrar empíricamente; *organización y análisis de los datos*; presentación de los *resultados,* mediante un proceso de *interpretación* de las operaciones estadísticas; y por último, vuelta a la *teoría,* mediante un proceso de *inducción.*

2. Una *teoría* se puede definir como un conjunto de proposiciones conectadas entre sí que presentan un grado de abstracción y generalización superior al de la realidad empírica. Una *hipótesis* es una proposición que implica la relación entre dos o más *conceptos,* situada a un nivel de abstracción y generalidad inferior al de la teoría. Las hipótesis permiten comprobar empíricamente la teoría.
3. Dado que la teoría es una interconexión de conceptos, una vez que estos conceptos (entidades abstractas) se concretan, también se concreta la red teórica que conecta los conceptos. Los conceptos se pueden transformar en entidades observables al traducirlos como propiedades de los objetos específicos estudiados, que se denominan unidades de análisis.
4. La *unidad de análisis* representa el objeto social al que se refieren las propiedades estudiadas. En la investigación social los tipos de unidad de análisis más comunes son el individuo, el conjunto de individuos, el grupo-organización-institución, el acontecimiento y el producto cultural.
5. Una *variable* es un concepto operacionalizado. Más en concreto, es la propiedad operacionalizada de un objeto, pues el concepto debe referirse a un objeto, convirtiéndose en una propiedad del mismo. La diferencia entre concepto, propiedad y variable es la misma que hay entre el peso (concepto), el peso de un objeto (propiedad) y el peso del objeto medido con una balanza (variable). Así como los conceptos son los ladrillos de la teoría, las variables son la parte fundamental del análisis empírico.
6. Las variables pueden agruparse en tres clases: nominales, ordinales y cardinales. Cuando una propiedad asume estados discretos no ordenables, hablamos de *variables nominales,* y el procedimiento de operacionalización es la *clasificación.* Las únicas operaciones posibles entre las modalidades de variables nominales son las de igualdad y desigualdad.

   Cuando la propiedad asume estados discretos ordenables, hablamos de *variables ordinales,* y el procedimiento de operacionalización es la *ordenación.* La existencia de un orden nos permite establecer relaciones de igualdad y desigualdad, así como relaciones de orden (mayor y menor que) entre las modalidades de las variables.

   Cuando los procedimientos de operacionalización son el *cómputo* y la *medición* obtenemos *variables cardinales.* El cómputo se aplica cuando la propiedad adopta estados discretos y existe una unidad de cálculo. La medición se aplica cuando la propiedad es continua y existe una unidad de medida. En el caso de las variables cardinales los números que identifican las modalidades (los «valores» de la variable) tienen pleno significado numérico y esto nos permite aplicarles las cuatro operaciones aritméticas.

7. Cuando los conceptos se colocan a un alto nivel de generalidad, puede resultar difícil operacionalizarlos. En este caso utilizamos los *indicadores*, que son conceptos más simples, específicos, y que se puede traducir más fácilmente en términos empíricos. Los indicadores están conectados con los conceptos generales por afinidades de significado.
8. En el proceso que lleva del concepto a la variable se pueden cometer varios tipos de error, y estos errores representan el desfase entre el concepto (teórico) y la variable (empírica). Estos errores pueden ser sistemáticos o aleatorios, y pueden surgir en la fase teórica, durante la cual se seleccionan los indicadores (error de indicación), o en la fase empírica, cuando se registran los indicadores (error de operacionalización).
9. La fiabilidad indica el grado en que un determinado procedimiento de transformación de un concepto en variable produce los mismos resultados en tests repetidos, con los mismos instrumentos empíricos (estabilidad), o con otros equivalentes (equivalencia). La validez indica el grado en que un determinado procedimiento de transformación de un concepto en variable operacionaliza efectivamente ese concepto (y no otros).

3.11. LECTURAS COMPLEMENTARIAS

Sobre la lógica de la deducción científica y sobre la conexión entre teoría, hipótesis y conceptos resulta muy útil el texto clásico de A.L. Stinchcombe, *Constructing Social Theories* (Harcourt, Brace & World, 1968, 303 pp.), sobre todo el capítulo 2. Para un análisis introductorio pero esmerado de la conexión entre teoría e investigación en el contexto de las teorías sociales vigentes véase la obra de D. Layder, *Sociological Practice: Linking Theory and Social Research* (Sage 1998, 191 pp.).

Sobre el problema de la medición remitimos al ensayo clásico de Duncan *Notes on Social Measurement: Historical and Critical* (Russel Sage, 1984, 256 pp.). Sobre medición y sobre el paso de teoría a hipótesis y conceptos, remitimos a las partes I y II del volumen de E. J. Pedhazur Elazar y L. Pedhazur Schmelkin, *Measurement Design and Analysis: An Integrated Approach* (Lawrence Erlbaum, 1991, 819 pp.). Sobre la relación crítica entre indicadores sociales y teorías sociales, remitimos a la obra de M. Carley, *Social Measurement and Social Indicators* (Allen & Unwin, 1981, 195 pp.). Sobre validez y fiabilidad, véase la obra de E. G. Carmines y R. A. Séller, *Reliability and Validity Assessment* (Sage, 1983, 70 pp.).

Sobre el problema del «error total» en la investigación social (y, en particular, en la encuesta por muestreo), remitimos al libro de R. M. Groves, *Survey Errors and Survey Costs* (Wiley, 1989, 590 pp.), que contiene una exposición muy completa de los errores de muestreo, no-respuesta, falta de cobertura y operacionalización inadecuada, así como al libro de P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz y S. Sudman (eds.), *Measurement Errors in Surveys* (Wiley, 1991, 760 pp.), una selección de ensayos sobre los distintos tipos de error y los instrumentos disponibles para evaluarlos.

CAPÍTULO 4

CAUSALIDAD Y EXPERIMENTO

CONCEPTO DE CAUSA 107

CONFIRMACIÓN EMPÍRICA DE LA RELACIÓN CAUSAL 108

4.2.1. Covariación entre variables independientes y dependientes 109
4.2.2. Dirección causal 110
4.2.3. Control de las variables superfluas 110

ANÁLISIS DE LA COVARIACIÓN Y EXPERIMENTO 111

4.3.1. Análisis de la covariación 112
4.3.2. Experimento 113

EL EXPERIMENTO EN LAS CIENCIAS SOCIALES 115

4.4.1. La solución científica 116
4.4.2. La solución estadística 118

EXPERIMENTOS DE LABORATORIO Y EXPERIMENTOS DE CAMPO 120

4.5.1. Experimentos de laboratorio 121
4.5.2. Experimentos de campo 125

EXPERIMENTOS Y CUASI-EXPERIMENTOS 126

4.6.1. Experimentos verdaderos 129
4.6.1.1. Diseño «sólo después» a dos (o más) grupos 129
4.6.1.2. Diseño «antes-después» a dos (o más) grupos 130
4.6.1.3. Diseño de Solomon a cuatro grupos 132
4.6.1.4. Diseño factorial 132
4.6.2. Cuasi-experimentos 134
4.6.2.1. Diseño «antes-después» a un solo grupo 135
4.6.2.2. Diseños en serie temporal interrumpida 136
4.6.2.3. Diseño «antes-después» a dos grupos sin asignación causal 137
4.6.2.4. Diseño «ex postfacto» 138

METODOLOGÍA Y TÉCNICAS DE INVESTIGACIÓN SOCIAL

VENTAJAS Y LIMITACIONES DEL EXPERIMENTO
SÍNTESIS DEL CAPÍTULO 4
LECTURAS COMPLEMENTARIAS

En este capítulo hablaremos del concepto de causa, un concepto muy importante en todo razonamiento científico pero al mismo tiempo muy difícil de traducir en términos empíricos. A continuación trataremos el «experimento», que constituye el instrumento fundamental creado por las ciencias naturales para el estudio del nexo causal y discutiremos el modo de aplicar el experimento a las ciencias sociales. La segunda parte del capítulo está dedicada a la ilustración de los distintos diseños de experimentos que se usan en la investigación social.

4.1. CONCEPTO DE CAUSA

Desde siempre, el hombre se preguntado por las causas de todo que ocurre en el mundo: desde la influencia de los astros sobre la vida en la Tierra, hasta el efecto de la lluvia sobre el crecimiento del trigo, los humanos han tratado de entender cómo un fenómeno podía influir en otro. Sin embargo, aunque el concepto de causa está íntimamente ligado a la idea de ciencia, parece ser uno de los más controvertidos en el plano filosófico y uno de los más difíciles de traducir en términos operativos, expuesto a esa «divergencia entre el lenguaje de la teoría y el de la ciencia que nunca puede se puede salvar totalmente» [Blalock, 1961, 5]. El problema de la causalidad existe en todas las ciencias, pero resulta especialmente relevante en las ciencias sociales, ya que éstas pocas veces pueden servirse del instrumento principal para evaluar empíricamente una relación causal: el experimento.

No entraremos en el debate filosófico sobre el *concepto de causa*. Los filósofos han discutido sobre él durante siglos: desde la doctrina aristotélica, que considera que la noción de causa es el fundamento de la ciencia y su principio de inteligibilidad, pasando por la crítica de Hume, en el siglo XVIII, sobre la imposibilidad de demostrar la conexión causal necesaria y deducir el efecto a partir de la causa y la arbitrariedad de toda previsión, hasta los avances más recientes de la ciencia, en los que los conceptos de función y probabilidad han sustituido al concepto tradicional de causalidad determinista.

En general, se suele aceptar la opinión de que «el pensamiento causal pertenece al nivel teórico y que las leyes causales no se pueden demostrar empíricamente» [Blalock, 1961: 6]. Como escribe Mario Bunge, para expresar el concepto de causa no es suficiente el enunciado:

Si C, entonces E,

la relación e

Si C, entonces (y sólo entonces) E siempre

relación, el nexo causal designa una forma de «generar» cosas a partir de otras cosas» [Bunge, 1959: 46]. Por ello Bunge propone el siguiente enunciado para expresar el nexo causal:

Si C, entonces (y sólo entonces) E siempre, producido por C

El elemento clave que añade este enunciado a los anteriores está en la idea de *producción:* no se limita a afirmar la existencia de una «conjunción constante» entre dos fenómenos, sino que afirma que «más allá de ir acompañado por la causa, el efecto es *generado* por ella» [*ibídem,* 47] [1].

Sin embargo, como escribe Blalock [1961: 10], «la acción de "producir" se refiere a un proceso ontológico [...], tiene una realidad independiente del observador y de sus percepciones [...]. No podemos, de ningún modo, observar o medir estas acciones [de producción]. Probablemente, como mucho, podemos observar algunas covariaciones y series temporales. Pero el simple hecho de que *X* e *Y varíen* conjuntamente de una forma previsible, y que una transformación en *X* preceda siempre a una transformación en *Y,* no nos garantiza que *X* produzca una transformación en *Y*». Por ello, añade Blalock, el concepto de «previsión» (en lugar de causa), se utiliza con frecuencia para sortear la objeción empirista a la idea de causa. Y por tanto la noción de causa queda confinada al reino de la teoría, y debemos aceptar la imposibilidad de comprobar empíricamente las leyes causales.

Aunque en el plano empírico nunca podremos llegar a «probar» de modo definitivo la existencia de una ley causal, la existencia teórica de un mecanismo causal conlleva consecuencias observables en el plano empírico: el hecho de cotejarlas empíricamente no podrá darnos una respuesta definitiva sobre la existencia del nexo causal, nunca eliminará del todo la incertidumbre sobre su existencia y corrección, pero sí nos permitirá corroborar nuestra hipótesis teórica. En otras palabras, nunca podremos decir, *en el plano empírico,* que la variación de *X* «produce» la variación de *Y.* Pero *si observamos que una variación de X va seguida de manera sistemática por una variación de Y, manteniéndose constantes todas las demás causas posibles de Y, disponemos de un fuerte elemento empírico de confirmación de la hipótesis de que X es la causa de Y.*

4.2. CONFIRMACIÓN EMPÍRICA DE LA RELACIÓN CAUSAL

La afirmación anterior nos permite comprender cuáles son las implicaciones empíricas que se derivan de la existencia hipotética de un mecanismo causal. Para poder confirmar empíricamente una hipótesis de relación causal entre dos variables, debe-

[1] En esta breve discusión hemos seguido el planteamiento de Bunge, que atribuye al concepto de causa los atributos de *necesaria, suficiente* y *universal.* Se trata de una visión rigurosa pero restrictiva; sin embargo, no entraremos en el debate en torno a esta cuestión.

mos poder disponer de elementos empíricos sobre tres aspectos: la covariación entre las variables independiente y dependiente, la dirección de la causalidad y el control sobre otras posibles causas.

Antes debemos introducir una cuestión terminológica, relativa a la distinción entre *variables dependientes* y *variables independientes*. En una relación asimétrica entre dos variables, es decir, cuando una variable influye en la otra, pero esta influencia no es recíproca, llamamos variable independiente a la variable que influye, y variable dependiente a la variable que recibe la influencia. En la relación entre clase social e ideas políticas, la clase es la variable independiente, y las ideas políticas la dependiente. Si la relación de dependencia es de tipo causal, *la causa es la variable independiente y el efecto es la variable dependiente* [2]. En una relación bivariada, indicaremos la variable independiente con la letra *X*, y la dependiente con la letra *Y*.

4.2.1. Covariación entre variables independientes y dependientes

En primer lugar, debe ser posible observar una variación de la variable independiente. Por ejemplo, si queremos tener elementos empíricos que confirmen la afirmación teórica de que el individualismo social produce una elevada tasa de suicidios, debemos poder observar situaciones (sociedades, grupos sociales) con distintos grados de individualismo, por ejemplo, sociedades protestantes y sociedades católicas; personas integradas o no en redes sociales, como las personas casadas con hijos y las no casadas, etc. No podríamos sacar conclusiones sobre la influencia de la cohesión social en la tasa de suicidios si observáramos situaciones en las que la variable independiente, es decir, la cohesión social, se mantiene constante. Al variar la variable independiente, debemos poder observar *simultáneamente* una variación de la variable dependiente. En el lenguaje estadístico se dice que debemos poder observar una «covariación» entre las dos variables, es decir, cuando varía una, varía también la otra. Si la teoría afirma la existencia de un nexo causal entre el individualismo social y la tasa de suicidio, debemos poder observar que en las sociedades con un mayor grado de individualismo se da también una tasa superior de suicidios, y viceversa.

---

odelo,

4.2.2. Dirección causal

Es necesario poder observar que al variar la variable independiente se produce una variación de la variable dependiente, pero que no se produce el proceso inverso. Este hecho se puede establecer empíricamente de dos maneras. Mediante la *manipulación de la variable independiente:* si el investigador puede hacer variar la variable *X,* y al hacerlo observa una variación de la variable *Y,* no hay duda de que, si existe un nexo causal, su dirección es de *X* a *Y,* y no de *Y* a *X.*

Esto puede hacerse sólo en el caso del experimento, que, como veremos, prevé la posibilidad de una variación artificial o manipulación de una de las dos variables. Cuando esto no es posible, podemos establecer la dirección del nexo causal mediante el criterio de la *sucesión temporal,* que consiste en observar que la variación de la variable independiente *X* precede a la variación de la variable dependiente *Y.* Si decimos que una socialización con valores religiosos (*X*) genera en la edad adulta una actitud más intolerante hacia otras ideologías (*Y*), el hecho de que la primera variable preceda en el tiempo a la segunda establece la dirección del nexo causal. Así, si se observa una covariación entre título académico y retribución laboral y se puede interpretar como una relación causal, la única dirección causal posible es que el título académico influye en el futuro profesional.

Hay que añadir que ciertas direcciones causales deben excluirse por *imposibilidad lógica.* Si decimos que existe un nexo causal entre clase social e ideas políticas, por lógica, la única dirección causal posible es de la primera variable a la segunda, ya que es imposible que un individuo cambie de clase social como consecuencia de una modificación de sus ideas políticas.

4.2.3. Control de las variables superfluas

Cuando provocamos un cambio en la variable independiente *X,* debemos asegurarnos de que ello no producirá cambios en otras variables que podrían estar correlacionadas con *X.* Es decir, tenemos que asegurarnos de que no son esas otras variables las verdaderas causas del cambio en la variable dependiente *Y.* Por ejemplo, si la tasa de suicidios por regiones aumenta al pasar de regiones católicas a regiones protestantes, pero, al mismo tiempo, todas las religiones protestantes son alemanas, y las católicas, francesas, no podremos discernir si la causa de la variación en la tasa de suicidios es la religión o la nacionalidad.

Por tanto, este tercer aspecto es necesario para poder hablar de control empírico de la relación causal. La observación del primer aspecto, la covariación, no es suficiente para poder hablar de causalidad, porque covariación, no significa causalidad. Por una parte, ambos conceptos se encuentran en planos distintos: el concepto de causalidad pertenece al ámbito teórico, mientras que el de covariación pertenece al empírico. Por otra parte, una covariación *nunca* puede ser la única prueba empírica para afirmar la existencia de una relación causal. Ya hemos hablado de esto anteriormente, al citar a Bunge y señalar que la noción de causalidad incluye la idea de producción y no la mera «conjunción constante» o asociación recurrente. En otras palabras: *puede existir covariación sin que exista causalidad.*

Merece la pena detenerse en las observaciones irónicas de George Bernard Shaw a este respecto, en el prólogo de la obra de teatro *El dilema del doctor,* sobre lo que él llama las ilusiones de la estadística: «Sería fácil probar que llevar un sombrero de copa y un paraguas hace que se desarrolle el tórax, alarga la vida y confiere una relativa inmunidad a las enfermedades, porque las estadísticas muestran que quienes que poseen esos dos objetos tienen el tórax más desarrollado, están más sanos y viven más tiempo que las personas que no pueden ni soñar con poseerlos» [Shaw, 1941; 54]. Otro ejemplo lo encontramos en Italia, donde en ciertas zonas existe una alta correlación entre el consumo de helados y los votos del Partido Radical (en los municipios donde el consumo de helados es mayor al Partido Radical recibe más votos, y viceversa), y sin embargo, a nadie se le ocurre decir que existe un nexo causal entre estas dos variables [3].

En conclusión: si es cierta la afirmación teórica de que *X* es causa de *Y,* entonces será posible observar, en el plano empírico, que una variación de *X,* manteniendo constantes el resto de posibles causas de *Y,* va acompañada de una variación de *Y.* Pero ¿cómo podemos conseguir empíricamente la condición *ceteris paribus,* es decir, «manteniendo constantes el resto de posibles causas de *Y*»? La respuesta depende de si aplicamos la lógica del análisis de la covariación o la lógica del experimento.

4.3. ANÁLISIS DE LA COVARIACIÓN Y EXPERIMENTO

Los científicos disponen de dos técnicas básicas para comprobar una afirmación causal: el *análisis de la covariación* en su manifestación natural y la situación artificial del *experimento.* Decimos que el análisis de la covariación es *natural* porque consiste en analizar las covariaciones tal y como se presentan en situaciones sociales no manipuladas por el investigador y que el experimento es *artificial,* ya que tiene lugar en el laboratorio, donde el investigador produce la variación de la variable independiente.

Ilustraremos la diferencia entre estas dos formas de análisis presentando dos estudios de una misma cuestión, el impacto de la propaganda televisiva sobre las tendencias políticas de los ciudadanos, con los dos procedimientos. Imaginemos una situación en la que hay dos candidatos a la presidencia del país, A y B, que durante la campaña electoral han utilizado la televisión para comunicar sus programas y lograr el apoyo del electorado. Sabemos que sobre el voto influyen numerosas variables «independientes», como la edad, el género, la educación, la clase social, la religión, las ideas políticas de la familia, etc., pero supongamos que no nos interesa la influencia de estas variables, sino sólo el efecto de la propaganda televisiva. Vamos a suponer que *X* es la variable «exposición a la propaganda televisiva», que presenta dos alternativas: «ha visto / no ha visto la campaña electoral en televisión», e *Y* es la variable dependiente «voto», que se divide en dos opciones: «voto para el candidato *A* / para el candidato *B*». El objetivo del investigador es relacio-

[3] Es este caso hay una tercera variable, el grado de desarrollo urbanístico, que influye tanto en el consumo de helados como en los votos del Partido Radical.

nar las variaciones de las dos variables *X* e *Y*, para observar si al variar la exposición a la propaganda televisiva *X* varía también el voto *Y*.

4.3.1. Análisis de la covariación

En un diseño de investigación planteado según la lógica del análisis de la covariación, se entrevistará a una muestra de sujetos, se les hará algunas preguntas sobre las características sociodemográficas de base (edad, género, estudios, etc.), se les preguntará si han visto la campaña electoral en televisión y a qué candidato han votado. Si encontramos una correlación entre las dos variables, es decir, si constatamos que quienes han estado más expuestos a la propaganda televisiva han votado al candidato *A*, ¿podríamos afirmar que la televisión ha influido en el voto en esa dirección?

La respuesta a esta pregunta es *negativa*. De hecho, los dos grupos de ciudadanos, quienes han visto la televisión y quienes no la han visto, no son distintos entre sí sólo por esta variable: probablemente, los que más ven la televisión son las personas que están más tiempo en casa, mujeres más que hombres, ancianos más que jóvenes, etc. Por ejemplo, si las personas que ven más la televisión tienen edades superiores a la media, no podemos saber, a partir de la correlación observada entre exposición a la televisión y voto para el candidato *A*, si la verdadera causa del voto es la edad o la exposición a la propaganda televisiva. Las dos variables se «confunden»: quienes más ven la televisión son también las personas más ancianas. Cabe entonces preguntarse si la causa de votar al candidato *A* es porque son ancianos o porque ven más la televisión.

Aunque hayamos observado una covariación entre *X* (exposición a la propaganda televisiva) e *Y* (voto), podría ser que la verdadera causa sea la edad, que influye en las dos variables: en la audiencia televisiva (los más ancianos ven más la televisión) y en el voto (los más ancianos son más conservadores). En este caso, la covariación entre audiencia televisiva y voto es un ejemplo clásico de lo que en sociología se denomina *relación espuria*, que consiste en una covariación entre dos variables *X* e *Y* que no se deriva de un nexo causal entre las dos, sino de la influencia sobre ambas de una tercera variable *Z* (en el ejemplo anterior, la edad): la variación de *Z* provoca la variación simultánea de *X* e *Y* sin que entre las dos exista un nexo causal.

En un caso como éste, el investigador tiene dos modos de asegurarse de que la relación entre *X* e *Y* no se debe a la acción de *Z* sobre las dos variables: *a*) el *control*, es decir, la transformación de las variables extrañas en constantes, y *b*) la *depuración*, es decir, la determinación por procedimientos matemáticos y la consecuente eliminación de los efectos de las variables extrañas.

En el primer caso se trata de *mantener bajo control* las variables que podrían provocar alteraciones. Si las personas que ven más la televisión por lo general son más ancianas que las que no la ven, por lo que la variable *Z* (edad) altera la relación entre *X* (exposición a la televisión) e *Y* (voto), bastará con mantener constante la variable *Z* para eliminar este efecto. Para conseguirlo analizaremos la covariación entre *X* e *Y* en grupos de individuos con la misma edad. Por ejemplo, los entrevistados pueden dividirse en jóvenes, adultos y ancianos, y dentro de estos tres grupos observaremos la covariación entre propaganda televisiva y voto. Si la relación está

presente en todos y cada uno de los grupos, podemos descartar la sospecha de que quienes más han visto la televisión son también los más ancianos, puesto que hemos comparado grupos en los que la edad era constante. Obviamente, el procedimiento se complica si queremos mantener bajo control muchas variables al mismo tiempo, pues tendremos que hacer comparaciones entre grupos cada vez más pequeños.

Este problema se puede superar mediante la *depuración*, también denominada «control estadístico», de los efectos de las terceras variables, que se realiza mediante cálculos estadísticos: correlación parcial, si se quiere controlar una sola variable, y regresión múltiple, o técnicas similares de estadística multivariada, si se quiere controlar más de una variable.

4.3.2. Experimento

La segunda técnica para comprobar una relación causal entre variables parte de un modo distinto de producir los datos. Ante la pregunta teórica sobre la influencia de la propaganda televisiva en el voto, el investigador plantea un diseño de investigación con las siguientes características: selecciona una muestra de 200 individuos, los subdivide *de manera aleatoria*[4] en dos grupos de 100 personas cada uno, y pide a un grupo que siga la campaña electoral por televisión, y al otro que se mantenga totalmente al margen de ella. Si tras una entrevista sobre el voto de los sujetos estudiados se constata una covariación entre la exposición a la campaña televisiva (X) y el voto (Y), y resulta, por ejemplo, que en el grupo que ha visto la campaña por televisión ha habido muchos más votos para el candidato A, dicha covariación ya no está expuesta a los riesgos de «espuriedad» del caso anterior.

Por término medio, el grupo expuesto a la televisión no tiene una edad superior, ni mayor presencia femenina, ni un nivel de formación inferior al grupo no expuesto a la televisión: de hecho, la asignación de los individuos a cada uno de los grupos se realizó deliberadamente «al azar», para procurar que los dos grupos fueran *por término medio* iguales en cuanto a las características de los sujetos que los forman. Por lo tanto, la variación media del voto (variable dependiente Y) entre ambos grupos puede atribuirse razonablemente al único elemento que los diferencia: la exposición a la propaganda televisiva (variable independiente X).

En el primer caso se trataba de un *análisis de la covariación*, en el segundo, de un *experimento*. ¿En qué difieren estos dos enfoques? En los dos casos se estudia la covariación entre una variable independiente X (la supuesta causa) y una variable dependiente Y (el supuesto efecto). Pero en el primer caso el investigador *observa* y analiza el modo en que las variaciones de X se corresponden con variaciones de Y en un contexto natural. En el segundo caso, en cambio, el investigador *produce* una variación de X en una *situación controlada* y a continuación mide la variación de Y.

---

[4] Nótese que cuando en un contexto científico hablamos de elección «al azar», este término no tiene el significado que se le da en el lenguaje común (una elección improvisada), sino que se refiere a un procedimiento muy preciso que implica una extracción aleatoria (o una técnica matemática equivalente, como el empleo de tablas de números aleatorias). Volveremos sobre ello en el capítulo 8.

Decimos que «produce» una variación de *X* porque la manipula desde el exterior, es decir, la hace variar de manera artificial (en el ejemplo decide exponer o no exponer al sujeto a la propaganda televisiva). Es una «situación controlada», porque mediante la distribución por sorteo de los sujetos en los dos grupos, el investigador tiene bajo control el resto de variables [5], es decir, garantiza que los dos grupos se diferencian, salvo oscilaciones accidentales, sólo en la variación de la variable independiente, es decir, la exposición a la propaganda televisiva.

En el primer caso la intervención del investigador se produce después de la recopilación de los datos, que se limita a analizar. En el segundo caso el investigador controla la producción de los datos, que se produce en una situación artificial creada por él. En definitiva, la idea que subyace al experimento es la siguiente: dada la hipótesis de que *X* es la causa de *Y, si provocamos* una variación en los valores de *X* sobre un determinado número de sujetos y *mantenemos constantes* todas las demás causas posibles de variación de *Y,* debe ser posible observar una variación de *Y* en esos mismos sujetos. La *manipulación* de la variable independiente y el *control* de las terceras variables son, por tanto, los dos elementos que caracterizan al experimento y lo diferencian del análisis de la covariación.

Nótese que, en el experimento, al asignar por sorteo los sujetos al grupo experimental y al grupo de control, se controlan *todas* las variables que podrían ser fuente de variación, también las no registradas o desconocidas por el investigador, y efectivamente los dos grupos que se comparan se diferencian sólo por la variable experimental; en cambio, los procedimientos de control estadístico del análisis de la covariación se pueden aplicar sólo a un número *finito y explícitamente conocido* de variables, es decir, a variables preestablecidas por el investigador y registradas en la fase de la recopilación de los datos.

Aunque el experimento constituye la respuesta más rigurosa al problema del estudio empírico de la relación causal, en la mayoría de las investigaciones sociales es más frecuente la observación de covariaciones que la construcción de experimentos. Por lo general, las variables sociales no son manipulables, y por tanto el investigador raras veces está en condiciones de provocar una variación de la variable independiente. Por ejemplo, si queremos estudiar la relación causal entre individualismo social (expresado, por ejemplo, mediante el indicador «religión dominante») y tasa de suicidios, no podemos tomar una sociedad católica, registrar la tasa de suicidio, después «hacer variar» la variable religión convirtiendo a toda la población al protestantismo, y registrar de nuevo la tasa de suicidios.

Sin embargo, hay situaciones, sobre todo en investigaciones en psicología, en las que es posible esta intervención manipuladora del investigador, y por tanto se pueden realizar experimentos. El resto de este capítulo, lo dedicaremos a estas situaciones de investigación.

---

[5] Se trata de una situación ideal; de hecho, nunca podemos estar totalmente seguros, ni siquiera en los experimentos de las ciencias físicas, de que se mantienen constantes *todas* las demás posibles causas de variación de la variable dependiente.

4.4. EL EXPERIMENTO EN LAS CIENCIAS SOCIALES

Podemos definir el experimento como *una forma de experiencia sobre hechos naturales que se produce como consecuencia de una intervención modificadora y deliberada por parte del hombre. El experimento se diferencia de la forma de experiencia consistente en la observación de los hechos en su desarrollo natural.*

El experimento encuentra su formulación filosófica y sus primeras aplicaciones científicas en el siglo XVII. A principios del mismo, Bacon distinguió entre la *experiencia observadora* y la *experiencia provocada* por la intervención manipuladora del hombre, y Galileo reconoció que el experimento era la base del conocimiento científico moderno. El experimento nace en las ciencias naturales; su aplicación a las ciencias sociales conlleva ciertas adaptaciones o, en todo caso, unas formas de realización particulares. Los experimentos que se pueden realizar en física, por ejemplo, sobre un conjunto de moléculas de un gas determinado indistinguibles entre sí, no puede extrapolarse a las ciencias sociales, en las que la unidad de análisis está constituida por individuos diferentes entre sí. En este apartado describiremos la forma específica que adopta el experimento en las ciencias sociales [6].

Imaginemos una población compuesta por unidades de análisis que llamaremos u; X es la variable independiente e Y la variable dependiente. Para mayor claridad en la exposición, suponemos que la variable X puede tener sólo dos valores: $X = c$ (control) y $X = t$ (tratamiento) [7].

Para seguir con el ejemplo anterior, supongamos que X es la exposición a la propaganda política televisiva, e Y es el comportamiento electoral. Imaginemos la siguiente situación (totalmente irrealizable): que sea posible registrar el valor de Y sobre una misma unidad, con $X = c$, y $X = t$. En nuestro ejemplo, imaginemos que es posible registrar el voto de un mismo individuo en caso de haber estado expuesto a la propaganda televisiva ($X = t$), y en caso de no haberlo estado ($X = c$). Si Y_t es el voto en caso de haber estado expuesto a la campaña, e Y_c el voto en el caso de no haber estado expuesto a la campaña, el efecto causal de la campaña se puede regis-

---

[6] Las primeras aplicaciones del experimento a las ciencias sociales y el origen del concepto moderno de diseño experimental se encuentran en los trabajos realizados en Inglaterra en los años veinte por Sir Ronald Fisher [1935], que llevó a cabo experimentos de agricultura en los que aplicó la aleatorización para controlar las variables extrañas y formuló los procedimientos del cálculo factorial, elaborando sus fundamentos estadísticos y estableciendo las bases de la aplicación «estadística» del método experimental (opuesta a la «científica» de los físicos), que se extendería rápidamente al campo de la psicología.

[7] En las páginas siguientes, al hacer referencia a la terminología de la psicología experimental, llamaremos tratamiento (t) y control (c) a los dos posibles valores de la variable independiente X; entenderemos por tratamiento la exposición de la unidad de análisis a un valor de X, y por control la exposición a un valor distinto que sirve de referencia. En muchos casos este valor de referencia puede ser «ausencia de tratamiento», pero no siempre (también podría ser un tratamiento distinto). Por ejemplo, citaremos el caso de individuos expuestos ($X = t$) y no expuestos ($X = c$) a propaganda electoral, y el caso de las matemáticas enseñadas con la teoría de los conjuntos ($X = t$) y al modo tradicional ($X = c$).

trar mediante la diferencia de comportamiento en ambas situaciones, es decir:

Efecto causal $t = Y_t(u) - Y_c(u)$ sobre la misma unidad u, al mismo tiempo

Obviamente, es imposible observar el valor de Y_t e Y_c en la misma unidad al mismo tiempo, ya que un individuo puede estar o no estar expuesto a la propaganda electoral, pero no ambas opciones al mismo tiempo. De ello se deriva la incapacidad para observar el efecto de X sobre Y. Éste es el denominado *problema fundamental de la inferencia causal*[8]. De la imposibilidad de una observación simultánea sobre la misma unidad Y_t e Y_c se deriva la incapacidad de observar el efecto de X sobre Y, y, por tanto, de comprobar empíricamente la presencia de un nexo causal. Quiero subrayar que tanto en las ciencias sociales como en las ciencias naturales no existe una solución para este problema. No existe proyecto de investigación, unidad de análisis, tipo de dato o tipo de variable, que permitan superar este obstáculo. Ni siquiera en la situación más «objetiva» de la física, el investigador puede resolver el problema fundamental de la inferencia causal. Ésta es la razón fundamental por la que es posible la comprobación empírica de una inferencia causal.

Sin embargo, aunque no sea posible encontrar una solución empírica definitiva y segura para el problema de la inferencia causal, ello no implica que no se puedan alcanzar soluciones parciales. En particular, existen dos vías practicables para ello, que Holland llama *solución científica* y *solución estadística*.

4.4.1. La solución científica

La tesis de invariación consiste en suponer:

— La estabilidad temporal: el valor de Y_c se puede sustituir por una medida del mismo Y_c en un momento anterior.
— La irrelevancia del procedimiento de medición: el valor de Y_t no es influido por la medición previa de Y_c sobre la misma unidad.

Con estas suposiciones, Y_c e Y_t ya no se tienen que registrar de manera simultánea sobre la misma unidad (lo cual es imposible), sino que Y se puede registrar en un momento anterior. En otras palabras, se registra Y_c, se modifica X y se registra Y_t. La diferencia entre los dos valores de Y revelará el efecto causal.

Este modo de proceder es frecuente en las ciencias naturales. Veamos un ejemplo. En 1850 el físico inglés James Prescott Joule, para comprobar empíricamente el principio de conservación de la energía, según el cual calor y trabajo son formas equivalentes de energía, construyó un instrumento en el que unos pesos descenden-

[8] En esta parte sigo el planteamiento de la causalidad de la escuela estadística que aplica el denominado «modelo de Rubin» [Rubin, 1974], y en particular, la contribución de Holland [1986].

tes hacía girar unas palas en un contenedor de agua. Se suponía que el trabajo mecánico producido por las palas haría aumentar la temperatura del agua. El experimento permitió a Joule no sólo demostrar la existencia de este nexo causal, sino también determinar la magnitud de la relación existente entre calor y energía, estableciendo que una caloría equivale a 4.186 julios[9]. Si Y_c es la temperatura del agua antes del «tratamiento», la aplicación de energía mecánica, e Y_t, después del tratamiento, y si u es la unidad de análisis sobre la que hemos realizado el experimento, que en este caso es la masa de agua sometida a la variación de energía, pero que en las ciencias sociales será un individuo, una institución, etc., podemos decir que el efecto causal, en una determinada unidad u, está representado por la diferencia algebraica[10]:

Efecto causal $t = Y_t(u) - Y_c(u)$ sobre la misma unidad, en tiempos distintos

Por tanto, podemos decir que la diferencia entre los dos valores de la variable dependiente Y representa el efecto causal, sólo si es sostenible la tesis de invariación, que es *indemostrable*. En el caso del experimento de Joule la tesis de invariación es sostenible. Se acepta la *estabilidad temporal* de Y_c porque, en una situación de laboratorio, no existen otras formas de variación de energía, aparte de la de X. El valor de Y_c, es decir, la temperatura del agua antes del movimiento de las palas, debería mantenerse estable si X no varía. Por eso la temperatura del agua antes de la rotación de las palas puede considerarse idéntica a la temperatura del agua que se habría registrado al final del experimento *si* las palas no se hubieran movido. Es posible, además, aceptar la irrelevancia del primer registro, ya que la primera medición de la temperatura del agua no influye en la segunda medición.

La tesis de equivalencia establece que dos unidades u y v son equivalentes, es decir, iguales para todos los aspectos relevantes. Es decir, que:

$$Y_c(u) = Y_c(v) \quad Y_t(u) = Y_t(v)$$

En este caso, la medición de Y_c (valor de Y cuando $X = c$) se realiza sobre una unidad (v), y la medición de Y_t (valor de Y cuando $X = t$), sobre la otra unidad (u). Y el efecto causal vendrá dado por $Y_t(u) - Y_c(v)$.

Supongamos que se quiere observar el efecto de un determinado fertilizante sobre una producción agrícola. La lógica del experimento anterior, es decir, de la tesis de invariación, nos llevaría a tomar un terreno, cultivarlo un año sin tratamiento fertilizante, medir la cantidad de producto $Y_c(u)$, y al año siguiente tratar el terreno con

fertilizante y medir de nuevo la cantidad producida $Y_t(u)$. Obviamente, en un experimento de este tipo resulta inaceptable la condición de estabilidad temporal de $Y_c(u)$: entre el primer y el segundo año intervienen otros factores que afectan a la productividad del campo, por ejemplo, las condiciones meteorológicas, por lo que la $Y_c(u)$ del primer año, (producción sin fertilizante), no puede considerarse invariable en el segundo año, (producción que habríamos obtenido en el segundo año si no hubiéramos utilizado el fertilizante), para la comparación con $Y_t(u)$.

En este caso la tesis de la invariación no es sostenible, pero sí la de equivalencia, pues podemos tomar dos unidades iguales para los aspectos relevantes. Pensemos, por ejemplo, en dos campos colindantes, donde las variables relevantes a efectos de productividad (condiciones climáticas, terreno, irrigación, etc.) sean iguales: la unidad v no se somete al tratamiento fertilizante, y sobre ella se mide $Y_c(v)$; la unidad u se somete al tratamiento, y sobre ella se mide $Y_t(u)$. El efecto causal vendrá dado por:

Efecto causal $t = Y_t(u) - Y_c(v)$ sobre dos unidades u y v, al mismo tiempo

También en este caso la tesis de equivalencia es indemostrable.

4.4.2. La solución estadística

Tanto la tesis de la invariación como la de la equivalencia suelen ser inaceptables en las ciencias sociales. Pensemos en el ejemplo del efecto de la propaganda televisiva sobre el voto. Y pensemos en la posibilidad de aplicar la tesis de la invariación en un experimento en el que se le pregunta a un individuo cuál es su intención de voto antes de la exposición al debate televisivo entre los dos candidatos (registro de Y_c o variable dependiente «voto», antes del tratamiento), se lo expone al debate televisivo y se registra de nuevo su intención de voto (registro de Y_t).

La estabilidad temporal de Y_c es poco fiable: en el tiempo transcurrido entre las dos encuestas el sujeto podría haber modificado su intención de voto, no por el efecto del debate televisivo, sino por otros motivos, por ejemplo, debido a un intercambio de opiniones con su pareja, la lectura del periódico, etc. La no influencia de la primera medición también es difícil de sostener: si recuerda su respuesta en la primera entrevista, el individuo podría contestar del mismo modo en la segunda entrevista aunque hubiera cambiado de opinión, para no parecer incoherente [11].

La tesis de la equivalencia de las unidades tampoco es admisible en las ciencias sociales, ya que no es posible encontrar dos unidades equivalentes. En el ejemplo agrícola es presumible que dos campos contiguos derivados de la división en dos de un único terreno se puedan considerar iguales para todas las variables ligadas a la productividad, como la composición del terreno, la irrigación, las variaciones meteorológicas, la exposición al sol, etc. Y también aunque no nos encontremos en una

[11] Los «efectos-memoria» (como en este caso) y los «efectos aprendizaje» (como la repetición de un test de habilidad después de un tratamiento) son los efectos de perturbación más conocidos cuando una encuesta se realiza repetidamente sobre los mismos sujetos. Volveremos sobre este problema en el apartado 6 de este capítulo, en relación con los experimentos «antes-después».

situación de tal similitud, es posible, en las ciencias naturales, identificar entre las características de una unidad aquellas que pueden influir en una variable dependiente Y, y convertir dos unidades en equivalentes en relación con esas características. Si, por ejemplo, queremos medir los efectos de la corrosión de dos ácidos diferentes sobre una lámina de metal, podemos saber cuáles son las características de la lámina que influyen en la corrosión, por ejemplo, el pulido de la superficie metálica, la forma de la lámina, la estructura microcristalina del metal, etc., y plantear un experimento en el que dos láminas de metal idénticas y en las mismas condiciones experimentales se someten a la acción corrosiva de los dos ácidos.

Pero en las ciencias sociales no se pueden encontrar unidades idénticas; la imposibilidad intrínseca de que se repitan, ya sean individuos, instituciones, conjuntos de individuos, etc., representa una de las diferencias epistemológicas fundamentales entre las ciencias sociales y las ciencias naturales: todos los átomos de carbono son iguales entre sí, pero no existen dos individuos idénticos. Así mismo resulta imposible identificar de manera unívoca las variables agentes sobre una determinada variable dependiente Y, para poder encontrar dos unidades al menos similares desde este punto de vista limitado. Mientras que en el ejemplo del cultivo del trigo, una vez identificado el tipo de terreno, el tipo de semilla, la irrigación y las condiciones meteorológicas, se puede presumir razonablemente que se han identificado las principales variables agentes sobre la productividad de un campo determinado, en el caso del sujeto humano hay un número casi infinito de variables que pueden influir en su comportamiento.

La solución a esta problemática es la *solución estadística.* No es posible encontrar dos individuos idénticos a los que exponer a dos valores distintos de la variable X para ver cómo varía la variable Y entre ellos. Pero es posible obtener dos grupos de individuos que sean estadísticamente ***equivalentes***, es decir, que difieran sólo en aspectos accidentales, para poder exponer a uno de ellos a un valor de X y al otro a un valor de X diferente, sobre los cuales registrar la variación de Y. Para obtener grupos con estas características se lleva a cabo un proceso de aleatorización: la muestra inicial de sujetos a los que aplicaremos el experimento se subdivide en dos grupos, y los sujetos se distribuyen «al azar», es decir, mediante sorteo, en un grupo o en el otro. Los dos grupos obtenidos serán equivalentes *en la media* de *todas* las variables que caracterizan a los individuos, desde las variables conocidas y visibles, como el sexo y la edad, hasta las no visibles e incluso inconscientes, excepto en las diferencias accidentales, pequeñas y debidas a la casualidad. Si sometemos a uno de los dos grupos a un valor (c) y al otro a un valor diferente (t) de la variable X y observamos en los dos grupos los resultados medios de las variables Y, estaremos en disposición de cuantificar el efecto causal, que esta vez será el efecto causal *medio:*

$$\text{Efecto causal medio } T = E(Y_t) - E(Y_c) \text{ sobre dos grupos «aleatorizados»}$$

donde E es el «valor esperado», interpretable como «valor medio» ($E(Y_t)$ es la media de los valores Y_t en el grupo de sujetos sometidos a $X = t$).

En psicología, cuando el valor $X = t$ de la variable independiente corresponde a la exposición de un grupo a un determinado tratamiento, y el valor $X = c$ corresponde a

la no exposición al tratamiento, al primer grupo se le asigna el nombre de *grupo experimental,* y al segundo, el de *grupo de control* [12].

Llegados a este punto, las diferencias entre el experimento en las ciencias naturales y en las ciencias sociales deberían estar claras, así como la especificidad del segundo, aunque el procedimiento de aleatorización que hemos descrito como característico del experimento en las ciencias sociales también se puede aplicar en las ciencias naturales. También deberían estar claras las razones que llevaron a escoger el diseño de investigación indicado en el ejemplo sobre la influencia de la propaganda televisiva en el voto (la distribución aleatoria de las personas estudiadas en dos grupos, la exposición de un solo grupo a la propaganda televisiva, etc.).

Una condición previa para la realización de un experimento es la capacidad de control del investigador sobre la variable independiente. Tras haber distribuido a los sujetos en los grupos, debe poder atribuir a los grupos los diferentes valores de la variable X. En otras palabras, debe poder manipular los valores de la variable independiente X. Si esto no es posible, no se trataría de un «verdadero experimento». Nótese en fin que, para simplificar, hemos supuesto hasta ahora que la variable X toma sólo dos valores (presencia/ausencia), de lo que se derivan dos únicos grupos (grupo experimental y grupo de control). Como veremos, la lógica del experimento no cambia aunque la variable X tome más de dos valores (si fueran tres, los individuos se distribuirían en tres grupos, y así sucesivamente). Tampoco cambia la lógica del experimento cuando hay más de una variable independiente (hablaremos de ello más adelante en el contexto del diseño factorial).

4.5. EXPERIMENTOS DE LABORATORIO Y EXPERIMENTOS DE CAMPO

En este apartado y en el siguiente hablaremos del experimento en las ciencias sociales, intentando visualizar mediante la presentación de algunos estudios, las formas

[12] El procedimiento de aleatorización asegura la equivalencia de los dos grupos y permite aplicar a los grupos la lógica del supuesto de equivalencia mencionada antes. También puede utilizarse para aplicar a los grupos la lógica de la tesis de invariación. En el primer caso, Y_c e Y_t se registran en dos grupos diferentes (en uno $X = c$ y en el otro $X = t$), y la aleatorización sirve para controlar la invariación del grupo en el tiempo. Si, por ejemplo, registramos la intención de voto Y_c de un grupo de ciudadanos, después les exponemos al debate televisivo y después registramos de nuevo la intención de voto Y_t de estos sujetos, la diferencia media $Y_t - Y_c$ se puede interpretar como el efecto de la exposición al debate y de otras variables presentes en el tiempo transcurrido entre las dos encuestas, como la difusión de noticias positivas sobre la economía del país, que hacen aumentar la popularidad del candidato del partido que gobierna. Si estudiamos un grupo de control, que no exponemos al debate televisivo, y le pedimos su intención de voto en los mismos momentos que se lo pedimos al grupo experimental, éste también estará influido por las variables extrañas, pero *no* por el debate televisivo. El grupo de control permite identificar el efecto medio de las variables extrañas, que, restado a la variación media en la intención de voto del grupo experimental, permite aislar el efecto debido *sólo* al debate televisivo. La «solución estadística» se puede aplicar para resolver los problemas planteados por la imposibilidad de aceptar la tesis de equivalencia y la tesis de invariación.

de aplicación de este método. No entraremos en los detalles técnicos de cada uno de los experimentos, ni en la vastísima tipología de los distintos diseños experimentales, limitándonos a la presentación de los tipos fundamentales. Este campo de la investigación social ha experimentado un notable desarrollo, dando lugar a una disciplina casi autónoma, por lo que se recomiendo consultar los manuales especializados para profundizar en el tema.

En nuestra presentación distinguiremos por una parte los experimentos de laboratorio de los experimentos de campo, y por otra, los «experimentos verdaderos» de los cuasiexperimentos. La primera distinción se refiere al contexto en el que se realiza el experimento. La segunda concierne a los procedimientos de realización y distingue entre experimentos «verdaderos», es decir, los que cumplen todos los requisitos del método, y los casos en que no se respeta alguno de los requisitos fundamentales del método experimental.

4.5.1. Experimentos de laboratorio

La diferencia entre experimentos de laboratorio y experimentos de campo es sencilla: es la distinción entre el experimento realizado en una situación artificial y el experimento realizado en un contexto de vida real. Para comprender la diferencia pensemos en dos experimentos orientados a estudiar la perturbación que produce el ruido en la concentración mental. En el primero, realizado en un laboratorio, se asigna a unos sujetos tareas que requieren concentración mental en presencia de varias formas de ruidos producidos de manera artificial. En el segundo, realizado en una escuela, se registra en un test el rendimiento de alumnos de clases afectadas por el ruido del tráfico y clases no afectadas por dicha perturbación.

El elemento que caracteriza al experimento de laboratorio es el control, presente en dos aspectos:

a) la posibilidad de mantener fuera del ambiente experimental variables e influencias externas no deseadas, y
b) la posibilidad de que el investigador establezca hasta en los más mínimos detalles las condiciones de realización del experimento.

En lo que respecta al primer punto, el laboratorio persigue, también en las ciencias sociales, el objetivo del *aislamiento* del ambiente experimental, a fin de minimizar la acción de factores externos, como interferencias con las actividades cotidianas, la influencia de los roles sociales y las relaciones dentro de los grupos, etc. En ciencias naturales este objetivo se consigue con operaciones como la esterilización, la insonorización, el aislamiento térmico, las condiciones de vacío, etc.

En cuanto al segundo punto, en el laboratorio el investigador puede «construir» su propio experimento, estableciendo todas las modalidades de su desarrollo. Desde este punto de vista, el laboratorio se caracteriza por una gran *flexibilidad,* ya que permite al investigador producir, y repetir, situaciones experimentales que se diferencian entre sí sólo por un pequeño detalle: por ejemplo, en el caso de los ruidos, puede experimentar si un ruido con intermitencia regular provoca un efecto distinto

al de un ruido con intermitencia irregular o uno continuo; y así sucesivamente.

En resumen, el laboratorio experimental de las ciencias sociales, igual que el laboratorio de las ciencias naturales, permite al investigador eliminar de la realidad social unos factores específicos, aislándolos del contexto en el que se producen, para replantearlos en una situación libre de accidentalidades. El laboratorio permite *crear situaciones* y *realizar observaciones* que en el entorno natural no son posibles: pensemos en la creación de dinámicas de grupo especiales, o en la posibilidad de estudiar mecanismos de interacción gestual y verbal entre personas con una cámara de vídeo o un espejo unidireccional.

Conviene subrayar que si bien las dos condiciones que caracterizan el experimento pueden conseguirse tanto en el laboratorio como en experimentos de campo, es más fácil realizarlas en el laboratorio, precisamente por las posibilidades que éste ofrece de ejercer control sobre el entorno.

Analizaremos las situaciones de laboratorio a partir de la clasificación de Aronson, Brewer y Carlsmith [1985], y retomada en parte por Judd, Smith y Kidder [1991], que clasifica los experimentos en tres grupos, según la tarea asignada a los sujetos: los estudios en los que se requiere la *implicación* de los sujetos estudiados, aquellos en los que se les pide su *opinión,* y aquellos en los que se requiere la *observación* sobre los sujetos experimentales.

En *los estudios de implicación* la intervención manipuladora del investigador actúa sobre los propios sujetos, que por efecto de la dinámica experimental son inducidos a conductas que pueden ser distintas de las que mostrarían si actuaran de manera espontánea. En los años sesenta Stanley Milgram realizó una célebre serie de experimentos de este tipo en la universidad de Yale sobre la obediencia a la autoridad [Milgram, 1974]. El objetivo del experimento era estudiar el mecanismo de la obediencia; para hacer tal cosa en un laboratorio de psicología, una persona con autoridad, en este caso un profesor universitario, inducía a un sujeto a realizar actos que iban cada vez más en contra de su conciencia. En estas circunstancias se invitaba al laboratorio a dos personas para participar en un estudio sobre «memoria y aprendizaje» y el investigador que dirigía la prueba explicaba que se trataba de un estudio sobre los efectos del castigo en el aprendizaje.

A un sujeto se le asignaba el papel de profesor, al otro el de alumno. El alumno era conducido a una habitación y atado a una silla, con un electrodo en el pulso; tenía que aprender de memoria una lista de asociaciones de palabras, y por cada error que cometiera, el profesor, con el fin de evaluar el efecto del castigo sobre el aprendizaje, debía administrarle una descarga de intensidad creciente (de 15 a 450 voltios, pulsando botones con varias inscripciones que iban desde «descarga ligera» hasta «descarga peligrosa»). En realidad, el profesor creía estar participando en un experimento sobre el aprendizaje, mientras que el alumno era un colaborador del investigador y no recibía ninguna descarga (sólo lo fingía). El experimento estaba orientado a comprender hasta qué punto una persona está dispuesta a obedecer a la autoridad (el investigador universitario, que insistía en la realización del experimento), incluso ante el creciente dolor provocado en el alumno, y sus insistentes peticiones de que se suspendiera el experimento.

El experimento se repitió con distintos sujetos (más de 100) de diferentes edades,

géneros y niveles sociales, y con diferentes modalidades experimentales, como contacto físico, oral, ausencia de contacto entre profesor y alumno, presencia o no de colaboradores ficticios del profesor, que según los casos alentaban la continuación o declaraban que no se sentían capaces de seguir con el experimento, etc. La serie de experimentos permitió aclarar algunos de los mecanismos de obediencia a la autoridad. Lo que dejó perplejos a los propios investigadores fue «la voluntad frenética por parte de personas adultas de alcanzar el grado máximo de obediencia a la autoridad, [y ello no con] individuos sádicos y asociales [...], [sino con] personas normales, representativas de varias clases sociales [...]; [es decir, que] individuos normales, que no están motivados por ningún tipo de agresividad especial, podrían llegar a convertirse en cómplices de un proceso de destrucción» [Milgram, 1974: 5-6].

En cuanto a los *experimentos en los que se pide la opinión* a los sujetos, podemos citar un estudio de Pheterson, Kiesler y Goldberg [1971] sobre los prejuicios hacia las mujeres. Según estos autores, una de las causas de la dificultad que tienen las mujeres para conseguir el mismo éxito social que los hombres es un prejuicio negativo hacia las mujeres por parte de las propias mujeres. En este experimento se mostraron ocho cuadros a 120 mujeres; las variables independientes (X) manipuladas fueron en este caso dos: el género del artista (a la mitad de las mujeres se les dijo que el autor era un hombre y a la otra mitad que era una mujer), y el éxito del cuadro (a la mitad de las mujeres se les dijo que el cuadro había obtenido un premio y a la otra mitad se le dijo que simplemente se había presentado a un concurso).

A las mujeres se las hacía pasar una por una a una sala equipada con un proyector; a continuación se les hacía leer una breve nota biográfica sobre el artista, se les mostraba una diapositiva del cuadro, y se les pedía que completaran un cuestionario con cinco preguntas de evaluación sobre el cuadro. Los resultados mostraron que las mujeres daban una valoración superior a los mismos cuadros cuando éstos eran atribuidos a hombres, si eran cuadros que no habían ganado un premio; en cambio, este prejuicio negativo hacia la artista-mujer no se daban con los cuadros que supuestamente habían ganado un premio. Ello demuestra el hecho de que «la mujer que compite es infravalorada por las otras mujeres [...]. Incluso un trabajo equivalente al de un hombre obtiene una valoración inferior, a menos que haya recibido el reconocimiento oficial». Por lo tanto, las mujeres tienen un prejuicio negativo hacia las otras mujeres, salvo cuando éstas tienen éxito, hasta el punto de que, según escriben los autores, «una mujer que tiene éxito puede ser sobrevalorada [...]. Probablemente si durante nuestro experimento hubiéramos atribuido a algunos autores un éxito extraordinario, las mujeres con estas características habrían obtenido una valoración incluso superior a la de los hombres» [*ibídem,* 117-118].

En cuanto a los *estudios de observación,* podemos citar las investigaciones realizadas por Albert Bandura en los años sesenta sobre los efectos de imitación de comportamientos agresivos inducida en los niños por escenas de violencia emitidas por televisión. Entre sus múltiples experimentos citamos uno publicado en 1963, que tenía el objetivo de estudiar la imitación de modelos agresivos y evaluar el impacto de las distintas formas de comunicación de la violencia [Bandura, Ross y Ross, 1963]. Ésta podía ser representada en directo por actores, o presentarse en forma de filmaciones televisivas de esas escenas grabadas de la realidad, o en forma

de dibujos animados. La hipótesis que se quería comprobar era que cuanto más lejanas de la realidad son las representaciones, menor es su impacto.

Los sujetos, un grupo de 96 niños de entre tres y cinco años, se dividieron en cuatro grupos de 24 unidades cada uno. La distribución de los sujetos en los distintos grupos fue casual, con tal de que se cumplieran los requisitos de que cada grupo estuviera constituido por la mitad de niños y la mitad de niñas y que en cada grupo hubiera el mismo número de niños agresivos y no agresivos. Esto se había registrado previamente mediante la observación de su conducta, asignándoles una puntuación en una escala de agresividad. Uno de los grupos actuó como grupo de control y no fue expuesto a ningún estímulo. Los otros tres grupos fueron expuestos a los tres estímulos, escena representada, filmación y dibujos animados. En la escena representada un actor realizaba una serie de actos violentos e inusuales, es decir, actos que un niño difícilmente realizaría si no es por imitación, contra un muñeco de goma (le daba puñetazos, patadas, lo lanzaba por el aire, etc.). El segundo grupo era expuesto a una filmación televisiva de los mismos actos. Al tercer grupo se le representaba una escena similar en forma de dibujos animados, en los que el agresor del muñeco de goma era un personaje común de los dibujos animados infantiles (un gato), y el contexto de la acción contenía elementos que reforzaban la irrealidad de la escena, como un prado y fondos claramente dibujados, casas, árboles, pájaros, etc. de colores brillantes y de diseño imaginativo, música típica de los dibujos animados, etc.

A continuación el niño era sometido a una experiencia de frustración: se le hacía pasar por una sala llena de juguetes muy atractivos, pero que no podía tocar; luego pasaba a una sala donde tenía diferentes opciones de juegos, la mitad eran juegos agresivos, como el muñeco de goma, un saco de arena para dar puñetazos, pistolas que disparan pequeños proyectiles de goma, etc,, y la otra mitad eran juegos no agresivos. El comportamiento del niño se observaba por un espejo unidireccional, se dividía en unidades de pocos segundos, y en cada una de ellas se clasificaba como agresivo o no agresivo, con una subdivisión posterior del comportamiento agresivo en imitativo, es decir, repetición de las acciones agresivas vistas en las escenas previas, y comportamiento no imitativo.

Los resultados mostraron una notable influencia del efecto de la imitación. Los niños expuestos a las escenas violentas se mostraron el doble de agresivos que el grupo de control. También se constató claramente que la exposición a escenas de agresividad no sólo aumentaba la probabilidad de conductas agresivas, sino que además estructuraba su forma, ya que la mayoría de los actos violentos de los niños eran imitaciones de los que habían visto en las representaciones. En cambio, la hipótesis de la relación entre el grado de realismo de la escena y la conducta agresiva se confirmó sólo parcialmente: los niños que vieron escenas reales se mostraron más agresivos que los niños que vieron dibujos animados, pero la diferencia fue moderada.

Los investigadores se fijaron también en la influencia del género de los niños (con una mayor propensión a la agresividad en los varones), en la relación entre la predisposición del niño a la agresividad (medida por la puntuación de conductas anteriores) y las conductas observadas y algunos aspectos más. Con los resultados

de este experimento, Bandura logró importantes avances en su investigación sobre la influencia de la comunicación, sobre todo televisiva, de episodios de violencia, considerados no sólo como refuerzo de la predisposición a conductas agresivas, sino también como elemento de construcción del comportamiento social, por sugerir formas de actuación que los sujetos humanos no adoptarían de manera espontánea.

4.5.2. Experimentos de campo

Entre los experimentos de campo citaremos dos ejemplos. El primero son los estudios de Sherif [1967] sobre las dinámicas de grupo, y en particular sobre las modalidades de desarrollo del sentido del «nosotros» y del «ellos», es decir, sobre los mecanismos de formación de solidaridad hacia los miembros del propio grupo y hostilidad hacia quienes no forman parte del mismo. Sus experimentos duraban unas tres semanas y se realizaban en un campamento de verano; los sujetos eran niños americanos de 11-12 años, que antes del experimento no se conocían entre sí. El campamento de verano se dividía en cuatro fases.

En la primera fase los niños podían interactuar libremente y formar grupos de manera espontánea. En la segunda fase se dividían en dos grupos alojados en dormitorios diferentes y alejados entre sí, y se les encomendaban actividades separadas que requerían cooperación y superación de dificultades comunes. La nueva situación hacía que se desarrollaran fuertes vínculos de amistad dentro de los grupos, hasta el punto de que, aunque los grupos se formaron disgregando deliberadamente los grupos de amigos que se habían formado en la primera fase espontánea, en la nueva situación, el 90 % de los chicos elegía a sus amigos dentro de este grupo (rechazando, por tanto, las amistades anteriores). En la tercera fase se ponían en contacto los dos grupos mediante una serie de juegos competitivos, estructurados de tal forma que crearan un conflicto de intereses entre los grupos (grandes premios para el grupo ganador). La dinámica de la relación entre los grupos se transformaba rápidamente de amistosa rivalidad en abierta hostilidad.

En la cuarta fase a los dos grupos en conflicto se les asignaba tareas que no podían conseguir por separado, pues requerían la colaboración del otro grupo. Los dos grupos volvían a interactuar enseguida, la hostilidad desaparecía y se formaban vínculos que superaban los límites de los grupos anteriores. En este estudio, la variable independiente son los objetivos del grupo y la variable dependiente es el tipo de relaciones interpersonales. El experimento demuestra que las relaciones interpersonales están condicionadas por el contexto estructural en el que se encuentran y actúan las personas.

El segundo experimento de campo que citamos procede de la célebre serie de estudios realizados en los años sesenta por Rosenthal y Jacobson sobre la «profecía que se autocumple», es decir, el hecho de que las expectativas de A sobre la conducta de B pueden influir en B hasta el punto de que éste lleve a cabo efectivamente la conducta esperada [Rosenthal y Jacobson, 1968]. En particular, los dos autores aplicaron este principio a la relación profesor-alumno, afirmando que los niños de minorías étnicas de EEUU (afroamericanos, mexicanos, puertorriqueños, etc.) con frecuencia tienen un rendimiento escolar deficiente, porque eso es exactamente lo

que se espera de ellos. En la aplicación experimental, por cuestiones éticas, los autores crearon en los profesores expectativas positivas (en vez de negativas) con respecto a algunos alumnos, y demostraron que éstas inducían una mejora real en el rendimiento de los citados alumnos.

El experimento se estructuró como sigue: al final del año escolar, en una escuela de un barrio pobre y de inmigrantes de EEUU, los investigadores hicieron un test de inteligencia a todos los alumnos, sin revelar los resultados a los profesores. Al reiniciarse el año escolar los investigadores indicaron a los profesores los nombres de una serie de alumnos (una media de cinco por clase), de los que podía esperarse una alto rendimiento escolar; en realidad, éstos no eran los niños que habían obtenido las puntuaciones más altas en el test de inteligencia, sino niños elegidos al azar. Durante el año escolar se volvió a realizar el test en repetidas ocasiones, y los resultados indicaron que los niños de los que los profesores esperaban los mejores resultados tuvieron, en efecto, esos resultados. Esto puede atribuirse a alguna forma de interacción inconsciente entre el profesor y el alumno: «el tono de voz del profesor, su expresión facial, su forma de actuar [...] pueden ser los medios por los que, probablemente de forma inconsciente, el profesor comunica a los alumnos sus expectativas. Esta comunicación puede ayudar al alumno a cambiar su concepto de sí mismo [...], sus motivaciones y sus capacidades cognitivas» [*ibidem,* 23].

4.6. EXPERIMENTOS Y CUASIEXPERIMENTOS

La distinción entre experimentos de laboratorio y experimentos de campo se basa en el contexto en el que se desarrolla el experimento, sin considerar lo que se define como *diseño experimental,* es decir, las características técnicas y organizativas que definen el experimento: el número de grupos experimentales, el modo en que se han formado, el número y las modalidades de las variables independientes, la eventual repetición de las observaciones, etc. Estos elementos que caracterizan el experimento pueden ser muy variados y, al combinarlos entre sí, dan lugar a múltiples diseños experimentales posibles. En las páginas siguientes presentaremos los tipos más comunes, pero antes de pasar a esta descripción queremos detenernos brevemente en los dos elementos que caracterizan el método experimental: la *distribución de sujetos por grupos* y la *manipulación de la variable independiente.*

- *Distribución de sujetos por grupos.* En este aspecto, la *aleatorización* representa la vía más segura para obtener grupos que se diferencien entre sí, en todas las variables, sólo en pequeñas variaciones accidentales [13]. Así, tene-

[13] La *aleatorización* no tiene por qué implicar una separación física de los sujetos. Puede realizarse también de forma oculta. Por ejemplo, en un estudio sobre la eficacia de la propaganda a favor de la participación electoral, podría tomarse la lista de los electores de un distrito electoral, exponer una mitad a la propaganda y después de las elecciones comprobar sobre las listas electorales si la mitad expuesta a la propaganda ha tenido una tasa de participación electoral superior a la otra mitad.

mos grupos llamados *equivalentes.* Cuando la asignación casual no es posible, un criterio alternativo que puede seguirse es el *emparejamiento,* consistente en formar parejas [14] de sujetos idénticos sobre características que se consideran relevantes para la cuestión estudiada, asignando después un componente de la pareja al primer grupo y el otro al segundo grupo. En este criterio se inspiró, por ejemplo, la investigación, que por cierto no era un experimento, de Sampson y Laub comentada en el capítulo 2, cuya muestra constaba de dos grupos de 500 sujetos cada uno, elegidos de forma que cada uno de los sujetos del segundo grupo se correspondiera en el aspecto sociodemográfico con un sujeto del primer grupo, y lo que distinguía a ambos grupos era la presencia o ausencia de conducta delictiva.

La inferioridad de este procedimiento en comparación con la aleatorización casual se debe a que la equivalencia de los dos o más grupos se garantiza sólo en relación con las variables que sirven de base para la formación de las parejas, es decir, las variables que se mantienen bajo control, y no se sabe si existen otras diferencias importantes entre los sujetos que forman la pareja. Además, esta solución puede llegar a ser difícil de aplicar cuando dichas variables son numerosas, ya que puede suceder que para algunos sujetos no se encuentre un equivalente con el que formar la pareja. Esta técnica resulta útil cuando por algún motivo no se puede realizar la aleatorización, o cuando los grupos sometidos al experimento son pequeños. Recordemos que la aleatorización produce grupos que difieren entre sí por variaciones accidentales. Esas variaciones son inversamente proporcionales al tamaño de los grupos, y por lo tanto pueden ser elevadas si los grupos son pequeños. Entonces puede ser conveniente optar por el control total de pocas variables importantes (mediante el emparejamiento), antes que tener variaciones accidentales, más bien elevadas, sobre todas las variables (aleatorización).

También se utiliza un procedimiento que combina la aleatorización casual y el emparejamiento. Se forman parejas (tríos, grupos de cuatro...) de sujetos idénticos para algunas variables básicas, y después se asignan por sorteo a los dos (tres, cuatro...) conjuntos los componentes de cada pareja (trío, grupo de cuatro...). De esta forma se garantiza la identidad de los grupos respecto a las variables que sirven de base para la formación de parejas, mientras que para las otras variables se tiene sólo una variación accidental entre los grupos. Este procedimiento se aplicó para formar los grupos del experimento de Bandura: los niños de los cuatro grupos del experimento se repartieron en grupos de cuatro sujetos del mismo sexo que mostraban actitud agresiva/no agresiva, y así se consiguió que estas dos variables estuvieran «bajo control» en los cuatro grupos (en cada uno de ellos había seis niños agresivos,

[14] En lo sucesivo nos referimos a «parejas» de sujetos imaginando un experimento constituido sólo por dos grupos (experimental y de control). De hecho, dado que en el experimento el número de grupos es igual al número de modalidades de la variable independiente *X*, éstos pueden ser más de dos. En tal caso, el emparejamiento consistirá en formar grupos de tres, cuatro, o más, sujetos idénticos en cuanto a las características de las variables de control.

seis niños no agresivos, seis niñas agresivas y seis niñas no agresivas).

- *Manipulación de la variable independiente.* Normalmente (aunque no siempre), en el experimento de las ciencias sociales la variable independiente está representada por una variable nominal, y su variación entre los grupos experimentales consiste en que cada grupo presenta una modalidad o categoría distinta de dicha variable. Así, a una clase de alumnos se le enseña matemáticas con la teoría de los conjuntos y a otra, con el método tradicional, a un grupo se le dice que un determinado cuadro lo ha pintado un hombre, y a otro grupo, que lo ha pintado una mujer, a un grupo se le proyecta una filmación determinada y a otro no, tres departamentos de una fábrica son expuestos a tres condiciones de trabajo diferentes, y así sucesivamente.

Nótese que la variable independiente X puede variar *entre grupos* o variar en el tiempo *dentro del mismo grupo,* es decir, las distintas modalidades sobre X pueden ser adoptadas por grupos diferentes en un mismo momento, o bien por un mismo grupo en momentos sucesivos [15]. Para ilustrar esta diferencia compararemos el estudio de Bandura sobre la influencia de la violencia transmitida por televisión y el estudio de Sherif sobre las dinámicas de grupo.

En el primer caso, *distintos grupos* de niños fueron expuestos a varias modalidades de la variable independiente X: un grupo no era expuesto a escenas de violencia (X_1), y los otros tres grupos eran expuestos a distintas formas de comunicación de escenas de violencia (X_2, X_3 y X_4). En el segundo caso, *el mismo grupo* era expuesto a cuatro modalidades sucesivas de la variable X (estructura de la relación entre los grupos): X_1 (grupos formales ausentes), X_2 (grupos separados con actividades orientadas a reforzar la cohesión interna), X_3 (grupos separados en competición) y X_4 (grupos en colaboración).

En el primer experimento la variable X variaba entre los grupos (al mismo tiempo, pues el experimento se podía realizar con todos los grupos simultáneamente); en el segundo experimento la variable X variaba dentro del mismo grupo (en momentos sucesivos). La variación de la variable independiente sobre el mismo grupo de sujetos, registrando sucesivamente la variable dependiente, puede interpretarse como un sucedáneo de la aleatorización: en lugar de exponer a grupos equivalentes a distintas modalidades de la variable X, se expone al mismo grupo a esas modalidades, en diferentes momentos [16]. Sin embargo, como veremos, el procedimiento no está exento de riesgos, debido a las perturbaciones provocadas por la repetición del mismo test y por el tiempo que transcurre entre los tests sucesivos.

Volviendo a los dos elementos del experimento, manipulación de la variable independiente y aleatorización de los sujetos en los grupos, podemos concluir afir-

---

[15] Véase el apartado 5 del capítulo 3, donde se explica que una variable puede variar entre distintos casos (simultáneamente), o en el tiempo (sobre un mismo caso).

[16] Recordemos la distinción introducida en el apartado 4 entre la tesis de equivalencia y la tesis de invariación: la variación entre grupos se corresponde con la tesis de equivalencia, y la variación dentro del mismo grupo con la tesis de invariación.

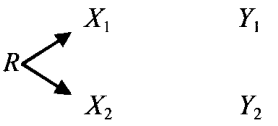
mando que el primer elemento no puede faltar. Si falta tendremos una investigación basada en el análisis de la covariación, aunque, como veremos, algunos autores introducen casos particulares de esta situación en los denominados experimento *ex post facto*. Si falta la segunda condición, tendremos los denominados *cuasiexperimentos*. En función de esta distinción presentamos los principales tipos de diseños experimentales, siguiendo el planteamiento clásico de Campbell y Stanley [1963], que propusieron la distinción entre experimentos y cuasiexperimentos.

4.6.1. Experimentos verdaderos

Para describir los distintos tipos de diseños experimentales emplearemos la representación gráfica de Campbell y Stanley, donde:

- R es la aleatorización de los sujetos en los grupos.
- X es la variable independiente («tratamiento» o «estímulo» en la terminología de la psicología experimental).
- Y es la variable dependiente («observación» o «respuesta»).

4.6.1.1. Diseño «sólo después» a dos (o más) grupos



Efecto causal: $Y_2 - Y_1$

El experimento se desarrolla en tres fases, representadas gráficamente en el esquema por las tres secuencias de las letras R, X e Y: *a*) los sujetos se dividen, mediante aleatorización, en dos grupos; *b*) se manipula la variable independiente de forma que en un grupo asuma la modalidad X_1 y en el otro la modalidad X_2, y *c*) se observa en los dos grupos el *valor medio* de la variable dependiente Y. El efecto causal inducido por la variación de X se mide por la diferencia $Y_2 - Y_1$. La denominación de «sólo después» atribuida a este diseño experimental se refiere al hecho de que la variable dependiente Y se mide *sólo después* de la exposición al estímulo experimental (la variación de la variable X), y no *antes* y *después* de dicha exposición, como sucede con otros diseños experimentales que veremos a continuación.

Éste es el más sencillo de los diseños considerados «experimentos verdaderos», pero contiene todos sus elementos fundamentales: aleatorización, exposición al estímulo (variación entre los grupos de la variable independiente), y registro de la variable dependiente tras la exposición. La aleatorización garantiza que, antes de la exposición al estímulo, los grupos sean *equivalentes,* es decir, que presenten *valores medios* iguales en todas las variables (salvo variaciones accidentales). Por tanto,

tras la exposición a los distintos estados de la variable independiente X, los dos grupos diferirán sólo en el valor de esta variable. De este modo, podemos atribuir al efecto de esta manipulación la diferencia entre las medias de los grupos de la variable dependiente ($Y_2 - Y_1$) registradas después de haber provocado la variación de X.

En el caso más simple las dos modalidades de la variable independiente X están representadas por «presencia/ausencia», es decir, una por la exposición al estímulo experimental, que da lugar al grupo experimental, y la otra por la no exposición al estímulo, que da lugar al grupo de control. Una aplicación de este diseño podría obtenerse reduciendo a dos los cuatro grupos del experimento de Bandura sobre la influencia de la violencia televisiva. Se trataría de: *a*) repartir a los niños en dos grupos; *b*) exponer a un grupo a la transmisión televisiva y al otro no, y *c*) examinar la conducta de los niños en la siguiente sesión de juego. Las diferencias medias de conducta entre los dos grupos son, en este caso, atribuibles a la exposición a las filmaciones televisivas, la única variable que diferencia a los dos grupos.

Cuando la variable X adopta más de dos estados hablamos de *diseño experimental a grupos múltiples*. Es el caso del experimento original de Bandura. La variable X adoptaba cuatro modalidades (ausencia de exposición, escenas reales, filmaciones, dibujos animados), y en tal caso, el esquema del experimento es el siguiente:

$$R \begin{cases} X_1 \text{ (grupo de control)} & Y_1 \\ X_2 \text{ (escenas reales)} & Y_2 \\ X_3 \text{ (filmaciones)} & Y_3 \\ X_4 \text{ (dibujos animados)} & Y_4 \end{cases}$$

4.6.1.2. Diseño «antes-después» a dos (o más) grupos

$$R \begin{cases} Y_1 & X_1 & Y_2 \\ Y_3 & X_2 & Y_4 \end{cases}$$

Efecto causal: $(Y_4 - Y_3) - (Y_2 - Y_1)$

En este diseño experimental, además del registro de la variable dependiente después de la exposición al estímulo, se realiza también un registro antes de la exposición,de ahí su nombre «antes-después». Los dos registros de la variable dependiente también se denominan *pretest* y *postest*. Por lo tanto, este diseño experimental enlaza las variaciones de la variable independiente *entre los grupos*, cuando comparamos los *postest* de los dos grupos, y las variaciones *dentro del mismo grupo en el tiempo*, cuando comparamos el *pretest* y el *postest* del mismo grupo.

Este tipo de diseño es muy frecuente en los estudios de evaluación, cuando se quiere controlar la eficacia de un determinado programa orientado a producir cambios en los sujetos. Si se quiere evaluar, por ejemplo, la eficacia de una campaña de alfabetización, se registran los conocimientos lingüísticos de los sujetos antes y

después de realizarse la campaña, o para valorar la eficacia de una campaña de propaganda política, se registran las tendencias políticas de los sujetos antes y después de la misma.

Estos objetivos también podrían alcanzarse con un diseño experimental de tipo «sólo después a dos grupos» descrito antes. Si queremos evaluar la influencia sobre la orientación política de un debate televisado entre los líderes de dos partidos opuestos, podremos proceder del siguiente modo: elegir un determinado número de electores, dividirlos en dos grupos por sorteo, exponer sólo a un grupo al debate televisivo y registrar la intención de voto de los sujetos al día siguiente del debate. Dado que los individuos fueron distribuidos «al azar» en los dos grupos, las tendencias políticas de ambos grupos antes del debate deberían ser iguales por término medio, salvo pequeñas variaciones accidentales. Si tras la exposición de un solo grupo al debate surgen diferencias suficientemente relevantes entre los dos grupos como para no poder considerarlas accidentales, éstas deben atribuirse al debate. Por tanto, el efecto causal viene dado por la diferencia entre las tendencias medias en los dos grupos: $Y_2 - Y_1$.

El diseño «antes-después a dos grupos» añade un nuevo registro a este esquema: el de la intención de voto antes de la exposición al debate televisivo *(pretest)*. En este caso la variación de tendencia del grupo experimental ($Y_4 - Y_3$) menos la variación de tendencia del grupo de control ($Y_2 - Y_1$) representa el efecto causal.

¿Qué ventaja presenta este diseño experimental con respecto al anterior o, dicho de otro modo, qué añade el *pretest*? Tal y como escriben Campbell y Stanley [1963, 25], el *pretest* «no es fundamental en los experimentos verdaderos», puesto que la aleatorización garantiza la equivalencia inicial de los dos grupos. Sin embargo, el *pretest* añade un control adicional sobre esta equivalencia. En el ejemplo anterior, la asignación de los sujetos a los dos grupos por sorteo debería dar lugar a dos grupos que tengan, antes de la exposición al estímulo experimental, una orientación política media equivalente.

Pero puede suceder, sobre todo si los grupos son pequeños, que las diferencias accidentales entre ambos grupos sean bastante elevadas, y que, como consecuencia, la diferencia $Y_4 - Y_3$ entre los dos *postests* se atribuya (sin que el investigador pueda controlarlo) a diferencias iniciales entre los dos grupos, más que al efecto del estímulo. El *pretest* permite controlar la equivalencia inicial efectiva de los grupos respecto a la variable dependiente y evaluar el impacto causal del estímulo aunque los dos grupos no sean equivalentes, ya que compara la *variación* de la actitud constatada en los dos grupos tras el estímulo, en lugar de comparar el valor absoluto de las tendencias de los dos grupos [17].

---

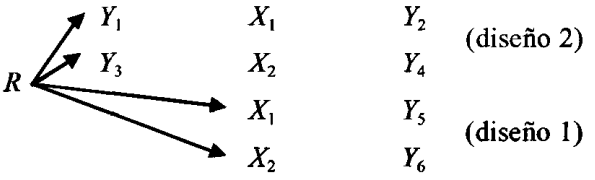
[17] En este diseño el grupo de control sirve para controlar estadísticamente la tesis de equivalencia (entre los dos grupos) y la de invariación (dentro del grupo experimental entre *pretest* y *postest*).

$(Y_4 - Y_2)$: efecto causal + posible efecto de no equivalencia igual a $Y_3 - Y_1$
$(Y_4 - Y_2)\ (Y_3 - Y_1)$: efecto causal libre de no equivalencia
$(Y_4 - Y_3)$: efecto causal + posible efecto de no invariación igual a $Y_2 - Y_1$
$(Y_4 - Y_3)\ (Y_2 - Y_1)$: efecto causal libre de no invariación
Observemos que $(Y_4 - Y_2)\ (Y_3 - Y_1) = (Y_4 - Y_3)\ (Y_2 - Y_1)$.

El diseño *pretest* presenta algunas desventajas. Puede influir en las respuestas del *postest*, sobre todo si los dos tests no se aplican con la suficiente distancia el uno del otro. Supongamos que queremos estudiar la influencia sobre los prejuicios raciales de un reportaje sobre la condición de las personas de color mediante *pretest* a un grupo experimental y grupo de control, exposición del grupo experimental a la filmación, y *postest* a ambos grupos. Las respuestas al *pretest* pueden influir en las respuestas al *postest*, puesto que en éste último el entrevistado puede advertir, dada la insistencia de las preguntas sobre lo que opina de las personas de color, que el test trata sobre los prejuicios raciales. En tal caso las respuestas del entrevistado no expresarán su opinión real, sino lo que se considera socialmente aceptable.

Podría afirmarse, no obstante, que este efecto de perturbación está presente en los dos grupos, ya que ambos realizan el *pretest* y el *postest*, y que, por tanto, no influye en las diferencias entre éstos. Y también podría ocurrir que la influencia del *pretest* sea distinta en cada uno de los grupos, con consecuencias sobre el *postest* que no pueden distinguirse de las diferencias producidas por la exposición al estímulo. Por ejemplo, quienes ven el reportaje sobre las personas de color pueden relacionarlo con las preguntas del *pretest*, percibir que el test trata sobre los prejuicios raciales, y responder de forma sesgada, mientras que los sujetos del grupo de control no expuestos al reportaje podrían no llegar a percibirlo. Estamos ante un efecto de interacción entre *pretest* y estímulo, que puede distorsionar el experimento. El diseño experimental que presentamos en el siguiente apartado pretende precisamente controlar este efecto.

4.6.1.3. Diseño de Solomon a cuatro grupos



Este diseño experimental es una combinación de los dos diseños anteriores: reúne las ventajas del primero (ausencia de interferencia del *pretest*) y las del segundo (disponibilidad del *pretest* como punto de partida previo al estímulo). Con diferencias simples entre los seis valores del *pretest* y del *postest* (Y_1 ... Y_6), se puede separar el efecto del estímulo del efecto de la interacción entre *pretest* y estímulo. Sin embargo, el lector ya habrá percibido la complejidad de un diseño como éste y el coste superior que conlleva el hecho de utilizar cuatro grupos para estudiar el efecto de dos estados de la variable independiente.

4.6.1.4. Diseño factorial

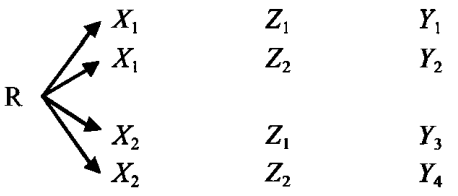
Hasta ahora hemos hablado de una única variable independiente (o estímulo) *X*. Hemos tratado sobre todo los casos en los que ésta adoptaba sólo dos modalidades

(X_1 y X_2), a menudo representadas por las opciones ausencia/presencia. Hemos precisado, no obstante, que todo lo dicho también es válido para el caso en que X_1 y X_2 sean dos modalidades cualesquiera de la variable X, y es fácilmente extensible al caso en que haya más de dos modalidades de X (diseño experimental a grupos múltiples).

En los *diseños factoriales* tenemos más de una variable independiente. Ya hemos citado experimentos con estas características, por ejemplo, en el estudio sobre los prejuicios hacia las mujeres había dos variables independientes, representadas por el género del artista y el éxito, y el experimento reveló que estas dos variables interactuaban entre sí.

Consideremos el siguiente experimento, sobre un tema similar. Costrich *et al.* [1975] formaron cuatro grupos de discusión, compuestos por 5-8 sujetos, dentro de los cuales incluyeron un individuo que fingía ser un sujeto experimental como los demás, pero que en realidad interpretaba un papel que le habían asignado los investigadores. Dos de los «actores» eran hombres, y dos mujeres; dos de ellos, un hombre y una mujer, adoptaban un papel de liderazgo dentro del grupo, y los otros dos, mostraban una actitud de sumisión. Tras la discusión, se pedía a los sujetos que dieran una puntuación de popularidad-simpatía a los componentes del grupo. La hipótesis de los investigadores era que la puntuación estaba condicionada por los estereotipos sociales sobre los comportamientos que se consideran adecuados en los hombres y en las mujeres.

En este experimento tenemos dos variables independientes: el género del actor X y su comportamiento Z. Se trata de dos variables dicotómicas (X_1 hombre, X_2 mujer; Z_1 dominante, Z_2 sumiso) que dan lugar a lo que en el lenguaje de los experimentos se llama diseño factorial 2 × 2, que significa que tenemos dos variables independientes (o *factores*), cada una de las cuales asume dos valores. Este diseño puede esquematizarse como sigue:



El esquema nos dice que los sujetos se han distribuido de manera aleatoria en cuatro grupos: en el primer grupo tenemos un actor hombre con comportamiento dominante (la variable X toma el valor 1 y la variable Z también), en el segundo grupo tenemos a un actor hombre con comportamiento sumiso (X_1 y Z_2), y así sucesivamente. Si la variable independiente Z tomara tres valores, por ejemplo, dominante, sumiso y neutral, el diseño sería 2 × 3 y se necesitarían 6 grupos, y seis actores, uno por cada grupo. Si quisiéramos añadir al diseño inicial 2 × 2 una tercera variable independiente, por ejemplo, la edad, que puede adoptar los valores, «joven» y «adulto», el diseño se convertiría en 2 × 2 × 2 y los grupos serían ocho. Como puede percibirse, el diseño experimental se va complicando a medida que aumenta el nú-

mero de variables independientes y el número de categorías de estas variables.

La principal ventaja del diseño factorial es que permite al investigador estudiar no sólo el efecto aislado de las variables independientes sobre la variable dependiente, lo que se podría hacer también con una secuencia de experimentos simples, cambiando el valor de la variable independiente, sino también el efecto de la interacción entre las variables independientes.

Por ejemplo, en el citado estudio de Costrich y otros autores no se registró una diferencia significativa en las puntuaciones de simpatía recibidas por actores y mujeres, ni se constataron diferencias significativas de simpatía hacia actores con un comportamiento dominante o sumiso. En otras palabras, las variables género y comportamiento, consideradas por separado, resultaron no influir en el índice de simpatía. Pero esto se debía a la interacción de dos fuerzas opuestas: la mujer dominante y el hombre sumiso resultaron ser muy antipáticos, en cambio, el hombre dominante y la mujer sumisa despertaron simpatía. En otras palabras, el grupo rechazaba al sujeto, hombre o mujer, cuando transgredía los estereotipos de género. En términos estadísticos esto se expresa diciendo que las variables «género» y «comportamiento» presentan un efecto principal no significativo, mientras que su *interacción* sí es muy significativa. Este resultado no se habría podido obtener realizando dos experimentos separados con una sola variable independiente: el único modo de obtenerlo es con un diseño factorial como el descrito.

4.6.2. Cuasiexperimentos

Como ya hemos dicho los cuasiexperimentos son «experimentos que tienen tratamiento, observaciones, *postest* y grupos experimentales [como los experimentos verdaderos], pero que no emplean la aleatorización para hacer comparaciones de las que deducir los efectos del tratamiento [...]. La comparación se basa, pues, en grupos no equivalentes que difieren entre sí también por otras características, aparte de diferir en el tratamiento cuyos efectos se están estudiando» [Cook y Campbell, 1979, 6].

El hecho de que el investigador ya no pueda partir de la equivalencia de los grupos antes de su exposición al estímulo representa evidentemente una importante limitación de la lógica experimental. Si inicialmente los grupos no son equivalentes, el investigador nunca podrá saber si las diferencias que constata entre los grupos en los valores de la variable dependiente se deben a la no equivalencia inicial o al efecto del estímulo. Dada la imposibilidad de aislar el efecto del estímulo experimental de todos los demás efectos, el investigador no puede hacer deducciones causales sobre el efecto de la variable independiente sobre la dependiente. Por ello, algunos autores niegan la validez de los cuasiexperimentos y los consideran una categoría híbrida y confusa.

En la investigación social a menudo es imposible aleatorizar a los sujetos en grupos, en particular en situaciones en las que los grupos son preconstruidos (como clases escolares, departamentos de trabajo, etc.), situaciones a las que, no obstante, se pueden aplicar diseños experimentales muy cercanos a la lógica de los «experimentos verdaderos» y que producen resultados interesantes que no se podrían conseguir con investigaciones basadas en el análisis de la covariación.

4.6.2.1. Diseño «antes-después» a un solo grupo

$$Y_1 \qquad X \qquad Y_2$$

Efecto registrado: $Y_2 - Y_1$

Por definición, si sólo hay un grupo, no podemos tener un «verdadero experimento». Sin embargo, este diseño experimental es muy importante como sucedáneo del diseño experimental «sólo después» a dos o más grupos, en el que teníamos dos grupos, uno expuesto al estímulo y uno no expuesto, y el posterior registro de la variable dependiente. Entonces podíamos comparar los registros Y_2, después de la exposición al estímulo, e Y_1, sin exposición al estímulo, atribuyendo la diferencia al efecto del estímulo.

También en el caso que estamos tratando ahora tenemos una observación de la variable dependiente sin exposición al estímulo (Y_1) y otra después de una exposición al estímulo (Y_2), la diferencia es que ahora las dos observaciones se realizan sobre el mismo grupo. Se trata, por tanto, de la variación de X en el tiempo y dentro del mismo grupo.

En otras palabras, en lugar de la tesis de equivalencia (entre grupos) aplicamos la tesis de la invariación (del mismo grupo). No obstante, sabemos que la tesis de la invariación requiere estabilidad temporal y ausencia de influencia del primer registro. Es necesario proceder de tal forma que entre las dos observaciones Y_1 e Y_2 no se produzca ningún otro hecho, aparte de las variaciones de X, que pueda influir también en Y, mezclándose con el efecto del estímulo. Así mismo, es necesario que el *pretest* no influya en el *postest*.

Volvemos al ejemplo del experimento sobre el efecto de la propaganda política en el voto. Si registramos la intención de voto de un grupo de personas, después las exponemos al estímulo del debate televisivo entre los candidatos y a continuación registramos de nuevo su intención de voto, no podemos afirmar con seguridad que la eventual variación en la orientación media del grupo se deba a la exposición al debate televisivo. Entre los dos registros pueden haberse producido hechos externos incontrolados (lectura de periódicos, intercambio de opiniones con familiares, noticias sobre la marcha de la economía, etc.), que podrían haber influido de forma determinante en la intención de voto de los sujetos.

Por estos motivos se deberá procurar que no transcurra mucho tiempo entre el *pretest* y el *postest*. Pero entonces nos enfrentamos a otro problema: este diseño implica una influencia potencial del *pretest* sobre el *postest*, sobre todo cuando no transcurre mucho tiempo entre ambos. De este modo, puede suceder que Y_2 difiera de Y_1 no por efecto del estímulo de X, sino por efectos de aprendizaje o de memoria, o incluso otros efectos debidos al hecho de que Y_1 ha precedido a Y_2.

Nótese que estas dos perturbaciones no estaban presentes en los «experimentos verdaderos» anteriores, con el diseño experimental 1: sólo había un test, por lo que no existía el problema del tiempo transcurrido entre *pretest* y *postest* ni el de la influencia del *pretest* sobre el *postest*. En cambio, dichas perturbaciones sí estaban

presentes, aunque anuladas por el grupo de control, en el diseño experimental «antes-después» a dos o más grupos: los dos efectos estaban presentes en la misma medida en el grupo experimental y en el grupo de control, y la diferencia de las puntuaciones entre ambos grupos anulaba esta diferencia.

Dicho esto, no se excluye que este tipo de diseño cuasiexperimental pueda ofrecer contribuciones útiles. Pensemos en todos los programas de evaluación que no tienen un grupo de control. No obstante, dadas sus limitaciones, es aconsejable recurrir a este diseño sólo cuando no es posible aplicar otros diseños experimentales o cuasiexperimentales.

4.6.2.2. Diseños en serie temporal interrumpida

$$Y_1 \quad Y_2 \quad Y_3 \quad X \quad Y_4 \quad Y_5 \quad Y_6$$

Este diseño también se realiza sobre un solo grupo, y está pensado para casos en los que la diferencia en el valor de Y antes y después de la exposición al estímulo se debe no tanto al efecto del estímulo, sino a una tendencia de Y que ya estaba en curso. Por tanto, este diseño no compara los valores medios, sino las *tendencias* de Y antes y después del estímulo.

El diseño consiste en una sucesión de registros de la variable dependiente Y. En un cierto punto de la secuencia temporal se produce una variación de la variable independiente X y se registra si este hecho introduce una variación en la evolución de la tendencia de Y. Imaginemos, por ejemplo, una escuela donde los indicadores de indisciplina están aumentando cada vez más. Supongamos que en un momento determinado se introduce una novedad importante en el centro (nuevo reglamento disciplinar, cambio del director, autogestión estudiantil, etc.), y que se observa el eventual cambio en los indicadores de disciplina.

Con respecto al diseño anterior (sólo dos registros de Y), este diseño tiene la ventaja de estar poco alterado por la influencia del *pretest* sobre el *postest,* dado que esta influencia está presente en todas las observaciones de Y, salvo en la primera, y por la concurrencia de acontecimientos externos incontrolados entre dos registros sucesivos, que podrían producirse en todos los intervalos entre Y_1 e Y_2, entre Y_2 e Y_3, entre Y_3 e Y_4, etc., y no sólo coincidiendo con la modificación de X. Por lo tanto, si se produce una variación en la evolución de la tendencia de Y entre Y_3 e Y_4 simultánea a la variación de X, es improbable que se deba al efecto del pretest Y_3, o a la concurrencia de otros factores desconocidos, por lo que parece aceptable atribuirla al efecto causal de X.

Este tipo de diseño es aplicable sólo en casos particulares, por ejemplo, para estudiar los efectos producidos por un cambio de normativa. Un ejemplo sería el estudio del controvertido efecto de la pornografía sobre los delitos sexuales. A mediados de los años sesenta en Dinamarca, una serie de medidas legislativas de liberalización de la prensa pornográfica condujo a una mayor difusión de ésta. Un simple análisis del antes y el después (comparando la tasa de delitos sexuales después de la entrada en vigor de la ley con la tasa de delitos antes de la misma), no

habría sido suficiente para evaluar el efecto de esta medida sobre los delitos sexuales, puesto que éstos estaban disminuyendo desde el año 1956. La disponibilidad de la serie temporal de los datos con relación a la variable dependiente, la tasa de delitos sexuales, permitió aplicar un «diseño en serie temporal interrumpida», que demostró la contribución positiva de la nueva normativa a la tendencia de disminución de los delitos sexuales [Kutchinsky 1973]. Otros estudios realizados con el mismo método fueron los de Wagenaar [1981] en Michigan, para evaluar el efecto del aumento de la edad mínima para adquirir bebidas alcohólicas sobre los accidentes de tráfico debidos al consumo de alcohol.

4.6.2.3. Diseño «antes-después» a dos grupos sin asignación causal

Y_1	X_1	Y_2
Y_3	X_2	Y_4

Efecto registrado: $(Y_4 - Y_3) - (Y_2 - Y_1)$

El esquema que ilustra este diseño cuasiexperimental es similar al esquema del diseño experimental «antes-después» a dos o más grupos, con la diferencia de que desaparece la letra *R* con las flechas dirigidas a los dos grupos, que indica la «aleatorización», y aparece una línea horizontal entre los grupos, que indica su origen separado y su no equivalencia.

Este diseño cuasiexperimental es muy frecuente y es un buen sustituto de los experimentos verdaderos cuando no es posible distribuir a los sujetos en grupos de manera aleatoria. Consiste en tomar dos grupos, realizar el *pretest*, aplicar el estímulo sólo a uno de los grupos, y después realizar el *postest* a los dos grupos.

En comparación con el caso anterior, la presencia del grupo de control elimina los efectos de perturbación debidos a hechos ocurridos entre el *pretest* y el *postest* y a la influencia del *pretest* sobre el *postest*, puesto que se trata de efectos presentes en ambos grupos que no influyen en las diferencias entre ellos. Obviamente, se mantiene el hecho de que los dos grupos no son equivalentes, pero el *pretest* proporciona información sobre las diferencias preexistentes entre los grupos, pues permite al investigador trabajar no sobre las puntuaciones absolutas de la variable dependiente, sino sobre la variación *postest/pretest*, que se ven afectadas sólo en parte por el nivel inicial del *pretest*.

De este modo, si queremos analizar el efecto sobre el aprendizaje escolar de un nuevo método didáctico, por ejemplo el estudio de un idioma extranjero con medios audiovisuales, podemos evaluar los resultados de estos nuevos medios comparando dos grupos de alumnos, uno de los cuales aprendería con medios audiovisuales y el otro con medios tradicionales, registrando el aumento de conocimientos con respecto a la situación anterior al inicio del experimento. El efecto sobre dicho aumento de las diferencias en el nivel inicial de ambos grupos debería ser moderada, sobre todo si se ha tenido la precaución de elegir dos grupos con un rendimiento similar.

4.6.2.4. Diseño «ex postfacto»

$$\frac{X_1}{Y_1} \quad \text{Emparejamiento} \quad \frac{X_2}{Y_2}$$

Cuando el investigador no puede distribuir los sujetos en grupos de manera aleatoria y tampoco puede controlar el estímulo, ya no podemos hablar de experimento. En tal caso, nos encontramos en la lógica del análisis de la covariación. Sin embargo, existen situaciones de investigación que, aunque carecen de los dos elementos que caracterizan al experimento, aleatorización y manipulación de la variable independiente, presentan un diseño muy similar al de los experimentos. Estos diseños se denominan *ex postfacto* (para subrayar que se basan en hechos que ya han sucedido) y, aunque no se puedan incluir ni en la categoría de los experimentos verdaderos ni en la de los cuasi-experimentos, en los manuales de metodología se suelen incluir en el capítulo de los experimentos, ya que se inspiran en la lógica experimental, tanto en el diseño como en el tratamiento de los datos.

Por lo general se plantean de la siguiente forma. Considerando que el hecho ya ha sucedido, es decir, que un determinado estímulo ya ha sido aplicado a un grupo de sujetos, se eligen otros tantos sujetos con características lo más parecidas posible a las del grupo anterior, pero que no han sido expuestos al estímulo, y se comparan las medias de la variable dependiente en ambos grupos. El esquema propuesto anteriormente ilustra este procedimiento: un grupo de personas ha sido expuesto a X_1 (que también puede significar no exposición al estímulo), y otro ha sido expuesto a X_2. A continuación, se procede al «emparejamiento», es decir, se seleccionan parejas que contengan un sujeto de cada uno de los dos grupos (uno expuesto a X_1 y el otro expuesto a X_2), que sean idénticos entre sí en una serie de características básicas definidas por el investigador. Una vez formados los dos grupos, se procede a la observación de la variable dependiente Y.

Entre los estudios basados en un diseño *ex post facto* podemos citar uno realizado por William Goldfarb [1945] sobre los efectos de la institucionalización en los niños. Goldfarb comparó a 40 niños que habían pasado dos años en un orfanato antes de ser adoptados por una familia, con otros tantos niños que, desde su nacimiento, habían crecido en familias con las mismas características de las primeras. El investigador constató que los niños que habían pasado por un orfanato sufrían trastornos de personalidad. Un planteamiento similar tienen los estudios sobre los gemelos a los que algún hecho de la vida ha separado y llevado a experiencias distintas, pero que se supone que son muy parecidos, por ser gemelos, en una serie de características básicas.

Otro ejemplo muy distinto de los anteriores son los estudios sobre los efectos de los terremotos en el estado psicológico de las personas que los han sufrido, que comparan pueblos afectados por un terremoto con pueblos lo más parecidos posible, pero que no se han visto afectados por catástrofes sísmicas.

4.7. VENTAJAS Y LIMITACIONES DEL EXPERIMENTO

El método experimental en las ciencias sociales presenta dos ventajas fundamentales. La primera es de carácter metodológico: se trata del método de investigación que permite abordar mejor el problema de la relación causal. La segunda se refiere a la naturaleza de los problemas que permite estudiar: con el método experimental podemos aislar de la vida cotidiana fenómenos específicos que en condiciones naturales no se podrían estudiar de forma tan sistemática, debido a la coexistencia de otros factores que los ocultan, confunden y distorsionan, y por el «ruido» de fondo presente en la vida real, que eclipsa los fenómenos menos llamativos.

Obviamente el experimento sólo se puede utilizar para estudiar determinados fenómenos y situaciones sociales. No puede realizarse cuando la variable independiente no es manipulable, ya sea por dificultades intrínsecas o por cuestiones éticas. Por lo general es aplicable a situaciones «micro» más que «macro».

El experimento consiste en la observación de una situación artificial cuyas características establece el propio investigador. Las causas perturbadoras se reducen al mínimo, las que producen el fenómeno se estudian con variaciones graduales y el experimento se puede repetir. Pero esta flexibilidad para determinar las condiciones constituye precisamente la mayor limitación del experimento. En términos generales, los inconvenientes del método experimental en sus aplicaciones a las ciencias sociales son dos: *su carácter artificial* y *su falta de representatividad.*

- El carácter artificial tiene dos aspectos: el *entorno* y la *reactividad de los sujetos* experimentales. En cuanto al entorno, pensemos, por ejemplo, en los experimentos de Milgram sobre la sumisión a la autoridad. Éstos se desarrollan en un laboratorio, estimulando una relación de autoridad artificial y muy distinta de las situaciones de autoridad/sumisión que encontramos en la vida real. Los efectos producidos en los niños por la exposición a contenidos televisivos violentos pueden ser distintos en un entorno artificial creado por el investigador, como el experimento de Bandura, en comparación con el medio natural. Este problema se plantea sobre todo en los experimentos de laboratorio y es menos relevante en los experimentos de campo.

  El segundo aspecto tiene que ver con la reacción del sujeto humano ante el hecho de sentirse observado. Si para hacer un experimento con un animal tenemos que encerrarlo en una jaula o inmovilizarlo, ello puede tener efectos desconocidos sobre el fenómeno estudiado. Al sentirse observado, el hombre puede comportarse de forma distinta a como se comportaría en un contexto natural. Podemos recordar una vez más el célebre caso de los experimentos de Hawthorne, cuando el hecho de saberse observadas llevó a las obreras a aumentar su productividad, independientemente de los cambios que se producían en las condiciones laborales.

  Una variante de este aspecto es el denominado *efecto del experimentador:* las expectativas de quien realiza el experimento pueden transmitirse, a través de sus actos inconscientes, a los sujetos experimentales, e influir en su comportamiento. Ya hemos visto en el experimento de Rosenthal y Jacobson que las

expectativas pueden influir en el comportamiento de los individuos objeto de las mismas. Rosenthal [1966] obtuvo resultados análogos sobre los animales. Realizó un experimento con ratas, diciendo a los experimentadores, en este caso, estudiantes, que unas ratas eran listas y otras tontas, aunque en realidad las ratas se habían repartido en dos grupos al azar. Los estudiantes que trabajaron con las ratas que se consideraban listas obtuvieron mejores resultados: estos estudiantes trataban a las ratas de forma más amistosa, las estimulaban con delicadeza, las animaban con la voz, en comparación con los estudiantes que trabajaron con las ratas consideradas tontas.

El problema de la reactividad es más evidente en los experimentos de laboratorio, donde resulta difícil, si no imposible, que los sujetos no sepan que están participando en un experimento. Pero la reactividad también constituye un elemento de perturbación grave en muchos experimentos de campo.

- El segundo problema es el de la *falta de representatividad,* en el sentido de que con frecuencia los resultados de un experimento no se pueden generalizar a toda la población o a segmentos de la misma distintos del estudiado. Esto es debido a dos cuestiones: la *amplitud* de la muestra y los criterios de *selección* de los sujetos experimentales.

  La primera cuestión se deriva del hecho de que normalmente los experimentos sólo pueden realizarse sobre muestras muy reducidas. En los ejemplos citados hemos hablado de grupos experimentales compuestos por pocos sujetos (en torno a una veintena de casos, pero a menudo, menos). Esto afecta a los experimentos de laboratorio, pero también a los experimentos de campo. No es posible hacer variar la variable independiente y mantener constantes el resto de las variables cuando la muestra de los casos estudiados es demasiado grande. Esta limitación conlleva que la muestra será es poco representativa. Sabemos que el error de muestreo y la amplitud de la muestra son inversamente proporcionales; por lo tanto, una muestra de pocos casos estará afectada por un error de muestreo elevado y será menos representativa de la población, con las consiguientes dificultades para generalizar los resultados obtenidos.

  La segunda cuestión es la selección de los sujetos. Ya que, por los motivos mencionados, no se puede aspirar a que los grupos experimentales sean muestras representativas de la población, con frecuencia el investigador renuncia a plantearse el problema de los criterios de selección de los sujetos a los que aplicará el experimento. Así, puede suceder que, en lugar escoger una muestra socialmente variada, escojan sujetos de un entorno social específico, por cuestiones de comodidad. Es muy frecuente, por ejemplo, que los sujetos experimentales sean estudiantes universitarios, un segmento de la población muy particular y poco representativo del conjunto de la sociedad.

  Hay que añadir, por otra parte, que el objetivo del experimento no es describir cómo se presenta un determinado fenómeno en la sociedad (para este objetivo existen otras técnicas de investigación social mucho más eficaces), sino analizar relaciones causa-efecto. Éstas se pueden estudiar también sobre situaciones y segmentos específicos de la población.

Estas consideraciones críticas sobre la técnica del experimento nos llevan a las siguientes conclusiones: el carácter artificial del contexto experimental implicaría que los resultados del experimento no se pueden generalizar a las situaciones más complejas del mundo real, y la falta de representatividad de los sujetos estudiados implicaría la imposibilidad de generalizar los resultados experimentales más allá de los pequeños grupos de sujetos que han participado en el experimento (como mucho, podrían extenderse al segmento social al que pertenecen dichos sujetos).

Pero la historia de la investigación social muestra que el método experimental han permitido realizar importantes avances. Además, hay cuestiones que se pueden estudiar sólo con este método. Leslie Kish aporta el relato de un hombre que bebió demasiado en cuatro ocasiones diferentes. Una vez bebió güisqui con soda, otra vez *bourbon* con soda, otra vez ron con soda, y por último vino con soda. En todas estas ocasiones acabó totalmente borracho, y aplicando una lógica aplastante, decidió que nunca volvería a tocar la soda [Kish, 1959, 333]. Y es que, si aplica la lógica del análisis de la covariación, resulta difícil no atribuir el efecto de la borrachera a la soda. Precisamente en una situación de este tipo es cuando la lógica del experimento resulta más útil: la flexibilidad de este método, la posibilidad de construir experimentos diferenciados entre sí por un solo elemento (en este caso, beber sólo soda), permiten emitir conclusiones sobre la causalidad que no serían posibles con el análisis de la covariación.

En conclusión, el experimento es un método de investigación social muy adecuado para el estudio de problemas específicos: análisis de grupos, dinámicas de interacción entre individuos, y todos los fenómenos limitados en cuanto a espacio, tiempo y número de personas implicadas. De ahí que el experimento se desarrolle sobre todo en el ámbito «micro» del análisis social, y que resulte muy útil en psicología y en psicología social.

4.8. SÍNTESIS DEL CAPÍTULO 4

1. Aunque la noción de causa-efecto es fundamental en el razonamiento científico, es una de las nociones más difíciles de traducir en términos operativos. Sin embargo, aunque la existencia de una ley causal nunca puede comprobarse empíricamente, el hecho de plantear una relación causal en el plano teórico, implica la observación de unos hechos. Estas observaciones no constituyen una prueba definitiva de la existencia del nexo causal, pero proporcionan una confirmación empírica de su existencia hipotética.
2. Para poder confirmar empíricamente la existencia de una relación causal entre dos variables, se precisan tres elementos empíricos: covariación entre variable independiente y variable dependiente, dirección del nexo causal y control de las variables extrañas.
3. Los científicos disponen de dos técnicas distintas para comprobar empíricamente las hipótesis causales: el *análisis de la covariación* y el *experimento*. En el primer caso, el investigador observa y analiza la relación entre las variaciones de X y las variaciones de Y en su desarrollo natural, excluyendo la influen-

cia de variables superfluas mediante la comparación entre grupos o el control estadístico. En el segundo caso, el investigador produce de manera artificial una variación de X en una situación controlada (la distribución aleatoria de los sujetos en grupos permite controlar las otras variables), y mide la consiguiente variación de Y.

4. Si queremos definir el efecto de la variable X sobre la variable Y, es decir, registrar las modalidades de Y cuando $X = t$ (tratamiento) y $X = c$ (control), no podemos hacerlo sobre la misma unidad y al mismo tiempo. Éste es «el problema fundamental de la deducción causal», que se puede resolver de dos maneras. La primera es la *solución científica*, que es aplicable cuando se cumple la tesis de la invariación (el valor de Y_c se sustituye por una medición de la misma Y_c registrada previamente sobre la misma unidad), o la tesis de equivalencia (el valor de Y_c se sustituye por una medición de la misma Y_c registrada sobre una unidad distinta pero equivalente). La segunda es la *solución estadística*, adecuada para las ciencias sociales, que no admiten la solución científica. La solución estadística es similar a la solución científica basada en la tesis equivalencia, con la diferencia de que utiliza dos grupos de individuos estadísticamente equivalentes, en lugar de dos unidades idénticas.
5. En las ciencias sociales los experimentos se dividen en *experimentos de laboratorio* y *experimentos de campo*. Los primeros se realizan en una situación artificial, los segundos en el contexto de la vida real. Los experimentos de laboratorio se clasifican en estudios de implicación, estudios de opinión y estudios de observación.
6. Un experimento debe tener dos características: manipulación de la variable independiente y distribución aleatoria de los sujetos en grupos (*aleatorización*). Si no se cumple la primera condición, no podemos hablar de experimento sino de estudio basado en el análisis de la covariación. Cuando no se cumple la segunda condición, hablamos de los denominados «cuasiexperimentos». Por tanto, podemos distinguir entre *experimentos verdaderos* y *cuasiexperimentos*.
7. El método experimental presenta dos ventajas. En primer lugar, permite abordar empíricamente el problema de la relación causa-efecto; en segundo lugar, permite aislar el fenómeno social objeto del estudio y centrar la observación en el mismo, eliminando la perturbaciones externas. Las limitaciones del método experimental son el carácter artificial de la situación (en cuanto al entorno y las reacciones de los sujetos implicados), y la escasa representatividad de las muestras estudiadas.

4.9. LECTURAS COMPLEMENTARIAS

Para una reflexión sobre el tema de la causalidad, su aplicación a la investigación social y la estadística, sugerimos el ensayo de J. H. Goldthorpe, *Causation, Statistics, and Sociology,* en Goldthorpe (2000).

Para el análisis de la relación causal cuando el diseño experimental no es viable, remitimos a la reseña (de planteamiento estadístico) de C. Winship y S. Morgan,

«The Estimation of Causal Effects from Observational Data» (*Annual Review of Sociology,* 1999).

Para una exposición clásica de los diseños experimentales y cuasiexperimentales recomendamos el ensayo clásico de D. T. Campbell y J. C. Stanley, *Experimental and Quasi-experimental Designs for Research* (Houghton Mifflin, 1966, 84 pp.). Existe una traducción del mismo al castellano. *Diseños experimentales y cuasi-experimentales en la investigación social* (1976). Buenos Aires: Amorrortu. Para una reflexión más profunda remitimos a la obra de T. D. Cook y D. T. Campbell, *Quasi-experimentation: Design and Analysis Issues for Field Settings* (Houghton Mifflin, 1979, 405 pp.).

Para una crítica del punto de vista clásico del experimento, con especial énfasis en el campo de la evaluación, véase el libro de R. Pawson y N. Tilley, *Realistic Evaluation* (Sage, 1997, 235 pp.).

En el texto de P. M. Sniderman y D. B. Grob, «Innovations in Experimental Design in Attitude Surveys», en *Annual Review of Sociology,* 1996, se presentan nuevas propuestas sobre la aplicación de la lógica del experimento a la encuesta por muestreo.

CAPÍTULO 5

LA ENCUESTA POR MUESTREO

LA ENCUESTA POR MUESTREO EN LA INVESTIGACIÓN SOCIAL 146

LA ESTANDARIZACIÓN O INVARIACIÓN DEL ESTÍMULO 150

5.2.1. Criterio objetivista y criterio constructivista 150
5.2.2. Enfoque uniformista y enfoque individualista 151
5.2.3. El objetivo del mínimo común denominador 152

. FIABILIDAD DEL COMPORTAMIENTO VERBAL 153

. FONDO Y FORMA DE LAS PREGUNTAS 156

5.4.1. Propiedades sociodemográficas, actitudes y comportamientos 157
5.4.2. Preguntas abiertas y preguntas cerradas 158

. FORMULACIÓN DE LAS PREGUNTAS 162

. BATERÍA DE PREGUNTAS 175

. MODALIDADES DE RECOPILACIÓN DE DATOS 178

5.7.1. Entrevistas cara a cara 179
5.7.2. Entrevistas telefónicas 181
5.7.3. Cuestionarios autocumplimentados 183
5.7.4. Entrevistas informatizadas 187

. FASES PREVIAS A LA RECOPILACIÓN DE DATOS 188

. ANÁLISIS SECUNDARIO Y ENCUESTAS SUCESIVAS 191

5.9.1. Análisis secundario 191
5.9.2. Encuestas sucesivas (diacrónicas) 194

. BASES DE DATOS EN ESPAÑA Y OTROS PAÍSES 196

5.10.1. Estudios transversales repetidos 197
5.10.2. Estudios longitudinales 201

. NUEVAS TENDENCIAS Y NUEVOS RIESGOS 202

. SÍNTESIS DEL CAPÍTULO 5 204

. LECTURAS COMPLEMENTARIAS 206

En este capítulo presentaremos la encuesta por muestreo, la técnica de investigación social cuantitativa más difundida, con sus aspectos teóricos y de aplicación. En primer lugar, discutiremos los problemas de fondo que plantea el hecho de «recopilar información mediante preguntas». Seguidamente abordaremos los problemas prácticos que surgen en el desarrollo de una investigación con esta técnica, desde la formulación del cuestionario hasta su aplicación y la recopilación de datos. También presentaremos de manera resumida, un catálogo de las principales bases de datos y archivos existentes en España, la Unión Europea y EEUU.

5.1. LA ENCUESTA POR MUESTREO EN LA INVESTIGACIÓN SOCIAL

En la vida cotidiana, cuando queremos conocer un determinado fenómeno social, ya sea individual (por ejemplo, la relación del paciente con el médico) o colectivo (por ejemplo, el comportamiento de la multitud en un estadio), disponemos de dos formas básicas de recopilar información: observar y preguntar. Si la observación es la vía más directa e inmediata para estudiar los comportamientos manifiestos, la interrogación es la vía obligada para explorar motivaciones, actitudes, creencias, sentimientos, percepciones y expectativas.

La estrategia de recopilar información preguntando directamente a los sujetos implicados en los fenómenos sociales estudiados abarca situaciones muy diversas, desde la encuesta periodística hasta el estudio de mercado, el sondeo electoral, el censo, etc. Conviene por tanto definir ese instrumento de investigación específico que hemos llamado «encuesta por muestreo».

Una encuesta por muestreo es un modo de *obtener información: a) preguntando, b) a los individuos que son objeto de la investigación, c) que forman parte de una muestra representativa, d) mediante un procedimiento estandarizado de cuestionario, e) con el fin de estudiar las relaciones existentes entre las variables.* A continuación, vamos a analizar los cinco elementos que componen esta definición. En primer lugar, la encuesta por muestreo conlleva la formulación de preguntas, que suelen ser orales (el encuestado las responde de forma oral), aunque a veces se presentan por escrito. Pueden ser encuestas cara a cara, entrevistas telefónicas o cumplimentación de fichas informativas: todas estas opciones tienen el denominador común de las preguntas.

Estas preguntas se plantean directamente a los individuos que son objeto de la investigación. En una investigación sobre la condición social de los presos, los entrevistados serán los propios presos; en una investigación sobre la afición al fútbol, los entrevistados serán los propios aficionados, etc. Desde este punto de vista el planteamiento de la encuesta por muestreo difiere de la investigación con entrevistas a los denominados «observadores privilegiados». Por ejemplo, la investigación sobre las cárceles podría realizarse entrevistando, no a los reclusos, sino a personas que, por su trabajo, tienen un conocimiento profundo del problema: directores de cárceles, jueces de vigilancia, asistentes sociales y psicólogos,

etc.; del mismo modo, una investigación sobre la afición al fútbol podría realizarse entrevistando a periodistas, miembros de grupos ultra, etc. Pero ya no sería lo que hemos denominado «encuesta por muestreo».

Dado que normalmente la población objeto del estudio está formada por un conjunto muy numeroso de sujetos, la imposibilidad de preguntarles a todos nos obliga a elegir una muestra sobre la que recopilar la información. En la definición anterior hemos dicho que esta muestra es «representativa», y eso significa que debe reproducir, a escala reducida, las características de la población objeto del estudio, para poder generalizar los resultados obtenidos en la muestra a la población total. Por lo tanto, los sujetos de la muestra se deben elegir siguiendo reglas muy precisas, establecidas por la estadística (hablaremos de ello en el capítulo 8), y el tamaño de la muestra debe ser considerable, por ello no incluimos en esta categoría los estudios exploratorios realizados sobre números limitados de sujetos.

Hemos dicho que el procedimiento para preguntar a los sujetos de la muestra debe estar estandarizado. Eso significa que a todos los sujetos se les plantean las mismas preguntas formuladas de idéntico modo. La estandarización del estímulo es una característica fundamental de la encuesta por muestreo, que permite comparar las respuestas y analizarlas con técnicas estadísticas. Si en una investigación sobre una muestra de jóvenes preguntamos a una parte de ellos qué sentimientos les inspira el himno nacional, y a otros qué sentimientos les inspira la bandera, no podemos saber si las diferencias en sus respuestas se deben a diferencias reales entre los entrevistados con respecto al sentido de identidad nacional, o si se deben al hecho de que los estímulos recibidos son distintos. Del mismo modo, si preguntamos a algunos individuos sobre sus prácticas religiosas y a otros no, nunca podremos analizar sobre la totalidad de la muestra la posible existencia de una relación entre, por ejemplo, prácticas religiosas y orientación política.

La entrevista a un individuo consta de dos elementos: la pregunta y la respuesta. Ambas se pueden expresar de forma estandarizada o de forma libre. Imaginemos una entrevista en la que a todos los sujetos se les plantea la misma pregunta, con idéntica formulación, por ejemplo: «¿le interesa a usted la política?», y éstos sólo pueden responder eligiendo una de las cuatro opciones preestablecidas, por ejemplo, mucho, bastante, poco o nada. En este caso, tanto la pregunta como la respuesta están estandarizadas. Otra posibilidad es plantear una pregunta estandarizada que el entrevistado puede responder libremente (por ejemplo, la respuesta a la pregunta anterior podría ser: «Depende, me interesa en campaña electoral porque todos hablan de ella y también tengo que decidir qué votar, pero fuera de estas ocasiones no me interesa en absoluto»). La tercera posibilidad es que la pregunta tampoco esté estandarizada, por ejemplo, cuando el entrevistador sabe que, en el curso de la entrevista, debe hablar del interés por la política, pero puede formular las preguntas como considere más oportuno, según el curso de la propia entrevista y las características del encuestado.

Combinando entre sí las características estandarizadas/libres de pregunta y respuesta, se obtiene la tipología de la figura 5.1, que da lugar a tres instrumentos de investigación distintos (el cuarto, en efecto, no existe):

Resumen 5.1. LA MATRIZ DE DATOS

Para poder analizar la evidencia empírica con técnicas estadísticas es necesario que tanto las preguntas como las respuestas estén estandarizadas, es decir, con un esquema de clasificación común a todos los sujetos. Eso significa que después de la recopilación de datos, obtendremos la denominada *matriz-de-datos,* que es la base de todos los cálculos estadísticos posteriores. Una vez que hemos registrado la misma información sobre todos los casos estudiados, podemos organizar esta información en una matriz rectangular de números, la denominada matriz de «observaciones por variables» ($O \times V$). En esta matriz, las observaciones se sitúan en las filas, y las variables se sitúan en las columnas, y en cada celda, que es el punto de intersección entre una fila y una columna, tendremos un *dato,* es decir, el valor que adopta una variable determinada para una observación concreta. Por ejemplo, una matriz de 200 (filas) × 50 (columnas) representaría 200 observaciones (pongamos que sean individuos) sobre las que se han registrado 50 variables (las mismas para todas las observaciones). En este ejemplo, cada fila de la matriz representa un individuo y cada columna de la matriz representa una variable.

— El *cuestionario,* cuando tanto la pregunta como la respuesta están estandarizadas.
— La *entrevista estructurada,* cuando sólo la pregunta está estandarizada, mientras que la respuesta es libre.
— La *entrevista libre,* cuando ni la pregunta ni la respuesta están estandarizadas.

En el primer caso, el entrevistador dispone de un texto escrito que reproduce la formulación exacta de las preguntas y, para cada una de ellas, la lista de posibles respuestas y su correspondiente codificación; en el segundo dispone de los textos de las preguntas, y debe atenerse a la formulación y al orden de los mismos; en el tercero sólo dispone de una guía para la entrevista, una lista de puntos y de temas que debe tratar, en un orden determinado y según la formulación que decide el propio entrevistador. Lo que cambia, por lo tanto, en las tres situaciones descritas, es el grado de libertad con el que se desarrolla la entrevista.

En este capítulo trataremos sólo el primer tipo, el cuestionario. Las entrevistas estructuradas y libres se estudiarán más adelante (en el capítulo 10), en la parte del libro dedicada a la investigación cualitativa.

La encuesta por muestreo tiene una larga historia en la investigación social. Marx y Weber ya aplicaron la idea de obtener información pidiéndola directamente

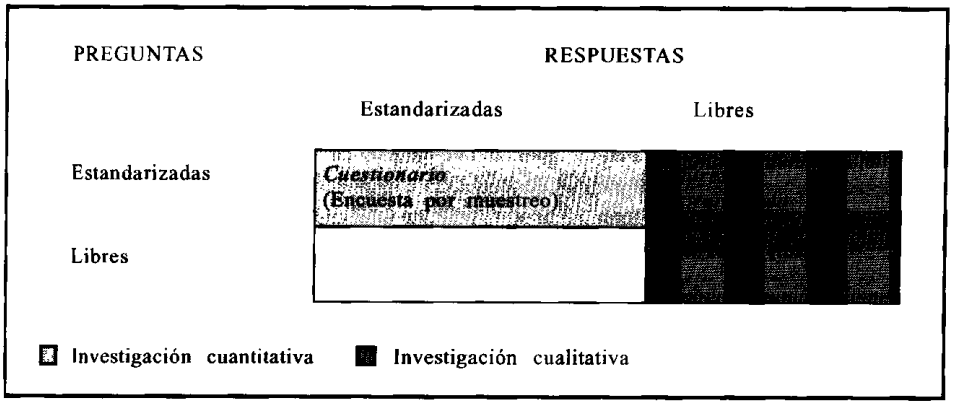


Figura 5.1. Instrumentos de investigación mediante preguntas.

a los propios sujetos estudiados. En 1880 Marx envió 25.000 copias de un cuestionario a los lectores de la *Révue socialiste,* con preguntas abiertas sobre sus condiciones de vida (sin embargo, recibió tan pocas respuestas que ni siquiera llegó a analizar los datos [cfr. Bottomore y Rubel, 1956]).

Weber también recurrió varias veces a los cuestionarios entre 1880 y 1910 para estudiar problemas sociales, como las condiciones de trabajo en los campos de Prusia oriental (con cuestionarios enviados por correo a propietarios agrícolas y a pastores protestantes), y los efectos de trabajar en la industria pesada sobre la estructura de la personalidad y la forma de vida de los obreros (con cuestionarios planteados tanto a observadores privilegiados como a una muestra de obreros [cfr. Lazarsfeld y Oberschall, 1965]). Podemos citar también las encuestas que se realizaron sobre el impulso de los movimientos de reforma social que tuvieron lugar entre los siglos XIX y XX, sobre todo en Inglaterra, con el objetivo de recopilar información sobre las condiciones de vida de los pobres, los obreros y los marginados. Estos primeros estudios partían de la premisa de que hasta que no se midiera la naturaleza y el alcance de estos problemas, poco podía hacerse para solucionarlos.

Pero el verdadero salto cualitativo en el ámbito de las técnicas de encuesta se produjo con la consolidación del concepto de representatividad y la introducción de los procedimientos de muestreo. Es decir, cuando se afianzó el principio de que para conocer la distribución de un número determinado de variables en una población concreta, no era necesario estudiar a toda la población, sino que un estudio de una muestra seleccionada convenientemente podía producir resultados igual de exactos que los obtenidos por un estudio realizado sobre el total de la población (o incluso más exactos, dado que los recursos no utilizados gracias a la reducción del alcance de la investigación podían emplearse para mejorar la calidad de la misma). Este principio, sistematizado teóricamente en los textos del estadístico polaco Jerzy Neyman hacia mediados de los años treinta, se consolidó rápidamente y allanó el camino para el extraordinario éxito que tendría la técnica de la encuesta por muestreo.

5.2. LA ESTANDARIZACIÓN O INVARIACIÓN DEL ESTÍMULO

Los problemas fundamentales a los que se enfrenta el investigador que ha decidido utilizar la encuesta por muestreo guardan relación con las diferencias de base (ontológicas, epistemológicas y metodológicas) existentes entre los dos paradigmas básicos de la investigación social: el planteamiento de la tradición positivista (con sus variantes neopostivista y postpositivista), y el planteamiento que hemos denominado interpretativo. A continuación comentaremos dos dilemas relacionados con estos dos planteamientos diferentes.

5.2.1. Criterio objetivista y criterio constructivista

El primer dilema enfrenta a quienes consideran que la realidad social existe fuera del investigador y éste puede conocerla de manera plena y objetiva, y quienes consideran que el propio acto de conocer la realidad conlleva una alteración de la misma. Esta alteración hace que la realidad sólo pueda ser la generada por el proceso interactivo que se establece entre el sujeto que estudia y el sujeto estudiado. Tenemos, pues, por una parte, la posición que podríamos definir como *objetivista* (el dato social se puede registrar objetivamente, con un procedimiento no muy distinto al de la «observación» en las ciencias naturales), y, por otra, la posición que podríamos definir como *constructivista,* para la cual el dato social no es observado, recogido o registrado, sino «construido» o «generado» por la interacción entre el sujeto que estudia y el sujeto estudiado.

Si trasladamos este dilema al terreno de la encuesta por muestreo, nos encontramos con la cuestión de la relación entre el entrevistado y el entrevistador, contraponiendo la investigación aséptica e impersonal con la interacción empática entre el investigador y el sujeto. Según el criterio objetivista, la relación entre el investigador y el sujeto debe ser impersonal. El entrevistador debe procurar ante todo no alterar el estado del sujeto estudiado. La interacción con este sujeto, que resulta innegable incluso para los defensores de la posición objetivista, se considera un hecho negativo, un mal necesario, que conviene limitar lo más posible. Por ello se han elaborado códigos de conducta a los que debe atenerse el entrevistador, orientados a garantizar la neutralidad y uniformidad de la relación. En una edición antigua del *Interviewer's Manual,* del instituto de encuestas más conocido de Estados Unidos, que durante décadas ha hecho las veces de escuela en este campo, se puede leer:

> Procure que nada en sus palabras o en su comportamiento denote crítica, sorpresa, aprobación o desaprobación [...], emplee un tono de voz normal, una forma de escuchar atenta, una actitud que no revele juicio alguno [...]. Si el encuestado le pide explicaciones sobre una palabra o sobre una frase, evite ayudarle y remita la responsabilidad de la definición al propio entrevistado [...]. Si por ejemplo, le pregunta [...] qué quiere decir con «discriminación», responda «exactamente lo que significa para usted» [...], y si insiste diciendo «no entiendo», pase a la siguiente pregunta.
>
> [Isr 1976, 11-13]

Eso no significa, según los defensores de este planteamiento, expresado en algunos manuales de investigación clásicos de los años cincuenta y sesenta, que el entrevistador deba ser frío y distante. Al contrario: «debe hacer entender al encuestado que lo está haciendo bien» [*ibidem,* 13]; debe «mostrar un interés vivo y empático por los problemas de la persona con la que habla [...], debe acercarse al entrevistado y mostrarle confianza» [Goode y Hatt, 1952, 190]. Pero al mismo tiempo no debe excederse en su familiaridad. Según Hyman: «Es preferible un cierto grado de formalidad burocrática, de distancia social. Cuando la relación supera un cierto límite, se hace demasiado íntima y el entrevistado puede estar tentado de cumplir las expectativas del entrevistador [...], sobre todo en el caso de que conozca poco el tema en cuestión [...] y no tenga una opinión personal propia» [Hyman, 1954, 48].

Es lógico y obvio que estas prescripciones a favor de lo que podríamos llamar «conversaciones de ascensor» hayan suscitado no pocas críticas. La asepsia de la relación es un mito carente de fundamento. En lo que concierne al entrevistador, éste «no sólo está en el mundo, sino que forma parte del mundo», tiene reacciones, percepciones, puntos de vista y esquemas mentales propios que, para bien o para mal, siempre lo condicionarán y harán imposible el papel de «registrador» neutral. En cuanto al entrevistado, existe una reactividad en el sujeto estudiado, que desde el momento en que se convierte en objeto de investigación sufre una alteración, ya no es igual a sí mismo, se siente sometido a examen, quiere quedar bien, está emocionado o irritado, contrariado o satisfecho, etc. La relación cognitiva ya no puede ser, por tanto, de observación-investigación, sino que se convierte inevitablemente en una interacción. El entrevistador no se limita a registrar la respuesta de su interlocutor, sino que participa directamente en la producción de la misma.

5.2.2. Enfoque uniformista y enfoque individualista

El segundo dilema enfrenta a las posiciones que podríamos llamar *uniformista* e *individualista.* El punto de vista uniformista considera que existen, si no leyes como las que gobiernan el mundo físico, al menos ciertas uniformidades empíricas, ciertas regularidades en los fenómenos sociales y en los comportamientos humanos, que permiten clasificarlos y estandarizarlos. La perspectiva individualista, por su parte, destaca la importancia de las diferencias entre individuos, la irreductibilidad fundamental del sujeto humano a cualquier forma de generalización y estandarización, y la posibilidad de ser comprendido por el investigador sólo en la medida en que éste consiga establecer una relación empática con aquél. En resumen, la primera postura sostiene que existen ciertas regularidades que permiten estandarizar los actos de los sujetos individuales, mientras que la segunda considera que cada caso, cada acto social, es un acontecimiento único.

Este problema nos conduce al tema de la *estandarización del instrumento de investigación.* El cuestionario limita al entrevistador, que debe plantear las mismas preguntas formuladas de manera idéntica a todos los sujetos, y cuando plantea respuestas cerradas también limita al entrevistado, que está obligado a escoger entre las respuestas preestablecidas. En el manual del entrevistador antes citado también podemos leer lo siguiente: «No modifique la formulación de las preguntas. Evite no

sólo los cambios deliberados, sino también los involuntarios que pasan inadvertidos [...]. Podría suceder que en un esfuerzo por hacer más fluida la conversación añada alguna palabra [...]. Varios experimentos han mostrado que incluso la más ligera modificación en las palabras puede influir en la respuesta» [Isr, 1976, 11].

Según los críticos de este enfoque, el cuestionario o esquema estandarizado de investigación, tiene dos inconvenientes graves. En primer lugar, el cuestionario es el mismo para todos los sujetos y se aplica a todos ellos del mismo modo, como si todos los destinatarios tuvieran la misma sensibilidad, rapidez y madurez; ignora la diversidad de los individuos en la sociedad. En segundo lugar, el cuestionario obliga al sujeto a situarse en un nivel determinado, con lo cual el sujeto no puede informar sobre la realidad en la que vive fuera de ese nivel impuesto por el cuestionario. En otras palabras, el cuestionario uniforma a los individuos, los sitúa al nivel del hombre medio. Además, como ha escrito Galtung, la encuesta por muestreo excluye a los grupos sociales periféricos: «los analfabetos, los ancianos, los aislados, los indigentes, los vagabundos, quienes se encuentran fuera de la geografía del país» [Galtung, 1967, 154]. Éstos grupos quedan excluidos porque no aparecen en las listas del censo, o porque no consienten ser entrevistados, o porque no entienden las preguntas o no se identifican con las opciones de respuesta previstas por el investigador para el ciudadano medio.

5.2.3. El objetivo del mínimo común denominador

El objetivo fundamental de la postura *objetivista-uniformista* es evidente. Las soluciones que se dan al primer dilema (despersonalizar la relación entrevistado-entrevistador) y al segundo (estandarizar las preguntas y las respuestas), llevan al mismo punto: la neutralidad del instrumento de investigación (del que también forma parte el entrevistador), es decir, utilizando un lenguaje conductista, la *invariación del estímulo*. Esto se hace para asegurar que las respuestas son comparables entre sí. Las respuestas se consideran comparables porque todos los encuestados han contestado a las mismas preguntas y han sido entrevistados en situaciones prácticamente idénticas.

Pero, ¿la invariación formal del estímulo, suponiendo que se pueda conseguir, debe implicar igualdad de condiciones de todos los entrevistados en relación con la pregunta planteada? Los autores de uno de los manuales de metodología más difundidos en los años setenta en Estados Unidos escriben lo siguiente: «¿Nos permite la naturaleza impersonal del cuestionario afirmar que existe uniformidad en la medición entre un caso y otro? Desde el punto de vista psicológico esta uniformidad puede ser más aparente que real; la misma pregunta estándar puede tener distintos significados para las distintas personas: para unos puede ser conocida, para otros totalmente incomprensible» [Selltiz, Wrightsman y Cook, 1976, 295]. En pocas palabras, ¿quién nos garantiza que la *uniformidad de los estímulos* se corresponda con una *uniformidad de los significados*?

En este punto se plantea la cuestión del significado que una misma pregunta, o una misma palabra, puede tener para los distintos individuos. Este problema ha sido ampliamente estudiado en psicología cognitiva. La pregunta planteada por el entre-

vistador no cae en un terreno aséptico, neutral y común a todos los encuestados, sino en un terreno que varía entre los distintos sujetos, en el que se funden la biografía del entrevistado, su cultura y sus experiencias anteriores; además, el significado de una misma expresión está muy influido por las circunstancias en las que se desarrolla la entrevista, así como por la forma en que el entrevistador se presenta y es percibido.

Cada individuo tiene un sistema de referencia propio y al recibir un estímulo lo lee y lo interpreta a la luz de sus experiencias anteriores. El lenguaje común de la pregunta estandarizada no garantiza que tenga el mismo significado para todos los entrevistados. Como escribe Cicourel [1964, 108], «las preguntas estandarizadas con respuesta cerrada ofrecen una solución para el problema del significado: ignorarlo».

¿Cómo podemos resolver esta situación? El dilema al que se enfrenta el investigador que ha decidido conocer la realidad social preguntando a sus protagonistas, surge de la oposición entre los dos paradigmas fundamentales que inspiran la investigación social, y se traduce en la elección entre una técnica de interrogación que maximiza la estandarización y la búsqueda de uniformidad, y una técnica que maximiza la individualidad del sujeto estudiado y su relación con el sujeto que estudia. En otras palabras, el investigador debe escoger entre cuestionario y entrevista no estructurada (o entrevista en profundidad).

Por tanto, en el momento en que el investigador elige el cuestionario y la encuesta por muestreo como instrumentos para recopilar información sobre los sujetos, toma una decisión importante: da prioridad a la búsqueda de uniformidad por encima de la búsqueda de individualidad, a la búsqueda de aquello que tienen en común los individuos por encima de aquello que los diferencia. Con esta elección, el investigador decide limitarse a estudiar ese mínimo común denominador en el comportamiento que se puede uniformar, clasificar y comparar, a pesar de la individualidad de los sujetos, y que se puede registrar, a pesar de la variabilidad de las dinámicas entrevistado-entrevistador. El *mínimo* común denominador puede parecer limitado en relación con el objetivo de alcanzar la *máxima* comprensión del comportamiento humano. Sin duda lo es, y constituye una de las limitaciones de la investigación cuantitativa, que, sin embargo, opta conscientemente por trabajar con números grandes, en lugar de profundizar trabajando con números pequeños.

5.3. FIABILIDAD DEL COMPORTAMIENTO VERBAL

Existe una segunda objeción de fondo, que es menos radical que la anterior porque no concierne a los fundamentos epistemológicos de la encuesta por muestreo, sino a su viabilidad empírica, pero que llevada al extremo cuestiona la razón de ser de esta técnica. Dicha objeción puede expresarse con la siguiente pregunta: «¿El comportamiento verbal es una fuente fiable para el estudio de la realidad social?».

Muchos científicos sociales han expresado su pesimismo sobre la posibilidad de comprender plenamente la realidad social a partir de las respuestas obtenidas preguntando a las personas. Hace tiempo que los antropólogos abandonaron la idea de

que las respuestas a las preguntas constituyen evidencia estable en el estudio del comportamiento humano, después de aprender que lo que las personas afirman que han hecho o harán no se corresponde con su comportamiento real.

Analizaremos el problema de la fiabilidad del comportamiento verbal desde dos puntos de vista. El primero tiene que ver con el principio de que *las respuestas sean socialmente deseables*. El segundo concierne a la *falta de opiniones* (o las *pseudo-opiniones*).

Respuestas socialmente deseables. En cada cultura, la sociedad atribuye una valoración concreta a una actitud o conducta individual determinada. Existen ciertos atributos o comportamientos individuales que son juzgados negativamente por las normas colectivas de una sociedad determinada (por ejemplo, la pobreza, el alcoholismo o la drogadicción, las relaciones sexuales extramatrimoniales, etc.) y otros que, en cambio, son valorados positivamente (por ejemplo, la honestidad, la laboriosidad, la religiosidad, etc.). Si una actitud (o una conducta) tiene una connotación positiva o negativa fuerte en una cultura determinada, cualquier pregunta sobre la misma puede dar lugar a respuestas muy distorsionadas, ya que el encuestado puede ser reacio a revelar opiniones o conductas que considera inconfesables y estar tentado de dar la mejor imagen posible de sí mismo, aunque sea poco veraz.

La literatura sociológica ofrece numerosos ejemplos de preguntas afectadas de distorsión sistemática según si las respuestas son más o menos deseables socialmente. Sudman y Bradburn citan cuatro casos en los que el investigador podía controlar el comportamiento real gracias a la existencia de datos oficiales registrados. Los comportamientos considerados fueron los siguientes: haber votado en las elecciones, ir habitualmente a la biblioteca (tener un carné de usuario), haber sido condenado por un delito de quiebra fraudulenta, y haber sido multado por conducir en estado de embriaguez [1].

Para los dos primeros comportamientos, socialmente deseables, en las entrevistas se obtuvo un porcentaje de respuestas positivas superior a los datos reales, respectivamente de 40 y 20 puntos porcentuales más. Para los otros dos comportamientos, socialmente reprobables, el resultado fue el contrario: entre un tercio y la mitad de las personas condenadas por quiebra fraudulenta o multadas por conducir en estado de embriaguez no declaró la verdad en las entrevistas. En Italia también se realizó un experimento similar a éste, encuestando a personas que según las listas electorales no habían ido a votar en las últimas elecciones: el 55 por 100 de ellos declaró que había votado. Conviene subrayar que el porcentaje más elevado de respuestas falsas correspondió a individuos con estudios superiores, es decir, aquellos que eran más conscientes de que su conducta era reprobable [lc. en Pitrone 1984, 127].

---

[1] Los controles del comportamiento real se hacían *a priori* (se entrevistaba a personas de listas judiciales de condenados por quiebra fraudulenta o a las que se habían impuesto multas por conducir en estado de embriaguez, sin que supieran que el investigador conocía sus antecedentes en cuestión) o *a posteriori* (después de la entrevista se controlaba en las listas electorales si las personas realmente habían ido a votar, o si estaban registradas en la biblioteca indicada por ellos).

Falta de opiniones. Un segundo problema del estudio de las actitudes mediante preguntas a los sujetos es lo que en inglés se denomina *non-attitudes* y que podríamos traducir como *falta de opiniones.* En las encuestas de investigación social se suelen plantear preguntas sobre temas complejos: si es mejor que el gobierno intervenga en la economía del país o si es preferible un liberalismo económico total; si es mejor aumentar los impuestos para ofrecer mejores servicios sociales o si conviene más reducir ambos, etc. Estas preguntas se suelen presentar en forma de «batería», es decir, una serie de afirmaciones para que el encuestado indique si está de acuerdo o no con cada una de ellas.

Puede ser que un cierto número de entrevistados no haya reflexionado nunca sobre algunos de estos temas y que, por tanto, no tenga una opinión al respecto. Sin embargo, la dinámica de la entrevista conlleva una especie de *presión por responder,* por la cual muchos entrevistados *eligen al azar* una de las respuestas posibles, debido también a la formulación de la «pregunta cerrada» [2]. En el mejor de los casos puede ocurrir que la opinión nazca en el momento mismo en que se plantea la pregunta (lo que hace que la respuesta sea sumamente volátil).

Aunque este problema se conoce desde hace tiempo, los investigadores sociales empezaron a prestarle atención a raíz de un célebre ensayo de Philip Converse publicado en 1970. Al estudiar los resultados de una encuesta longitudinal realizada entrevistando a las mismas personas en varios momentos distintos, Converse advirtió ciertos hechos inexplicables (como la baja correlación entre las respuestas dadas a la misma pregunta por los mismos individuos en dos encuestas sucesivas). Estos hechos sugerían que los entrevistados podían clasificarse en dos categorías muy distintas: las personas con respuestas estables y las que responden al azar.

Los primeros son individuos perfectamente coherentes en sus respuestas, que no varían con el paso del tiempo; los segundos son individuos «sin una opinión real sobre el tema en cuestión, pero que por alguna razón se sienten obligados a responder» y lo hacen al azar. «Al margen de nuestras intenciones, el cuestionario se percibe como un test de inteligencia, en el que las respuestas «No sé» o «No sé qué respuesta escoger» se consideran una confesión de incapacidad mental» [Converse, 1970, 175, 177]. De ahí que elijan al azar entre las distintas opciones de respuesta disponibles. Converse acuñó el término *non-attitude* para indicar esta falta de opinión.

El problema planteado por Converse ha sido estudiado empíricamente de otra manera, que consiste en plantear a una muestra de entrevistados una serie de preguntas sobre hechos inexistentes, con el objetivo de distinguir a los individuos que responden aunque no tengan una opinión al respecto (sería el caso de quienes dan una respuesta cualquiera en lugar de contestar «no sé»). Los ejemplos de este tipo de estudios son numerosos, pero aquí citamos sólo dos experimentos de Schuman y Presser [1981, 148 sq.], en los que se preguntó a los entrevistados si estaban a favor o en contra de dos medidas gubernamentales ficticias, una relacionada con la agri-

[2] En la pregunta cerrada el encuestado escoge la respuesta que mejor refleja su opinión, entre una serie de opciones propuestas por el entrevistador (ver también el apartado 5.3.2).

cultura, y otra concerniente al control monetario. También citamos una serie de experimentos similares de Bishop, Tuchfarber y Oldendick [1986], que formularon la pregunta sobre leyes inventadas de varios modos, incluyendo explícitamente en algunas formulaciones la opción de «no sé» entre las respuestas. En todos los experimentos, aproximadamente un tercio de las respuestas fueron «estoy a favor» o «estoy en contra» de la medida gubernamental inexistente, mientras que sólo dos tercios de las respuestas fueron «no sé», que en este caso era la única opción correcta. Esta evidencia pone en tela de juicio la fiabilidad de la técnica.

Llegados a este punto cabe plantearse otra cuestión, relacionada con la anterior: la pregunta estandarizada registra una opinión, pero no su *intensidad* ni su nivel de *convencimiento*. Una pregunta normal de cuestionario, en la que se plantea una afirmación a los entrevistados y se les pide que digan si están de acuerdo o no con ella, produce un cierto número de respuestas positivas y negativas de los encuestados. Pues bien, el investigador no puede distinguir cuáles de esas respuestas están basadas en opiniones profundamente arraigadas y sólidas y cuáles son simplemente opiniones superficiales, que podrían haber nacido en el mismo momento de plantearse la pregunta. El sociólogo está obligado a dar la misma importancia a las opiniones pasajeras o impresiones superficiales que pueden cambiar de un día para otro, y las opiniones consolidadas y enraizadas en la biografía del entrevistado.

Es evidente que existe una clara diferencia entre una opinión, por ejemplo, contraria al aborto, expresada de manera objetiva y desde una perspectiva puramente intelectual, y una opinión afectada por una implicación emocional que conlleve la formulación de manifestaciones antiabortistas. De manera similar las repercusiones conductuales de una misma opinión serán muy diferentes según el grado de implicación emocional del individuo con respecto a esa opinión.

La técnica del cuestionario no permite distinguir las opiniones fuertes y estables de las débiles y volubles. Cuando decimos que el 25 % ha declarado que votará a un determinado candidato, no sabemos en qué medida esta orientación es sólida y fiable y en qué medida podrá cambiar con la primera aparición televisiva del candidato seleccionado o de su contrincante. Desde este punto de vista podemos compartir la opinión de la periodista italiana Barbara Spinelli cuando afirma que «los sondeos [...] son esclavos de los sentimientos, de los humores momentáneos en vías de descomposición, de la sombra que roza repentinamente la frente cuando pasa una nube» [*La Stampa,* 1 de agosto 1994].

5.4. FONDO Y FORMA DE LAS PREGUNTAS

Después de esta reflexión general sobre los problemas de la recogida de datos mediante preguntas directas a los individuos que son objeto de la investigación, entramos ahora de lleno en los instrumentos empleados para tal fin. En este apartado trataremos el contenido de las preguntas y la distinción entre preguntas abiertas y preguntas cerradas.

Podría pensarse que un cuestionario es una secuencia banal y simple de preguntas y que su redacción es una operación trivial. Sin embargo, la formulación de un

cuestionario es una labor complicada y difícil, que requiere mucho empeño y atención. Se trata además de un trabajo que resulta difícil de formalizar y para el que no existen reglas precisas; como mucho se pueden hacer sugerencias derivadas de la experiencia de estudios previos realizados específicamente sobre esta cuestión.

En general, para la redacción de un buen cuestionario se precisa: *a*) experiencia del investigador; *b*) conocimiento de la población a la que va destinado el cuestionario, y *c*) claridad de las hipótesis de investigación. Para los investigadores con poca experiencia el primer punto se podría sustituir por la consulta de cuestionarios utilizados anteriormente por otros investigadores sobre el mismo tema o temas similares. Como recuerdan Sudman y Bradburn [1982, 14] en uno de los manuales más conocidos sobre la formulación del cuestionario, copiar preguntas de otros cuestionarios no es plagio, al contrario, se trata de una práctica recomendable en el ámbito científico, ya que permite acumular conocimientos y hacer comparaciones a lo largo del tiempo.

5.4.1. Propiedades sociodemográficas, actitudes y comportamientos

Empezamos con el tema de las preguntas, para el cual los autores han propuesto diversas clasificaciones. Con algunas variaciones, éstas se pueden reducir en última instancia a tres bloques: propiedades sociodemográficas básicas, actitudes y comportamientos.

Preguntas relativas a propiedades sociodemográficas básicas. Se refieren a la descripción de las características sociales básicas del individuo, e incluyen sus características permanentes, como las demográficas (género, edad, lugar de nacimiento, etc.) y las connotaciones sociales heredadas de la familia o, en todo caso, estabilizadas en la edad juvenil (como la clase social de origen y los estudios), y también las características que, aunque no sean permanentes en sentido estricto, definen al individuo en un periodo concreto de su vida, como la profesión, el estado civil, la situación geográfica y el tamaño del municipio de residencia, etc. Casi todas estas preguntas se plantean en todas las encuestas por muestreo (sea cual sea el objeto de la investigación) y para ellas existen formulaciones estándar que conviene adoptar.

Preguntas relativas a actitudes. Se refieren a opiniones, motivaciones, orientaciones, sentimientos, juicios, valores. Son las propiedades de los individuos que se suelen observar mejor con la encuesta por muestreo: la pregunta directa al individuo parece de hecho la única vía practicable para poder conocer sus pensamientos. También es el campo más difícil de explorar: es más fácil responder sobre una cuestión cuando ha generado una conducta que cuando se trata estrictamente de una opinión. Por ejemplo, es más fácil preguntar si se ha votado o no al referéndum sobre el aborto, que preguntar si se está a favor o en contra del aborto, sin un comportamiento consecuente como el voto en el referéndum.

Las actitudes suelen ser complejas y multidimensionales: el sujeto en cuestión puede estar a favor del aborto en determinadas situaciones y en contra del mismo en otras, o puede no tener una opinión definida al respecto. Además, con frecuencia las motivaciones de nuestros actos son inconscientes para nosotros mismos (alguien pue-

de ir con regularidad a la iglesia por motivos que ni siquiera conoce). Por último, las actitudes varían mucho entre las personas en cuanto a la intensidad, y este aspecto no es fácil de registrar con un cuestionario. Por estas razones, y sobre todo por esa intrínseca complejidad y ambigüedad que hemos mencionado, las preguntas sobre actitudes son las más difíciles de formular y las respuestas pueden verse influidas fácilmente por la formulación de la pregunta, por su posición en el cuestionario, por la manera en que las plantea el entrevistador, etc.

Preguntas relativas a comportamientos. Estas preguntas pretenden registrar lo que el entrevistado dice que hace o ha hecho. Se trata, por tanto, del terreno de las «acciones», un terreno mucho más sólido que el anterior, al menos por dos razones. En primer lugar, los comportamientos, a diferencia de las actitudes y los estados psicológicos y mentales, son *inequívocos.* Una acción puede haberse producido o no, y, por tanto, existe una respuesta terminante para la pregunta relativa a ella, por contraste con las actitudes, que muchas veces no son claras ni siquiera para el propio sujeto que las adopta. En segundo lugar, los comportamientos con *observables.* Una acción puede ser observada por una tercera persona y dejar una huella objetiva (frente a la subjetividad propia de las actitudes): si una persona ha secundado una huelga, su comportamiento es conocido por sus compañeros de trabajo; si ha votado, el hecho consta en las listas electorales, etc. Se pueden comprobar empíricamente, aunque ello no implica que las respuestas se comprobadas efectivamente. De este modo, la existencia de una respuesta inequívoca hace que sea más fácil responder a las preguntas sobre comportamientos, y la posibilidad de comprobar su veracidad hace que sea más difícil mentir que en el caso de las preguntas sobre actitudes.

5.4.2. Preguntas abiertas y preguntas cerradas

Una de las primeras decisiones que debe tomar el investigador en el momento de formular una pregunta es si la plantea como pregunta abierta o cerrada (en realidad deberíamos llamarla «pregunta a respuesta abierta» y «pregunta a respuesta cerrada», pero hoy en día es más común el uso de la forma abreviada). La pregunta abierta es aquella que el entrevistador plantea dejando libertad total al entrevistado para la formulación de la respuesta (que el entrevistador transcribe íntegramente). La pregunta cerrada, en cambio, es aquella en la que el entrevistado dispone de una serie de posibles respuestas entre las que debe elegir la que considere más apropiada (y el entrevistador se limita a marcar con una cruz la alternativa seleccionada).

Como ejemplo consideremos una pregunta planteada frecuentemente en los cuestionarios sociológicos y veamos sus dos variantes. Pregunta abierta: En su opinión, «¿cuáles son los problemas más graves a los que se enfrenta nuestro país en la actualidad?». Pregunta cerrada: «Le voy a leer una lista de problemas [...] (paro, inflación, criminalidad, inmigración, etc.). En su opinión, de los problemas mencionados, ¿cuáles son los más graves en nuestro país en la actualidad?».

Es evidente que la elección entre las dos formas de plantear la pregunta está directamente relacionada con la problemática debatida anteriormente sobre la estandarización de la encuesta. Se trata de decidir si, además de estandarizar la pregunta, queremos estandarizar también las respuestas a priori.

Analicemos primero la *pregunta abierta.* Ésta tiene la ventaja de ofrecer libertad de expresión. La diferencia fundamental entre los dos tipos reside en la espontaneidad de la respuesta frente a la limitación de las respuestas precodificadas. Pero, ¿estamos seguros de que no hay nada más? En el caso de la pregunta cerrada, con la elección de la respuesta entre las opciones propuestas se alcanza el objetivo de la encuesta en la investigación cuantitativa, es decir, la matriz de datos (clasificación de las respuestas en un esquema estandarizado comparable entre todos los entrevistados). En cambio, con la pregunta abierta, después de trascribir textualmente la respuesta del entrevistado, aún nos falta una fase antes de llegar a la matriz de datos, la denominada «codificación» de las respuestas, es decir, la clasificación en un número finito y limitado de categorías, para hacer posibles las comparaciones y los análisis cuantitativos posteriores. En resumen: las respuestas a las preguntas abiertas tienen que ser codificables, o sea, estandarizadas, para poder procesarlas en la matriz de datos.

No obstante, el proceso de codificación conlleva un alto grado de arbitrariedad por parte del codificador, que puede llegar a forzar la respuesta original, de manera similar a como lo hace a priori la pregunta cerrada. Efectivamente, las respuestas abiertas son a menudo contradictorias, incomprensibles, confusas, genéricas y ambiguas.

Tomemos como ejemplo la pregunta anterior sobre los «problemas más importantes del país». ¿Cómo clasificaríamos una respuesta como «el Estado»? ¿Habrá querido referirse el encuestado a la ineficacia de la administración pública, o al centralismo, porque está a favor de la descentralización, o a la corrupción de los cargos públicos?

Consideremos otra pregunta: «¿Cuáles son las características que considera más importantes de un trabajo?». Supongamos que la respuesta es abierta. ¿Cómo se debe clasificar una respuesta como «un trabajo que guste»? ¿O respuestas demasiado generales como «buenas condiciones de trabajo» (que puede implicar muchas cosas, como sueldo, posibilidades de ascenso, seguridad del puesto, horario [...])? ¿O respuestas que parten de un marco de referencia distinto y mencionan no ya una característica específica, sino un tipo de trabajo, como «un trabajo de funcionario»?

Todo cambia si el entrevistador interviene en la encuesta a modo de orientador. Sudman y Bradburn [1982, 150] citan el ejemplo siguiente:

> Pregunta: En su opinión, ¿cuáles son los problemas más importantes de nuestro país en la actualidad?
> Respuesta: Pues no sé, hay tantos.
> P.: De acuerdo, pero querría saber cuáles son los que usted considera más importantes.
> R.: Bien, desde luego la inflación y, después, los gastos del Estado.
> P.: Los gastos del Estado... ¿Qué quiere decir con eso? ¿Puede explicarlo mejor? ¿A qué se refiere cuando dice «los gastos del Estado»?
> R.: No acaban nunca, tenemos que reducir de algún modo los gastos estatales, hay demasiado despilfarro.
> P.: ¿Nada más?
> R.: No, creo que éstos son los problemas más importantes.

Con la orientación de un buen entrevistador, la pregunta abierta siempre produce un resultado no ambiguo, que está dentro del marco de referencia deseado por el investigador. Pero esta forma de proceder tiene un elevado coste y no es viable con muestras grandes. Una encuesta por muestreo normal implica al menos 1000 entrevistas y, por consiguiente, varias decenas de entrevistadores. Si la entrevista es cara a cara [3] y la muestra abarca todo el país, éstos deberán desplazarse a regiones distantes, por lo que su control no será fácil. Circourel [1964, 107] se pregunta cómo podemos «controlar que todos los entrevistadores [...] utilizan el mismo marco teórico de referencia e interpretan cada hecho, persona entrevistada, etc. de forma idéntica, es decir, empleando las mismas estructuras de significado en contextos diferentes con las mismas reglas interpretativas» [4].

Por tanto, debido a motivos fundamentalmente prácticos, de costes y de viabilidad, la pregunta abierta se ha ido abandonando progresivamente en las encuestas por muestreo. No obstante, se sigue utilizando, como veremos más adelante, en investigaciones sobre muestras de dimensiones reducidas.

La *pregunta cerrada* presenta algunas *ventajas* que constituyen, como es natural, la otra cara de la moneda en relación con lo que acabamos de decir sobre la pregunta abierta. A continuación presentamos una breve lista de estas ventajas:

a) La pregunta cerrada ofrece el mismo marco de referencia para todos los entrevistados. Pensemos en la siguiente pregunta: «¿Cuáles son, según su opinión, las cosas más importantes que deben aprender los niños para estar preparados para la vida?». La pregunta se utilizó en un experimento en sus dos variantes: abierta y cerrada. Las opciones propuestas en la pregunta cerrada eran: obediencia, laboriosidad, responsabilidad, ayuda a los demás, etc., es decir, una serie de virtudes-valores. En las respuestas abiertas la más frecuente fue «sacar buenas notas», respuesta que obviamente queda fuera del marco de referencia previsto por los entrevistadores (que pretendían identificar los valores de referencia de los entrevistados) [Schuman y Presser, 1981, 109]. En otras palabras, frente al hecho de que una misma

---

[3] Si la entrevista es telefónica, el número de entrevistadores será menor y, al estar centralizados, será más fácil controlarlos. Ésta es una de las ventajas de esta técnica. Sin embargo, la rígida limitación de los tiempos inherente a la misma, hace imposible usarla para entrevistas en profundidad y orientadas por el entrevistador, como la presentada anteriormente.

[4] Una solución que parece combinar las ventajas de la pregunta abierta y la cerrada podría ser la llamada «pregunta codificada en el campo», que consiste en preguntas presentadas al entrevistado como preguntas abiertas, pero que el entrevistador codifica en el momento en que se formula la respuesta, en categorías preformuladas (que el entrevistador tiene en el cuestionario pero que el entrevistado no conoce). A pesar de sus aparentes ventajas, quienes han estudiado el problema a fondo recomiendan evitar esta solución. En efecto, existe «la ventaja de que el entrevistador puede pedir al entrevistado que explique su respuesta si no está seguro de que ésta pueda entrar en las categorías previstas»; pero esa ventaja queda anulada por las desventajas de que «no sabemos en qué medida los entrevistadores entienden correctamente las respuestas», y que «la presión generada por la situación de la entrevista puede provocar errores de codificación en mayor medida que si la codificación se hace a posteriori» [Sudman y Bradburn, 1982, 153].

pregunta puede tener interpretaciones distintas, la pregunta cerrada, gracias a sus respuestas preestablecidas, indica al entrevistado cuál es el marco de referencia presupuesto y ofrece el mismo marco a todos los entrevistados.

b) La pregunta cerrada facilita el recuerdo: cada alternativa propuesta sirve como apunte para el entrevistado.

c) La pregunta cerrada estimula el análisis y la reflexión, obliga a vencer la vaguedad y la ambigüedad, ayuda al entrevistado a distinguir opciones que aún están confusas en su mente y que, probablemente, si se le planteara una pregunta abierta las expresaría de forma poco clara. Por ejemplo, aunque a una pregunta abierta sobre un «buen» trabajo el entrevistado responderá instintivamente «un trabajo que dé satisfacciones», la pregunta cerrada le obligará a reflexionar y a distinguir entre el salario, el desarrollo de sus capacidades, las posibilidades de promoción, etc.

De forma análoga, podemos identificar también tres *limitaciones* básicas de la pregunta cerrada:

a) Excluye todas las demás respuestas posibles que el entrevistador no ha previsto. Al imponer a la situación estudiada el esquema conceptual del investigador se produce una «limitación anticipada del horizonte teórico», que ya hemos mencionado al hablar de los inconvenientes de la investigación cuantitativa. Como escribe Cicourel, «el cuestionario de preguntas con respuesta preestablecida excluye la posibilidad de obtener definiciones de la situación no previstas que pueden revelar los pensamientos íntimos y los sentimientos de los sujetos» [1964, 105].

b) Las opciones a elegir propuestas por la pregunta cerrada influyen en las respuestas. Las categorías preestablecidas conllevan el riesgo de sugerir la respuesta incluso a quienes no tienen una opinión formada sobre el tema. La pregunta cerrada propicia lo que hemos llamado «pseudo-opiniones», induce a responder al azar. Este problema no se plantea con las preguntas abiertas, puesto que el entrevistado no puede disfrazar su «no sé» recurriendo a las alternativas que se le proponen.

c) Las opciones de respuesta propuestas pueden no tener el mismo significado para todos los sujetos. La precodificación de las respuestas, que tiene como fin poder compararlas, puede ser una ilusión si cada individuo les atribuye un significado distinto. La pregunta cerrada requiere una doble interpretación-comprensión por parte del entrevistado, que debe entender no sólo la pregunta, sino también todas las posibles respuestas propuestas. Por tanto, hay muchas más posibilidades, en comparación con la pregunta abierta, de que haya interpretaciones distintas entre los distintos sujetos. En este sentido, hay que añadir que la pregunta cerrada disfraza la incomprensión de la pregunta, mientras que si la pregunta es abierta por lo general el entrevistador puede percibir si el entrevistado ha comprendido efectivamente el sentido de la pregunta.

De todo lo dicho hasta ahora, se deduce que la pregunta cerrada implica una condición fundamental: que todas las alternativas posibles estén presentes en el abanico de respuestas ofrecidas al entrevistado. La táctica de añadir al final de la lista la opción «otros» es necesaria, pero poco eficaz: de hecho, las alternativas presentadas al entrevistado atraen su atención, y él tiende a «elegir» entre esas propuestas; difícilmente, a menos que tenga las ideas muy claras y una posición firme al respecto, rechazará expresamente las respuestas citadas para ofrecer una distinta. Por consiguiente, la pregunta cerrada exige un estudio exploratorio sumamente cuidadoso, realizado con las mismas preguntas pero de forma abierta, sobre una muestra (no demasiado reducida) de sujetos extraída de la misma población a la que luego se aplicará el cuestionario, con el objetivo de identificar todas las posibles respuestas.

En algunos casos la pregunta cerrada no es viable: cuando las posibles respuestas no están totalmente claras para el investigador, cuando éstas son demasiado numerosas o tratan sobre cuestiones demasiado complejas (por ejemplo, una pregunta sobre las razones por las que un hijo tiene una relación conflictiva con sus padres), cuando los entrevistados tienen un nivel cultural muy bajo y están poco habituados al lenguaje abstracto de las respuestas precodificadas, cuando se trata de cuestiones delicadas, accesibles sólo mediante una entrevista en profundidad (por ejemplo, la conducta sexual, el consumo de drogas, etc.).

Así como las limitaciones de la pregunta abierta estaban claras, las de la pregunta cerrada también son evidentes. Sin embargo, cuando la muestra es muy numerosa, a partir de varios centenares de entrevistados, no queda otra alternativa que la pregunta cerrada. En un grupo de investigación pequeño, donde el responsable controla y dirige directamente la labor de los entrevistadores, y éstos se relacionan entre sí continuamente y trabajan con un número de entrevistas moderado, se podría pensar en entrevistas con preguntas abiertas, realizando lo que se denomina entrevista estructurada, que estudiaremos con detalle en el capítulo 10.

Como ya hemos dicho, esta forma de proceder es inviable en las encuestas por muestreo a gran escala. Cuando los sujetos estudiados son varios centenares o miles, distribuidos por un amplio territorio, y la investigación utiliza un número elevado de entrevistadores, se impone la opción de estandarizar el instrumento de investigación. Si no estandarizamos las preguntas y las respuestas, la labor de los entrevistadores y los procedimientos de codificación, quedaríamos abrumados por una enorme masa de datos no interpretables, incompletos, incongruentes e incodificables. Por este motivo, actualmente las encuestas por muestreo sobre muestras grandes se realizan con cuestionarios de preguntas casi en su totalidad cerradas.

5.5. FORMULACIÓN DE LAS PREGUNTAS

Hay un chiste que se suele citar en los manuales de técnicas de investigación social sobre la importancia de la forma en que se plantean las preguntas en los cuestionarios.

> Dos religiosos de congregaciones diferentes discuten sobre si es pecado fumar y rezar al mismo tiempo. Al no saber cómo resolver la cuestión, deciden planteársela a sus respectivos superiores. Cuando se vuelven a encontrar, el primero le pregunta al segundo: «¿Qué te ha dicho tu superior?». «Me ha dicho que es pecado». «Qué raro, replica el primero, el mío me ha dicho lo contrario». «¿Pero tú qué le has preguntado?». «Le he preguntado si es pecado fumar mientras se reza», dice el segundo. «¡Ah, responde el primero, yo le he preguntado si es pecado rezar mientras se fuma!».

Es obvio que la manera en que se formula la pregunta puede influir en la respuesta. Sin embargo, no está de más insistir sobre esta cuestión. Hasta los mínimos cambios en la expresión verbal de las preguntas pueden provocar variaciones sensibles en las respuestas. Uno de los primeros experimentos al respecto es el que realizó en los años treinta Rugg [1941]. Este investigador descubrió que la aceptación de medidas para limitar la libertad de expresión era un 21 % superior cuando se preguntaba: «¿Cree usted que en Estados Unidos se debería permitir dar discursos públicos contra la democracia?», que cuando se planteaba la misma cuestión con la pregunta «¿Cree usted que en Estados Unidos se debería prohibir dar discursos públicos contra la democracia?». En el primer caso hubo un 75 % de respuestas negativas (no permitir), mientras que en el segundo, sólo hubo un 54 % de respuestas afirmativas (prohibir). Aunque los significados de los dos verbos «no permitir» y «prohibir» son equivalentes, el segundo parece implicar un veto mucho más firme, probablemente porque «implica un ejercicio activo de oposición y no la simple negación de un apoyo activo» [Hipler y Schwarz, 1986, 89].

Schuman y Presser [1981, 284 sq.] citan otro ejemplo, donde la pregunta fue la siguiente: «Si se produjera en otra parte del mundo una situación como la de Vietnam, ¿cree usted que Estados Unidos deberían intervenir, o no debería hacerlo?». La misma pregunta se planteó a una segunda muestra, con el siguiente añadido (en cursiva): «¿[...] debería intervenir *para frenar una invasión comunista,* o no debería hacerlo?». En la primera formulación el 18 % de los entrevistados se manifestó a favor de la intervención, y la segunda recibió un 33 % de respuestas favorables (casi el doble).

En los manuales sobre la técnica del cuestionario pueden encontrarse innumerables ejemplos similares, que ilustran la sensibilidad de las respuestas al modo en que se formula la pregunta. Aunque es difícil elaborar unas reglas precisas sobre esta cuestión, a continuación ofrecemos una serie de sugerencias relacionadas con el lenguaje, la sintaxis y el contenido de las preguntas.

1. *Sencillez en el lenguaje.* Dada la estandarización de las preguntas, que impone que sean iguales para todos, es necesario emplear un lenguaje fácil de entender. Por ejemplo, en un cuestionario realizado en Italia que contenía la pregunta: «Muchos expertos consideran que la prisión preventiva es de gran utilidad para la ciudadanía porque impide que los delincuentes potenciales repitan sus delitos o escondan las pruebas. ¿Está usted a favor de ello?», resultó que menos del 50 % de los entrevistados conocía el significado exacto del término «prisión preventiva». En general, hay que tener presente que: *a*) el lenguaje del cuestionario debe adecuarse a las

características de la muestra estudiada. Podemos preguntar a trabajadores temporeros, inmigrantes, ancianos, estudiantes sobre los mismos temas, pero con distintas preguntas; *b*) el cuestionario autocumplimentado exige un lenguaje más sencillo que el cuestionario con entrevistador, ya que el primero no se puede acompañar con explicaciones, y *c*) incluso en presencia de un entrevistador no hay que apoyarse en sus intervenciones explicativas, ya que los entrevistados podrían avergonzarse de reconocer que no entienden algo, y muchos preferirán responder al azar antes que pedir explicaciones.

2. *Extensión de las preguntas.* En líneas generales, las preguntas, además de estar formuladas en un lenguaje sencillo, también deben ser breves. Las preguntas demasiado largas no sólo hacen que la entrevista completa sea demasiado extensa, sino que además pueden distraer al entrevistado del centro de la pregunta; además, puede suceder que al llegar al final de la pregunta el entrevistado haya olvidado el principio y base su respuesta sólo en la última parte de la misma. Hay que decir, sin embargo, que cuando se trata de cuestiones complicadas, puede ser preferible la pregunta más larga, ya que: *a*) facilita el recuerdo; *b*) da más tiempo al entrevistado para pensar (además, la extensión de la respuesta es proporcional a la extensión de la pregunta); y *c*) favorece una respuesta más organizada.

La pregunta más larga, por tanto, puede ser preferible cuando se trata de conocer cuestiones personales delicadas, o que requieran una reflexión atenta o que obliguen a recurrir a la memoria. En estos casos, se puede introducir la pregunta planteando la cuestión como un problema o citando ejemplos del tipo: «Como usted sabe, hay diferentes opiniones sobre la conveniencia de [...]. Hay quien piensa que [...], y quien piensa que [...]. ¿Con qué postura se identifica usted más?», en lugar de preguntar simplemente: «¿Está usted a favor o en contra de [...]?».

3. *Número de opciones de respuesta.* En las preguntas cerradas, las opciones de respuesta ofrecidas al entrevistado nunca pueden ser demasiado numerosas. Si se presentan oralmente, se recomienda que no sean más de cinco, pues de lo contrario será difícil recordar las primeras cuando el entrevistador llega a las últimas [Sudman y Bradburn, 1982, 172]. Para números superiores se suele presentar al entrevistado una tarjeta en la que se reproducen las respuestas por escrito, pero incluso en este caso no debe superarse una cantidad razonable de opciones.

4. *Expresiones en jerga.* Muchos grupos protegen celosamente su jerga y el empleo de la misma por parte de extraños podría irritar al entrevistado, o hacer que la pregunta le parezca ridícula. En general, el entrevistador se asegurará la cooperación de los entrevistados subrayando su profesionalidad científica, más que intentando adherirse a la subcultura de aquéllos [Bailey 1994].

5. *Términos ambiguos.* Es necesario evitar el uso de palabras con significado ambiguo. Por ejemplo, la pregunta «¿tiene un empleo estable?» implica un concepto de estabilidad del empleo que no tiene un significado unívoco: ¿se considera *estable una suplencia anual en una escuela?, ¿y un contrato que se renueva cada* año? Ante la pregunta, «¿cuántas habitaciones tiene su casa?», podríamos plantear qué se entiende por habitación y si el concepto incluye también el baño y la cocina.

6. *Palabras con una fuerte carga emocional.* Conviene evitar los términos con un componente emocional importante, sobre todo si es negativo. Si queremos saber

si un padre aplica castigos físicos a su hijo, tendremos que evitar preguntas directas como: «¿pega usted a su hijo?», ya que la palabra «pegar» tiene una connotación muy negativa. Así mismo, deben evitarse palabras como libertad, justicia, igualdad, comunismo, capitalismo, gran industria, etc. Es preferible utilizar locuciones en lugar de términos tan directos, aunque esto haga que la pregunta sea más larga.

7. *Preguntas de sintaxis compleja.* La pregunta debe tener una sintaxis sencilla. Hay que evitar, por ejemplo, las dobles negaciones, como se comprueba en la pregunta siguiente: «Diga si está de acuerdo o en desacuerdo con la afirmación siguiente: No es cierto que los trabajadores estén tan mal como dicen los sindicatos». Quien contesta «en desacuerdo» está en desacuerdo con la negación de que los trabajadores estén mal y por tanto considera que están mal... Como puede percibirse, esta pregunta es muy complicada y puede inducir a error.

En general, conviene evitar someter al juicio del entrevistado una frase negativa que contenga desaprobación o condena hacia un determinado objeto, personaje o conducta. En este caso, el entrevistado que aprueba esa conducta o ese personaje, debe responder negativamente, para negar la negación, y el que desaprueba esa conducta debería responder afirmativamente, para expresar su condena. Pero suele suceder que se declara «en desacuerdo» cuando su desacuerdo no se refiere a la afirmación, sino a la conducta propuesta. Este problema no se plantea si utilizamos exclusivamente frases afirmativas.

8. *Preguntas con respuesta no unívoca.* Deben evitarse las preguntas múltiples, es decir, que contengan más de un interrogante, y las preguntas sobre cuestiones que no estén lo suficientemente delimitadas. Por ejemplo, la pregunta: «¿sus padres eran religiosos?» resulta imposible de responder en caso de que el padre y la madre tuvieran posturas distintas al respecto; a la pregunta: «¿considera que la política económica del gobierno es justa y eficaz?», no sabrá qué responder quien la considera eficaz pero injusta. En estos casos el problema puede resolverse planteando dos preguntas separadas.

En cuanto a la delimitación de la cuestión, no podemos plantear la pregunta «¿está a favor o en contra del aborto?», sin distinguir los diferentes casos (uno puede estar en contra del aborto, pero aprobarlo en caso de malformación del feto o peligro para la vida de la madre), o preguntar a un joven «¿usted trabaja?» sin precisar cuándo, en verano, en invierno o todo el año, si se cuentan los trabajos ocasionales o los de pocas horas a la semana, como cuidar niños, etc.

9. *Preguntas no discriminantes.* Las preguntas deben construirse de tal forma que produzcan diferencias en las respuestas de todos los entrevistados. Una pregunta a la que el 90 % de los encuestados responden igual suele ser una pregunta inútil (salvo en casos particulares en los que se pretende aislar a una minoría específica). Por este motivo, al presentar las opciones de respuesta hay que evitar aquellas que atraigan de forma evidente una elevada proporción de consensos. Por ejemplo, si se formula la pregunta «¿de entre los siguientes grupos de personas, de cuáles se fía usted más?», seguida de una lista de categorías de personas (vecinos de casa, sacerdotes, compañeros de trabajo, profesores, etc.), no se debe incluir en esta lista a «familiares». De forma similar, si se pregunta «¿en qué país de los que siguen considera usted que se vive mejor?», no debe incluirse el propio país (a menos que

se quiera identificar explícitamente a los xenófilos).

10. *Preguntas tendenciosas*, también llamadas viciadas o de respuesta dirigida. A veces, incluso de manera inconsciente, el investigador construye una pregunta que, por la adjetivación empleada, por los ejemplos citados o por la combinación de palabras utilizada, predispone al entrevistado hacia una de las posibles opciones de respuesta, en vez de presentarlas de forma equilibrada. En una investigación realizada en Francia sobre cuestiones relacionadas con luchas sindicales se planteó la siguiente pregunta: «En nuestro país hay 700 religiosos que han dicho que el Evangelio es un mensaje para los *pobres* y los *oprimidos*, por ello *viven pobremente, trabajan en fábricas y se afilian* a sindicatos y organizaciones políticas para *ayudar* a los trabajadores a conseguir más *conquistas sociales*. ¿Cree usted que estos religiosos hacen bien?». Hemos marcado en cursiva siete expresiones que dan una connotación claramente positiva al comportamiento de los sujetos sobre los que se pide la opinión del entrevistado. Es evidente que esta pregunta está formulada de manera sesgada, lo que se deriva de la necesidad de presentar al entrevistado una descripción de la figura del religioso-obrero. Pero esa descripción tiene tantos elementos positivos que predispone al entrevistado hacia la respuesta afirmativa, a menos que éste ya tenga las ideas claras al respecto.

En este contexto podemos citar también la pregunta sobre el aborto. Se obtienen resultados muy distintos según si se plantea la pregunta desde el punto de vista de la elección de la maternidad responsable o del derecho a la vida por parte del feto.

Una forma de tendenciosidad consiste en formular una pregunta cerrada omitiendo algunas de las posibles respuestas en la lista de opciones (las opciones que no se presentan de manera explícita estarán infrarrepresentadas en las respuestas). Por ejemplo, en la pregunta «¿cómo pasa usted su tiempo libre, viendo la televisión o haciendo otra cosa?», o bien «algunos piensan que las mujeres deberían tener una posición igual a los hombres en puestos directivos y cargos públicos; otros piensan que el lugar de la mujer está en el hogar y que su tarea principal es criar a los hijos y cuidar de la familia, ¿qué opina usted?». Sin la segunda parte («otros piensan que...») la pregunta sería tendenciosa.

11. *Comportamientos supuestos.* Es indispensable no dar por supuestos ciertos comportamientos en los entrevistados. Por ejemplo, no se debe preguntar al entrevistado a quién votó en las últimas elecciones sin antes preguntarle si fue a votar. Las preguntas que permiten seleccionar a los encuestados antes de plantear una cuestión que no es relevante para todos se suelen llamar «preguntas filtro», y las preguntas que se plantean sólo si el sujeto ha dado una respuesta determinada a la pregunta filtro se llaman «preguntas condicionadas». Sin estos filtros puede ocurrir que, al sentirse obligado a contestar, el entrevistado dé una respuesta (al azar o según lo socialmente deseable), a una pregunta que no le concierne.

12. *Precisión en el tiempo.* En general conviene tener cuidado con las preguntas que se refieren al comportamiento habitual o que requieren el cálculo de la media en el tiempo. Por ejemplo, las preguntas del tipo «¿con qué frecuencia suele usted leer el periódico?» y «¿cuántas veces al mes va al cine?», presentan el riesgo de que el entrevistado responda según «lo que debe ser», según la imagen que tiene y quiere ofrecer de sí mismo, y no según su comportamiento real (ver también el

apartado 15 de esta serie). Por ello conviene añadir a estas preguntas una segunda referida a un momento temporal determinado, preguntando también si hoy (ayer) ha leído el periódico y si ha ido al cine en las dos últimas semanas. La precisión en un periodo de tiempo definido facilita el recuerdo y hace más difícil que el comportamiento ideal se imponga al real en la respuesta.

La precisión en el tiempo también sirve para preguntas relativas a hechos que cambian con el tiempo. Por ejemplo, sugerimos no preguntar simplemente «¿en qué trabajaba su padre?», sino «¿en qué trabajaba su padre cuando tenía usted catorce años?», puesto que el padre puede haber cambiado de trabajo en el curso del tiempo, y la falta de precisión temporal puede generar dificultades para el entrevistado.

13. *Concreción – abstracción.* Las consideraciones son similares a las del punto anterior. La pregunta abstracta puede dar lugar a respuestas genéricas, superficiales, o que reflejan las normas sociales más que el pensamiento real del sujeto. En cambio, la concreción facilita la reflexión y la identificación con el problema real, y excluye el malentendido. Por ejemplo, en un estudio realizado en Italia sobre la pena de muerte, a la pregunta general «¿cree usted que la pena de muerte debería existir para los casos de delitos de gravedad excepcional?», el 42 % de los entrevistados respondió «no». Sin embargo, cuando se proporcionaron ejemplos concretos de ese tipo de delitos, sólo un 29 % del total de los entrevistados manifestó su desacuerdo [Corbetta y Parisi, 1983].

Otro ejemplo de concreción de una actitud abstracta lo constituye la siguiente pregunta, usada para registrar «el sentido de la patria» en una muestra de jóvenes: «Nadie quiere que haya guerras; pero si estallara una, ¿crees que irías a combatir por tu país?». El sentimiento de patriotismo es tan abstracto y difícil de definir que sólo mediante la visualización de una situación concreta, aunque hipotética, podemos plantear una pregunta al respecto.

A veces, en lugar de una pregunta abstracta, puede ser útil relatar un episodio sobre el que se pide al encuestado que dé su opinión. Consideremos el siguiente ejemplo, donde en lugar de hacer una pregunta abstracta sobre el contraste entre rigor ético y relativismo oportunista, se ilustra un caso concreto.

Es por la mañana temprano, y el señor Moretti, de camino a su trabajo en bicicleta, sale de repente de una callejuela a una calle del centro de la ciudad. Justo en ese momento pasa un coche de gran cilindrada que, para evitar al ciclista, gira y acaba chocando contra un pilar. Moretti se declara culpable y promete pagar los daños de la carrocería. Moretti cuenta lo sucedido a sus compañeros de trabajo, que le aconsejan que no sea tonto y no pague nada, puesto que no había testigos: la reparación puede costarle muy cara, mientras que para el propietario del coche probablemente será un gasto de poca monta. Pero Moretti se niega a hacerlo, ya que piensa que si el conductor del coche no hubiera girado, él podría haber muerto. A continuación el entrevistador pregunta: «¿Está usted de acuerdo con Moretti o con sus amigos?» [Marradi, 1996, 34].

El objetivo de la concreción se aplica también a las respuestas. Por ejemplo, Groves [1989, 452 sq.] cita varios estudios sobre los distintos significados atribuidos por los entrevistados a los adverbios «a menudo», «bastante a menudo», etc.:

por ello se recomienda, en la medida de lo posible, sustituirlos por un abanico de tiempo bien definido y concreto.

14. *Comportamientos y actitudes.* Ya hemos dicho que normalmente en un cuestionario pueden incluirse preguntas sobre comportamientos y actitudes. Sin embargo, también hemos dicho que las actitudes son por naturaleza mucho más imprecisas y ambiguas y se prestan más a respuestas normativas que los comportamientos. Por tanto, es recomedable, si el tema lo permite, centrar la pregunta en comportamientos más que en opiniones. Por ejemplo, sobre el tema de la participación política, en lugar de preguntar al entrevistado si está al corriente de la política, será mejor preguntarle si lee noticias sobre política en los periódicos, o si ve el telediario. Así mismo, en lugar de preguntarle si considera correcto dedicar tiempo y dinero a actos de caridad y solidaridad social, será mejor preguntarle si hace donaciones o participa en organizaciones solidarias, etc.

Otra norma recomendable es escoger comportamientos sobre los que existe confirmación empírica. Por ejemplo, además de preguntar si está al corriente de la política (actitud) y si lee los periódicos (comportamiento), también puede hacerse una pregunta que implique un conocimiento específico, por ejemplo, si conoce un hecho reciente de la crónica política, o el nombre del líder de un partido o del sindicato de su ciudad. Como ejemplo concreto no se planteará la pregunta genérica «¿lee usted libros?», sino que se preguntará cuántos libros se han leído en los seis últimos meses y se pedirán los títulos de estos libros.

15. *Respuestas socialmente deseables.* Hasta ahora hemos tratado este problema desde un punto de vista general, presentándolo como una de las mayores dificultades de las encuestas. *Ahora vamos a ver cómo puede abordarse la cuestión desde un punto de vista técnico, en el momento de la creación del cuestionario.*

En primer lugar, para este problema específico repetimos la recomendación de formular las preguntas ligándolas lo más posible a casos concretos. La pregunta abstracta favorece respuestas genéricas y acordes con las normas sociales, mientras que los casos concretos obligan al entrevistado a posicionarse (recordemos la pregunta sobre la actitud con respecto a la pena de muerte, que es fácil rechazar en el plano general, porque probablemente ésa es la norma compartida, pero que se suele aceptar cuando se ponen ejemplos de casos concretos).

Es recomendable formular este tipo de preguntas de tal modo que incluso la respuesta menos deseable resulte aceptable, ofreciendo para ella una justificación. Por ejemplo, no preguntar «¿lee usted el periódico?», sino «¿suele usted encontrar tiempo para leer el periódico?», o bien «El 21 de abril de este año hubo elecciones. Muchas personas no van a votar porque están enfermas, o porque se encuentran muy lejos del lugar donde están empadronados, o porque consideran inútil votar, o por otros motivos. ¿Usted fue a votar?».

Otra recomendación sería considerar el comportamiento negativo como algo normal y extendido (y, por tanto, no desviado); por ejemplo: «Todo el mundo, en algún momento de su vida, ha pensado en el suicidio. Usted, en algún momento de depresión, ha pensado alguna vez [...], etc.?».

También se puede formular la pregunta de manera que todas las respuestas se presenten como posibilidades igualmente aceptables, para descartar que alguna de

ellas parezca más deseable que las otras. Por ejemplo: «Algunas personas dicen que fumar marihuana lleva al consumo de drogas duras; en cambio, otras dicen que la marihuana es totalmente inocua. ¿Qué opina usted?». Esta formulación «Algunas personas dicen que [...]. Otras piensan que [...] ¿Qué opina usted?» es muy recomendable para presentar distintas alternativas de forma equilibrada.

Otra opción es atribuir al entrevistado el comportamiento socialmente indeseado, dejándole a él la tarea de desmentirlo, si es el caso. Alfred Kinsey utilizó una estrategia de este tipo en su célebre estudio sobre el comportamiento sexual de los estadounidenses: en lugar de preguntar a las mujeres solteras si tenían relaciones sexuales, les preguntaba directamente con qué frecuencia las tenían [Kinsey *et al,* 1953].

Algunos autores proponen formular las preguntas en tercera persona, desplazando la atención sobre una persona distinta al entrevistado. Aunque es poco realista pensar que mediante un cuestionario se puede saber si un joven consume drogas, sí se podría captar alguna de las motivaciones de ese comportamiento con una pregunta indirecta como: «En la actualidad muchos jóvenes consumen drogas; ¿podrías decirme por qué motivos crees tú que lo hacen?».

Conviene subrayar que aunque todas estas tácticas pueden atenuar los efectos de las respuestas socialmente deseables, nunca conseguirán eliminarlos completamente: una pregunta sobre un tema con una expectativa social determinada implicará inevitablemente un grado de imparcialidad desconocido, salvo en los pocos casos en que es posible controlar el comportamiento.

16. *Preguntas embarazosas.* Hay una serie de cuestiones delicadas, como el comportamiento sexual, los ingresos, las conductas desviadas (consumo de drogas, alcoholismo), etc., que resulta muy difícil investigar con un instrumento como el cuestionario. Estos temas deberían estudiarse con preguntas abiertas, que permitan al entrevistado responder con sus propias palabras y dar explicaciones detalladas. No obstante, son cuestiones que sólo se pueden estudiar con entrevistas no estructuradas, en las que un entrevistador experimentado logra ganarse la confianza de los entrevistados. Por ejemplo, Kinsey, el primero que trató el tema del comportamiento sexual, utilizó como única herramienta para recopilar material empírico las entrevistas en profundidad, que realizó en su mayoría él mismo.

17. *Falta de opinión y respuesta «no sé».* Ya hemos hablado del problema de la falta de opinión, las preguntas que el entrevistado no se había planteado antes, la presión por dar una respuesta y las opiniones que nacen en el momento mismo de la pregunta.

El problema sólo puede resolverse dejando bien claro al entrevistado que «no sé» es una respuesta tan aceptable como las demás, por ejemplo incluyéndola como una de las posibles opciones. En lugar de preguntar «¿cree usted que muchos de los cargos públicos son corruptos o sólo unos pocos?», podríamos preguntar «¿cree usted que muchos de los cargos públicos son corruptos, o sólo unos pocos, o no tiene una opinión al respecto?». Una forma algo distinta consiste en anteponer a la pregunta verdadera una pregunta «filtro», como: «Hay quien afirma que muchos de los cargos públicos son corruptos, otros dicen que sólo lo es una pequeña minoría. ¿Tiene usted una opinión al respecto?». (Si la respuesta es afirmativa): «¿Con quién

está usted de acuerdo?». Schuman y Presser [1981, 116 sq.] han realizado experimentos formulando las preguntas con y sin filtros. La pregunta que hemos mencionado obtuvo el 4,2 % de respuestas «no sé» en la versión sin filtros, y casi el 20 % en la versión con filtros.

El sujeto inseguro que no tiene una opinión sobre el tema planteado y se siente incómodo respondiendo «no sé», probablemente contestará al azar o buscará un indicio cualquiera (en la formulación de la pregunta, en la actitud del entrevistador, en su tono de voz, en la expresión de su cara), para descubrir la respuesta «correcta». Por ello, el investigador deberá ser muy cuidadadoso y procurar formular la pregunta de forma neutral y evitar sugerencias directas o indirectas.

18. *Intensidad de las actitudes.* Además de saber si los entrevistados están a favor o en contra de las actitudes, también es importante conocer la intensidad de *sus posturas, dado que ésta determina el comportamiento. Por ejemplo, en Estados Unidos se viene constatando en los sondeos que una amplia mayoría de la población*, tres cuartos aproximadamente, está a favor del control de la posesión de armas; sin embargo, no se consigue aprobar una ley que regule esta cuestión y ello se atribuye, entre otras razones, al hecho de que la minoría que está en contra del control es mucho más combativa y activa que la mayoría favorable a dicho control.

Schuman y Presser [1981, 231 sq.] sugieren dividir la cuestión en varios aspectos: *intensidad, importancia y consecuencias conductuales.* En el caso del control de la posesión de armas, a la pregunta básica «¿estaría usted a favor de una ley que obligara a las personas a solicitar una licencia a la policía antes de comprar un arma, o está usted en contra de ello porque piensa que esa ley viola el derecho de los ciudadanos a poseer un arma?», ellos añaden una pregunta de intensidad: «en comparación con lo que usted piensa sobre otros temas de intervención pública, ¿sus sentimientos sobre el control de la posesión de armas son: muy intensos, bastante intensos, poco intensos, nada intensos?», una pregunta sobre la importancia que el entrevistado atribuye al tema: «cuando usted tiene que votar al candidato al Congreso, ¿qué importancia tiene para usted su postura sobre el control de la posesión de armas: muy importante, importante, poco importante, nada importante?», una pregunta sobre la relevancia conductual: «¿ha escrito usted alguna vez una carta a las autoridades o algún periódico, o ha enviado dinero a alguna asociación, en relación con el problema sobre el control de la posesión de armas?»). Este estudio reveló que la minoría contraria al control de la posesión de armas tenía sentimientos mucho más intensos hacia el problema que la mayoría favorable a dicho control.

Como puede verse en este ejemplo, para medir la intensidad de una actitud no basta con formular adecuadamente la pregunta, sino que es necesario profundizar con preguntas posteriores. Obviamente, esto conlleva costes y va en contra del principio general de economía del cuestionario. Por esta razón es tan frecuente *descuidar el registro de la intensidad; como mucho el investigador se limita a preguntar si se está muy, bastante, poco o nada de acuerdo con una afirmación*[5].

[5] Y con frecuencia en la elaboración de las preguntas se abandonan las distinciones entre «mucho» y «bastante», y las respuestas acaban siendo a favor y en contra.

19. *Aquiescencia.* Con este término nos referimos a la tendencia de los entrevistados a elegir las respuestas que expresan acuerdo, a dar contestaciones afirmativas, más que negativas. Este comportamiento es más frecuente en las personas menos instruidas.

Schuman y Presser dividieron una muestra de entrevistados en dos grupos. A uno le plantearon la pregunta «¿Está usted de acuerdo con la siguiente afirmación: los delitos y violaciones de la ley se producen más por culpa de los individuos que por culpa de la sociedad?». Al otro le hicieron la misma pregunta pero invirtiendo sociedad e individuos («¿[...] se producen más por culpa de la sociedad que de los individuos?»). En la primera formulación se obtuvo un 59 % de respuestas favorables (que atribuían la culpa a los individuos), y en la segunda, se obtuvo un 57 % de respuestas favorables (que atribuían culpa a la sociedad). Estos resultados son contradictorios. Los investigadores formularon entonces la pregunta de la siguiente forma: «¿En su opinión, quién es más culpable de los delitos y violaciones de la ley, la sociedad o los individuos?». Esta vez un 54 % de los entrevistados atribuyó la culpa a la sociedad, y un 46 % a los individuos [Schuman y Presser, 1981, 204 sq.].

Las preguntas se pueden formular teniendo en cuenta esta posible distorsión, siempre que quien elabora el cuestionario sea consciente del problema.

Una forma de distorsión en las respuestas similar a ésta es la llamada *response set*: frente a una batería de preguntas con el mismo tipo de opciones de respuesta (por ejemplo: totalmente, bastante, poco, nada de acuerdo), algunos entrevistados, por pereza (para acabar la entrevista antes) o por falta de opiniones, podrían responder siempre con la misma opción (por ejemplo, «bastante de acuerdo»), independientemente del contenido de las preguntas. Este problema se resuelve alternando la polaridad de las respuestas, es decir, formulándolas de manera que un individuo con ideas coherentes tenga que responder a unas preguntas en sentido positivo y a otras en sentido negativo para no contradecirse. Por ejemplo, en una batería de preguntas sobre política, se podrían formular dos preguntas consecutivas de tal manera que en la primera la postura conservadora coincida con la opción de respuesta «de acuerdo» y en la segunda con la opción de respuesta «en contra».

Pueden surgir problemas similares debidos a la aquiescencia o la pereza, cuando se presenta al entrevistado una lista de opciones, entre las cuales debe hacer una elección múltiple. Por ejemplo: «Entre las siguientes características, ¿cuáles considera que son las más importantes para un buen trabajo: posibilidad de promoción, salario, flexibilidad de horarios...? Indique las tres que considera más importantes». Se ha constatado que cuando el entrevistado lee las opciones posibles (por ejemplo, en una ficha), existe una tendencia a escoger las primeras de la lista; cuando el entrevistador las enumera oralmente, el encuestado tiende a elegir las últimas de la lista. Esta distorsión se puede evitar variando el orden de las alternativas de respuesta al pasar de una entrevista a otra[6].

20. *Efecto memoria.* Las preguntas relativas a hechos y comportamientos del

[6] Esta táctica no evita el posible error individual por parte del entrevistado, pero al menos evita un error sistemático, es decir, un error en la misma dirección en todos los entrevistados.

pasado conllevan dificultades específicas debidas a lagunas o distorsiones en el recuerdo. Se han sugerido algunas tácticas en la formulación del cuestionario para *reforzar la validez de las preguntas basadas en el recuerdo.*

La primera táctica consiste en establecer límites temporales para el recuerdo. Por ejemplo, en lugar de preguntar a una persona cuántas veces ha estado en el extranjero o si alguna vez ha sido víctima de un delito (robo, hurto, etc.), se situará la pregunta en un intervalo de tiempo definido (en el último año, en los últimos seis meses, etc.)[7]. La segunda estrategia consiste en utilizar puntos de referencia temporales de hechos más importantes que el estudiado para ubicar temporalmente un hecho pasado. Si, por ejemplo, se quiere saber cuándo ha comprado una lavadora un sujeto, se le preguntará si fue antes o después del nacimiento de su último hijo, antes o después de las vacaciones de verano, etc.

Otra manera de estimular el recuerdo consiste en ofrecer listas de posibles respuestas. Por ejemplo, en una pregunta sobre las estrategias empleadas para buscar empleo, el investigador no preguntó qué iniciativas se habían tomado, sino que presentó una lista de posibles estrategias, y para cada una de ellas preguntó al entrevistado si la había puesto en práctica o no. Se podría aplicar la misma estructura a una pregunta sobre la audiencia televisiva, usando una lista de los programas transmitidos.

Cuando el comportamiento a estudiar se refiere al pasado, pero también al presente y al futuro (por estar en desarrollo en el momento de la investigación), para evitar que el entrevistado recurra al recuerdo, se pueden utilizar diarios o instrumentos similares. Por ejemplo, los organismos estatales de estadística de todos los países utilizan diarios de gastos, estructurados por puntos, que se distribuyen a las familias estudiadas, a fin de registrar las cuentas económicas familiares[8].

Hay que añadir que, si en el momento de la entrevista la actitud del sujeto hubiera cambiado con respecto al pasado al que se refiere la pregunta, es muy probable que el sujeto, de manera inconsciente, proyecte su actitud actual al momento pasado.

21. *Orden de las preguntas.* Para concluir, reflexionaremos sobre los criterios para decidir el orden de presentación de las preguntas en un cuestionario. En este sentido conviene tener presente la dinámica de la relación entre el entrevistado y el entrevistador. Se trata de una relación asimétrica. Por una parte, está el entrevistador, que conoce perfectamente la situación de la entrevista, tiene amplia experiencia con ella, conoce su objetivo y puede prever su desarrollo. Por otra parte, está el entrevistado, que no sabe por qué le plantean esas preguntas, quién ha enviado al entrevistador, no entiende por qué lo ha elegido precisamente a él, tiene miedo de dar respuestas equivocadas, y vive la entrevista como una experiencia dominada por

---

[7] Recordemos que la memoria está afectada por el denominado «efecto telescopio», que hace que recordemos los episodios lejanos como si fueran más recientes de lo que son en realidad.

[8] Un registro de datos similar es el de los «balances de tiempo», que son unas fichas que se entregan a los sujetos para que anoten en ellas los horarios iniciales y finales de todas sus actividades en un intervalo de tiempo determinado (por lo general un día o una semana).

la desconfianza y la duda. La primera tarea del entrevistador será, por tanto, tranquilizar al entrevistado, hacerle entender que no tiene nada que temer. La segunda tarea consistirá en hacer que capte el mecanismo de la entrevista y la dinámica de la pregunta-respuesta.

De todo ello se deduce que la primera parte del cuestionario debe tener como objetivo hacer que el entrevistado se sienta cómodo y hacerle entender cómo funciona la entrevista. Para ello se recomienda por lo general empezar con las preguntas fáciles, no demasiado indiscretas ni personales, basadas en hechos más que en opiniones, preguntas que tienen el objetivo de tranquilizar e instruir.

En cambio, se recomienda situar las preguntas potencialmente embarazosas hacia la mitad del cuestionario, para que el entrevistador tenga tiempo de ganarse la confianza del entrevistado. A veces se sugiere situarlas al final del cuestionario, para minimizar los daños de una eventual interrupción de la entrevista y, sobre todo, para evitar el posible deterioro del clima de la conversación que podría provocar una pregunta entrometida planteada al principio.

El segundo criterio a tener en cuenta guarda relación con el *interés* y el *cansancio* del entrevistado. Es importante estructurar el cuestionario de tal forma que atraiga la atención del encuestado en todo momento. Se ha demostrado que la atención y el interés del sujeto siguen una curva creciente hasta la mitad de la entrevista aproximadamente, y después decrecen. Por ello es recomendable situar las preguntas más complicadas hacia la mitad de la entrevista, dejando para el final las preguntas más tediosas pero que no requieren reflexión, como las que conciernen a las características sociodemográficas. En cuanto a la extensión del cuestionario, es difícil establecer criterios generales, dado que este elemento depende en gran medida del tema y la población estudiada. Podemos decir que la duración media de una encuesta con un cuestionario estándar debería ser de unos 45 minutos en el caso de la entrevista cara a cara y unos 25 minutos en el caso de la entrevista telefónica.

El tercer criterio es el aspecto *secuencial* de la entrevista. Es necesario que los temas tratados en el cuestionario se desarrollen en un orden lógico, que la entrevista se presente como una conversación natural, sin cambios bruscos de temas (pasar de una cuestión a otra totalmente distinta), ni periodos temporales (pasar de la vida actual a la infancia y de ahí a la actualidad, etc.). Si se pasa de un argumento a otro conviene introducir el tema nuevo con frases como: «Pasamos ahora a una cuestión distinta, pero que [...]» o similares.

En el orden de las preguntas conviene pasar de preguntas generales a preguntas particulares, según la técnica denominada «de embudo», que consiste en formular primero preguntas amplias, para ir pasando paulatinamente a aspectos más específicos. Por ejemplo, para estudiar la influencia de las asociaciones de trabajadores en el voto, se podría formular el orden de preguntas de la siguiente forma: «¿Usted trabaja?». En caso de respuesta afirmativa: «¿Cómo asalariado o como autónomo?». Si la respuesta es «trabajo autónomo»: «¿Es usted miembro de alguna asociación de trabajadores?». En caso de respuesta afirmativa: «¿Esta asociación ha dado indicaciones a sus miembros sobre qué votar?». En caso de respuesta afirmativa: «¿Qué tipo de indicaciones?». Si ha sugerido que voten a un determinado partido o candidato: «¿Ha seguido usted esas indicaciones?».

El último punto a tratar en relación con el orden de las preguntas es el denominado *efecto contaminación,* que hace referencia al hecho de que, en ciertos casos, la respuesta a una pregunta puede estar influida por las preguntas que la preceden. Schuman y Presser [1981, 36 sq.] citan el siguiente ejemplo. La pregunta: «¿cree usted que debería estar legalmente permitido interrumpir el embarazo de una mujer que está casada y que no quiere tener más hijos?» obtuvo el 58 % de respuestas afirmativas cuando fue planteada por separado. Pero las respuestas afirmativas descendieron al 42 % cuando, en una segunda investigación, la misma pregunta se presentó integrada en una serie de preguntas sobre el aborto, tras ésta otra: «¿Cree usted que debería estar legalmente permitido interrumpir el embarazo de una mujer en el caso de que existieran muchas probabilidades de que el niño naciera con malformaciones?». En el segundo caso, la mención previa de una situación en la que el aborto parece particularmente justificado hace que el entrevistado se muestre más severo con una situación que no tiene una motivación similar. En este ejemplo, la primera pregunta ha predispuesto al entrevistado para hacer una interpretación distinta de la que habría hecho si no se hubiera presentado esta sucesión de preguntas.

Es posible que después de esta larga lista el lector se sienta abrumado ante la sensibilidad del instrumento-cuestionario en relación con la formulación de las preguntas, e incluso podría llegar a pensar que es una forma de recopilar información sobre la realidad social muy poco fiable y que está expuesta a las manipulaciones del investigador.

Es cierto que las respuestas están muy influidas por la manera en que se formulan las preguntas. Conviene señalar que esta dependencia es relevante sobre todo en el «análisis a una variable», es decir, en la lectura de los porcentajes de respuesta a cada una de las preguntas consideradas de forma aislada. Sin embargo, el investigador social, aunque está interesado en saber cuántos ciudadanos votan o cuántos son partidarios de la legalización de las drogas blandas, querrá estudiar sobre todo las relaciones entre las variables. Por ejemplo si la cuota de votantes o de partidarios varía según el género, la edad, la clase social, el nivel de estudios, las ideas políticas, el lugar de residencia, etc. El objetivo más general es identificar los determinantes del fenómeno social que se estudia, los factores que influyen o están correlacionados con el mismo. Desde esta perspectiva, aunque la pregunta esté formulada de manera que obtenga un porcentaje de respuestas «no sé» inferior a la realidad, o una distorsión hacia una de las opciones de respuesta, es presumible que las relaciones entre la variable estudiada y el género, la edad o la clase social, etc. no resulten muy alteradas, partiendo del supuesto de que tal distorsión afecte más o menos de la misma forma a todos los entrevistados[9].

Por consiguiente, el investigador social podrá aventurarse a hacer comparaciones y análisis de relaciones entre variables cuando las preguntas se planteen a todos los entrevistados de la misma forma, y en cambio deberá ser mucho más cauto a la

[9] Hay que añadir que este supuesto no está totalmente fundamentado, ya que las personas «más influidas por la formulación de las preguntas tienden a ser las menos instruidas y menos interesadas o menos implicadas en la cuestión particular abordada» [Schuman y Presser, 1981, 6].

hora de comparar resultados de encuestas distintas, sobre todo cuando las preguntas presenten formulaciones diferentes entre sí.

Por último, conviene recordar la importancia de que en todas las presentaciones de la evidencia extraída de un cuestionario, el investigador indique también la formulación exacta de la pregunta que los ha generado. Éste es un requisito fundamental para garantizar el rigor científico de esta técnica de producción de datos.

5.6. BATERÍA DE PREGUNTAS

En los cuestionarios es frecuente encontrar series de preguntas que, al tener la misma formulación (la misma pregunta introductoria y las mismas alternativas de respuesta, variando sólo el objeto al que hacen referencia), se presentan en un único bloque. Estas formulaciones compactas se conocen con el nombre de «batería de preguntas» y tienen como objetivo: *a*) ahorrar espacio (en el cuestionario) y tiempo (de entrevista), mediante la condensación de varias preguntas en una única formulación; *b*) facilitar la comprensión del mecanismo de respuesta (una vez entendido en la primera pregunta, queda entendido para todas las demás); *c*) aumentar la validez de la respuesta, dado que, al responder a cada pregunta, el entrevistado tiene en cuenta implícitamente las respuestas que ha dado a las otras preguntas de la batería, y *d*) permitir al investigador, en la fase del análisis de los datos, crear índices sintéticos que resuman en un único resultado las respuestas dadas a las distintas preguntas de la batería.

En la tabla 5.1 hemos reproducido varios ejemplos de estas baterías. En el primero se presenta al entrevistado una lista de situaciones, cuestiones, etc. a las que debe responder con un «sí» o un «no». En este ejemplo se trata de trastornos (que la persona puede tener o haber tenido), pero podrían ser también listas de bienes (que la persona puede poseer), de actos (que puede haber realizado), etc., presentados al entrevistado en forma de lista para facilitarle el recuerdo. A continuación, en la misma tabla, encontramos tres baterías de preguntas que se diferencian de la anterior en que la respuesta común contempla más de dos alternativas. Un ejemplo muy frecuente de este tipo son las preguntas con opciones de respuesta como muy, bastante, poco o nada satisfecho con una situación determinada, de acuerdo con una afirmación concreta, a favor o en contra de una opción política, etc.

Cuando, como sucede con las baterías, una misma pregunta se aplica a varios temas, es importante distinguir entre el caso en que se formula *en términos absolutos* y el caso en que se formula *en términos relativos*. En los ejemplos anteriores cada una de las preguntas de la batería se formulaba en términos absolutos, en el sentido de que cada elemento de la batería es una entidad autónoma, a la que el entrevistado puede responder de manera independiente con respecto a las otras preguntas. Por ejemplo, ante la pregunta: «¿tarda más tiempo del habitual en conciliar el sueño por la noche?», el entrevistado puede responder «sí» o «no», sea cual sea la posición de esta pregunta dentro de una batería que contiene otras preguntas similares. Lo mismo se puede decir, refiriéndonos a los ejemplos anteriores, de la pregunta: «¿su madre está al corriente de su homosexualidad?».

Tabla 5.1. Ejemplos de baterías de preguntas

Pregunta 1. Indique si en los últimos días ha sufrido alguno de los siguientes trastornos [a]

	No	Sí
— Tardo más tiempo del habitual en conciliar el sueño por la noche	1□	2□
- Me despierto a menudo en mitad de la noche	1□	2□
- Puedo concentrarme con facilidad cuando leo el periódico	1□	2□
— Me siento irritable y nervioso	1□	2□
— Durante gran parte del día tengo miedo de algo sin motivo	1□	2□
— Estoy asustado y al borde del pánico	1□	2□
- Tengo problemas digestivos	1□	2□
- A veces mi corazón late más rápido de lo normal	1□	2□
- A veces tengo vértigo y siento que me voy a desmayar	1□	2□
— Mis manos tiemblan tanto que los demás lo perciben	1□	2□

[a] Extraído del cuadro de depresión de Carroll [Robinson *et al*, 1991, 211], en los aspectos de insomnio, ansiedad psicológica y ansiedad somática (el tercer elemento presenta una dirección invertida con respecto a los otros).

Pregunta 2. ¿Qué familiares están al corriente de su homosexualidad? [a]

	Está al corriente	Lo sabe, pero finge no saberlo	Creo que lo sabe, pero nunca hemos hablado de ello	No lo sabe	Ausente (porque falleció o por otras causas)
Madre	1□	2□	3□	4□	5□
Padre	1□	2□	3□	4□	5□
Hermano/s	1□	2□	3□	4□	5□
Hermana/s	1□	2□	3□	4□	5□
Cónyuge	1□	2□	3□	4□	5□
Hijo/s	1□	2□	3□	4□	5□

[a] Investigación realizada sobre una muestra de homosexuales.

Pregunta 3. En relación con los siguientes aspectos de su trabajo, ¿se considera muy satisfecho, bastante, poco o nada satisfecho?

	Mucho	Bastante	Poco	Nada	No lo sé
Autonomía	1□	2□	3□	4□	5□
Responsabilidad	1□	2□	3□	4□	5□
Nivel técnico	1□	2□	3□	4□	5□
Salario	1□	2□	3□	4□	5□
Posibilidad de promoción	1□	2□	3□	4□	5□
Seguridad del puesto	1□	2□	3□	4□	5□

Pregunta 4. A continuación le voy a leer una serie de afirmaciones extraídas de razonamientos que se suelen hacer con frecuencia. Para cada una de estas afirmaciones debe decirme si está totalmente de acuerdo, bastante de acuerdo, más bien en desacuerdo o totalmente en desacuerdo.

	De acuerdo		En desacuerdo		
	Totalmente	Bastante	Más bien	Totalmente	No sé
— Cuando hay escasez de trabajo, los hombres deberían tener más derecho que las mujeres a conseguir un empleo	1□	2□	3□	4□	5□
— Cuando hay escasez de trabajo, debería ser obligatorio jubilarse antes de la edad establecida	1□	2□	3□	4□	5□
— Cuando hay escasez de trabajo, los empleadores deberían dar preferencia a los ciudadanos del país por encima de los inmigrantes	1□	2□	3□	4□	5□
— No es justo dar trabajo a los minusválidos mientras las personas sin minusvalías físicas no logran encontrar un empleo	1□	2□	3□	4□	5□

En cambio, una pregunta en términos relativos es aquella cuya respuesta nace de una comparación con las otras respuestas posibles. Consideremos por ejemplo la pregunta sobre los «problemas más importantes del país». Ésta puede presentarse en términos absolutos: «A continuación le voy a leer una serie de problemas de nuestro país. Para cada uno de ellos, conforme los vaya leyendo, deberá decirme si lo considera un problema muy importante, importante, poco importante o nada importante». A continuación se presenta una lista de problemas como el desempleo, la inflación, la delincuencia, la corrupción, etc. Esta misma pregunta puede plantearse también en términos relativos: «A continuación le voy a leer [...], etc. De estos problemas, indíqueme cuáles cree que son los tres más importantes».

Otro ejemplo sería una batería de preguntas orientadas a conocer la popularidad de ciertos personajes públicos. Con el planteamiento en términos absolutos se puede presentar al entrevistado una lista de políticos y pedirle que haga una evaluación de cada uno de ellos (por ejemplo, con una puntuación del 1 al 10). Para el planteamiento en términos relativos se podría pedir al entrevistado que ordene a los políticos de la lista según su preferencia, o que elija a los tres (o a uno, o dos) que más le atraigan.

La pregunta formulada en términos relativos, que conlleva una comparación, suele ser más adecuada para valorar el punto de vista del entrevistado, puesto que permite distinguir entre sus posturas. Por ejemplo, la mayoría de las personas segu-

ramente estará a favor de reducir los impuestos y de aumentar las prestaciones asistenciales del Estado. Sin embargo, si se les obliga a hacer una elección es probable que establezcan prioridades, revelando su opinión con mayor precisión. Hay que añadir, no obstante, que las preguntas que implican comparaciones entre una serie de estímulos propuestos son más difíciles de tratar desde el punto de vista de los cálculos estadísticos, por lo que se suele utilizar con menos frecuencia.

Las baterías de preguntas son especialmente frecuentes en el procedimiento de investigación social conocido como técnica de las escalas (*scaling*), que se aplica en particular para la denominada «medición de actitudes» (por ejemplo, las preguntas 3 y 4 de la tabla 5.1 tienen por finalidad registrar la actitud del entrevistado en relación con su trabajo y la discriminación laboral). En el caso de las escalas, de los cuatro objetivos mencionados para la organización de las preguntas en batería, el más importante es el cuarto, es decir, condensar las respuestas en un resultado único, que permita la operacionalización de esa actitud particular que se quiere estudiar. En el próximo capítulo profundizaremos sobre esta cuestión.

Por último, conviene destacar que la ventaja más importante de las baterías de preguntas, que consiste en la posibilidad de comprimir una serie de estímulos diversos en un tiempo y un espacio reducidos, conlleva también un cierto riesgo. El tono insistente de la batería y la repetición del esquema de respuesta pueden provocar fácilmente dos de los errores de respuesta más comunes: las respuestas emitidas al azar (o pseudo opiniones) y las respuestas que se repiten de manera mecánica *(response set)*. Por tanto, es conveniente que el investigador y el entrevistador presten una atención especial a la hora de formular y proponer este tipo de preguntas.

5.7. MODALIDADES DE RECOPILACIÓN DE DATOS

Existen tres modos de realizar un cuestionario: la entrevista cara a cara, la entrevista telefónica y el cuestionario autocumplimentado (dentro de éste, la técnica más importante es el cuestionario postal).

En Estados Unidos la forma más común de realizar cuestionarios fue durante décadas (al menos en el periodo que abarca desde principios de los años treinta hasta a los años setenta), la entrevista cara a cara. Lo normal era que el entrevistador se desplazara hasta la vivienda del entrevistado, se sentara frente a él y le planteara las preguntas. A partir de los años ochenta, varios fenómenos simultáneos, como la difusión del teléfono, que en los países occidentales llegaba ya a todos los hogares, el rápido incremento del coste de las entrevistas, una cierta saturación de los ciudadanos, y la creciente desconfianza hacia los extraños debido al aumento de la delicuencia, llevaron a las agencias de sondeos a sustituir, al menos como procedimiento estándar, las entrevistas cara a cara por las entrevistas telefónicas.

Junto a estas dos modalidades siempre ha habido espacio para los cuestionarios autocumplimentados. Dentro de esta categoría se incluyen los cuestionarios enviados por correo, los entregados directamente a los sujetos interpelados (visitantes de un museo, clientes de grandes almacenes, etc.) y los que se aplican a un grupo de personas (clases escolares, departamentos de una empresa, etc.).

El modo de aplicar un cuestionario condiciona en parte el instrumento utilizado. Aunque lo dicho en apartados anteriores vale para todas las modalidades citadas, no hay que olvidar que, desde los primeros pasos de la concepción y preparación del cuestionario, el investigador debe tener las ideas muy claras sobre el procedimiento que va a seguir, ya que las soluciones a los diversos problemas pueden variar según la modalidad de recopilación de datos prevista. Además, cada una de las tres modalidades presenta problemas específicos, como vamos a ver en las páginas siguientes.

5.7.1. Entrevistas cara a cara

En la entrevista cara a cara el entrevistador desempeña un papel fundamental. La calidad de la entrevista depende en gran parte de su actuación y del modo en que plantea y conduce la relación con el entrevistado.

Ya hemos hablado del problema de la interacción entre el observador y el fenómeno observado, y hemos visto que representa uno de los mayores dilemas de la encuesta por muestreo. Recordemos que la técnica que estamos presentando en estas páginas es el cuestionario estandarizado, que se sitúa dentro de la corriente objetivista. En efecto, estamos hablando, en el caso más típico, de una investigación con al menos un millar de entrevistas y varias decenas de entrevistadores, lo cual significa que investigador y entrevistador no coinciden. Es imposible que los propios investigadores realicen las entrevistas, ya que en la mayor parte de las investigaciones se requiere una multitud de ellos. Por ejemplo, para 2000 sujetos, si cada entrevistador realiza 50 entrevistas, se necesitarán 40 entrevistadores.

Por tanto, es necesario reducir el efecto del entrevistador, estandarizando su comportamiento y precisando lo más posible su actuación durante la realización de las entrevistas, para lo cual se requiere una preparación previa del mismo. Nadie tiene la menor duda de que en investigaciones de gran magnitud el entrevistador debe *evitar cualquier comportamiento que pueda influir en el entrevistado.* Por tanto debe evitar expresar aprobación o desaprobación sobre las afirmaciones manifestadas por el sujeto y acostumbrarse a responder con frases o gestos neutrales cuando el sujeto que responde manifiesta el deseo de obtener alguna reacción por su parte. Al mismo tiempo, debe conseguir que el encuestado colabore de manera positiva, que no decaigan la atención y el interés del mismo, que entienda siempre el significado de las preguntas, que no se produzcan malentendidos y errores graves en las respuestas. De ello se deduce que el entrevistador desempeña un papel crítico. Desde el punto de vista de la relación interpersonal, deberá mostrar una actitud amistosa pero a la vez neutral, con una cierta permisividad que logre transmitir de manera implícita al entrevistado la idea de que no hay respuestas correctas o erróneas, que en el mundo existe una amplia variedad de opiniones y comportamientos, y que todos son igualmente aceptables.

Para conseguir estos objetivos los entrevistadores deben presentar una serie de rasgos que enumeramos a continuación:

- *Características.* Los manuales de investigación social dan mucha información sobre las características físicas de los entrevistadores, desde el género (se pre-

fiere a las mujeres), hasta el estatus social (que debe ser superior al de los entrevistados, pero no demasiado), la edad (preferiblemente intermedia), los estudios (de grado medio) y, en los países multirraciales, la raza. También se sugiere que sean personas aptas para un trabajo sin grandes aspiraciones profesionales, discontinuo en el tiempo, a tiempo parcial, con una retribución discreta, que se realiza en horas sueltas más que en las horas centrales del día. Un perfil ideal sería el de la mujer de mediana edad, casada, con estudios de grado medio, ama de casa y de clase media. Los manuales hablan incluso del aspecto externo del entrevistador: dado que en estos casos, la primera impresión puede ser determinante, se recomienda que dicho aspecto sea neutral, ni llamativo ni excéntrico.

No nos parece necesario dedicar demasiada atención a este tema, ya que puede resolverse simplemente con un poco de sentido común (y también porque las características del entrevistador ideal dependen mucho de las características de la muestra a entrevistar).

• *Expectativas.* Si el entrevistador sabe que el entrevistado está afiliado a un partido determinado, esperará que responda de manera coherente a las preguntas de contenido político; si percibe que el encuestado está poco informado, tenderá a plantear las preguntas con un tono apresurado, aceptando sin vacilar las respuestas evasivas y los «no sé». De forma análoga, si el entrevistador está convencido de que una pregunta es difícil, obtendrá probablemente un número elevado de respuestas «no sé», provocado por una transmisión de desconfianza involuntaria a los entrevistados y un menor esfuerzo por obtener una respuesta no evasiva.

Con frecuencia, estas expectativas se transmiten de forma inconsciente a la dinámica de la entrevista. Pueden transmitirse por el tono de voz del entrevistador, sus gestos faciales (miradas de incredulidad, asentimientos con la cabeza, etc.), énfasis al leer las preguntas o las opciones de respuesta. Todos estos factores son especialmente influyentes en el caso de los entrevistados indecisos, inseguros, sensibles a la tentación de complacer al entrevistador, sobre todo en las preguntas para las que no tienen una opinión clara.

• *Preparación.* De todo ello se deduce la importancia de la preparación del entrevistador. En primer lugar se le debe informar del tipo de interacción existente entre él y el entrevistado, de los mecanismos de aquiescencia e influencia inconsciente que suelen crearse. En segundo lugar, dado que el entrevistador no debe ser un autómata inerte e inactivo, sino que debe intervenir ante posibles malentendidos, peticiones de aclaraciones, etc., es necesario darle instrucciones detalladas sobre cómo debe comportarse en estos casos, a fin de limitar su variabilidad subjetiva e interpretativa y evitar que se den indicaciones discordantes. Los cuestionarios suele incluir instrucciones para los entrevistadores. Por ejemplo: «¿durante cuántos meses después de acabar los estudios estuvo buscando trabajo hasta encontrar un empleo estable?» [Para el entrevistador: contar los meses efectivos excluyendo el periodo de servicio militar]. En todo caso, se deberán dar instrucciones escritas, organizar reuniones preparatorias previas y reuniones de control durante la encuesta, y prever la intervención de supervisores a los que los entrevistadores puedan recurrir si es necesario.

• *Motivaciones.* Por último, es necesario tener en cuenta, precisamente por esos mecanismos de interacción invisible e influencia inconsciente a través de gestos,

miradas, tonos de voz, la disposición psicológica del entrevistador con respecto a su trabajo y la entrevista en cuestión. Una actitud pasiva, hastiada y desmotivada por parte del entrevistador tendrá una repercusión negativa en el entrevistado. Es necesario, por tanto, que el entrevistador esté convencido de la importancia de su trabajo y de la investigación, es decir, debe estar implicado en los objetivos de la encuesta. De ahí la importancia de que durante las reuniones preparatorias previas para los entrevistadores no se traten sólo las cuestiones prácticas, sino que se dedique también tiempo y atención a la exposición de los objetivos de la investigación.

5.7.2. Entrevistas telefónicas

Como ya hemos apuntado, en los últimos años las entrevistas telefónicas se han hecho cada vez más populares, hasta llegar a constituir, al menos en los países donde el teléfono llega a un porcentaje de la población muy elevado, la modalidad más difundida para la realización de cuestionarios. Las ventajas de las entrevistas telefónicas pueden resumirse brevemente en los siguientes puntos:

- Permiten realizar la encuesta con gran rapidez: con una buena organización es posible realizar una encuesta a nivel nacional sobre una muestra de dimensiones considerables (más de 1000 casos), en un solo día; en ciertas situaciones esto puede ser fundamental para la encuesta (por ejemplo, cuando se quiere valorar la reacción de la opinión pública ante un suceso de actualidad).
- Conllevan costes muy inferiores en comparación con la entrevista cara a cara: según las estimaciones, permiten un ahorro del 50 al 75 % [Groves, 1989, 526-538; Klecka y Tuchfarber, 1978].
- Provocan menos resistencia a la realización de la entrevista por parte del encuestado y permiten mayor garantía de anonimato: el teléfono entra directamente en los hogares, y hace posible acceder a una cuota de entrevistados que, por desconfianza, nunca dejarían entrar en su casa a un extraño; además, aunque en realidad no es anónima (el entrevistador siempre puede identificar el número de teléfono), el entrevistado percibe en ella una mayor garantía de anonimato en comparación con la entrevista cara a cara.
- Con el mismo coste permiten cubrir también a los entrevistados que residen en las zonas periféricas del país, y por tanto no requiere, como sucede con las entrevistas cara a cara, una concentración territorial de la muestra en áreas más fácilmente accesibles para los entrevistadores.
- Facilitan enormemente el trabajo de preparación de los entrevistadores y su supervisión. De hecho, estas entrevistas están centralizadas, por lo que los entrevistadores pueden formar un grupo cohesionado, localizado en un lugar concreto y fácilmente accesible para tareas de formación y supervisión. Además, mediante el uso de una centralita se puede controlar la marcha de las entrevistas.
- Permiten utilizar directamente el ordenador en la fase de recopilación de datos. El uso del teléfono suele ir asociado con el uso del ordenador (CATI,

computer assisted telephone interviewing): el entrevistador lee en la pantalla el texto de la entrevista, transcribe la respuesta directamente con el teclado y la almacena en la memoria, el ordenador gestiona el curso de la entrevista (por ejemplo, en las preguntas «embudo» o condicionadas pasa de manera automática a la pregunta correspondiente, sin que el entrevistador tenga que seguir las «flechas» sobre el cuestionario), y señala automáticamente las incongruencias evidentes. De esta forma, la entrevista se desarrolla con mucha más rapidez en comparación con la entrevista cara a cara, se reducen notablemente las posibilidades de error por parte del entrevistador y durante la fase de recopilación de datos pueden hacerse análisis elementales sobre los datos introducidos (por ejemplo, controles de frecuencia para corregir la muestra, comunicaciones de datos provisionales al cliente, etc.).

De forma análoga podemos resumir brevemente las desventajas de la entrevista telefónica:

- Al faltar el contacto personal, el entrevistado se siente menos implicado en la entrevista, de ahí la mayor incidencia de respuestas superficiales, emitidas al azar y apresuradas; el propio entrevistador siente esa mayor distancia y frialdad de la relación (en comparación con la entrevista cara a cara), y ello puede hacer que la entrevista se realice con menos cuidado.
- También debido a la falta de contacto personal, la relación con el entrevistado se deteriora más rápido, por lo que la entrevista telefónica debe ser mucho más corta que la entrevista cara a cara (unos 20 minutos para la primera frente a los 40 de la segunda).
- Resulta imposible usar material visual para la entrevista, como fichas, fotografías o dibujos que reproduzcan las opciones de respuesta cuando éstas son complicadas.
- El entrevistador no puede ver al entrevistado, ni su casa, y por tanto no puede registrar datos no verbales (por contraste con la entrevista cara a cara, en la que se le pide que anote el tipo de vivienda, el contexto de la entrevista, la clase social de la familia en función del mobiliario, etc.).
- Debido a la difusión desigual del teléfono, algunos sectores sociales más desfavorecidos quedan excluidos de este medio.
- Los ancianos y las personas con pocos estudios suelen estar infrarrepresentados en las entrevistas telefónicas, ya que en la familia siempre habrá un miembro más joven y con más formación que sustituya al anciano en apuros («de paso con mi marido o mi hijo»), aunque las instrucciones del entrevistador indiquen claramente que el entrevistado debe ser aquél [10]. En cambio, en la entrevista cara a cara, con insistencia y paciencia, un buen entrevistador puede conseguir que lo reciba incluso la anciana más indecisa.

[10] No se entrevista directamente a la persona que responde al teléfono, sino que se aplican diversos sistemas para escoger al azar a una de las personas de la familia (por ejemplo, «el miembro de la familia que cumple años antes», o similar).

- La limitación del tiempo disponible obliga a simplificar las preguntas, a formularlas de manera sintética y en última instancia elemental; además, por la ausencia de contacto personal, es difícil implicar a los entrevistados y obligarlos a colaborar más allá de las respuestas formales a preguntas formales.

De todas las limitaciones mencionadas, las más graves son la primera (falta de contacto) y la última (presión derivada de la limitación del tiempo). El ritmo apremiante no deja tiempo para la reflexión, la pregunta breve demanda implícitamente respuestas con las mismas características, el entrevistado siente que se le pide una respuesta breve, clara e inmediata, no hay lugar para la duda. En la comunicación telefónica no hay tiempos muertos y las pausas de silencio resultan embarazosas (el entrevistado no puede utilizar señales no verbales para que el entrevistador comprenda que está pensando en la respuesta).

Por otra parte, la ausencia de contacto personal en la entrevista telefónica hace que la relación sea más formal, menos comprometida; el entrevistado puede verse impulsado a responder de manera burocrática o al azar; es más fácil confesar que no hemos entendido a una persona que está sentada delante de nosotros que a una voz lejana que está esperando una respuesta para poder pasar a la siguiente pregunta. En una palabra, la presión por responder (sobre el entrevistado) y la ansiedad generada por el paso del tiempo (en el entrevistador) condicionan enormemente la dinámica de la entrevista telefónica y justifican la opinión generalizada de que las respuestas dadas por teléfono son más superficiales que las obtenidas en una entrevista cara a cara.

La entrevista telefónica obliga a simplificar los asuntos a tratar y exige preguntas cortas y respuestas con un número de opciones reducido. La imposibilidad de realizar entrevistas largas (por cuestiones de tiempo) o profundizar demasiado en los temas (por cuestiones de intensidad de la relación), hace que este instrumento sea poco adecuado para los casos en que la investigación requiere información detallada y profunda sobre los sujetos. Por otro lado, su instantaneidad, la inmediatez en la obtención de los resultados, la facilidad para acceder a grandes muestras de la población, la posibilidad de reunir en una sola encuesta temas dispares (pensemos en los estudios de mercado) y el consiguiente abaratamiento de los costes, hace que este instrumento sea muy adecuado para estudios a gran escala.

5.7.3. Cuestionarios autocumplimentados

A continuación vamos a hablar del cuestionario que el sujeto cumplimenta sin la presencia de un entrevistador. Pensemos en los cuestionarios que se reparten a los estudiantes de una clase, las personas reunidas en una asamblea, los usuarios de un servicio determinado, o en los cuestionarios que un encuestador distribuye entre las familias y pasa a recogerlos después de una semana, o el caso frecuente del cuestionario postal.

Una ventaja evidente de esta técnica es el enorme ahorro que supone para los costes de la investigación. Pensemos en un cuestionario dirigido a los visitantes de un acontecimiento (un festival, una exposición, una feria, etc.), que se reparte a la

entrada y se recoge a la salida: en un día un solo encuestador podrá recoger varios cientos de cuestionarios. Con entrevistas cara a cara habría sido necesario emplear a una decena de entrevistadores durante varios días para llegar a esta cifra.

La principal limitación de esta técnica también resulta evidente. Mientras que en los casos anteriores el cuestionario lo cumplimenta el entrevistador, que ha sido instruido para esa tarea y ha tenido tiempo para aprender la manera correcta de realizarla, y por tanto difícilmente cometerá errores, en el caso del cuestionario autocumplimentado no tenemos esta certeza. Entre los entrevistados podría haber alguna persona que no sepa exactamente lo que es un cuestionario, que nunca antes haya cumplimentado uno y que por tanto no sepa muy bien qué hacer con él, habrá personas con niveles de estudios muy dispares y probablemente muchos de ellos estarán poco motivados para hacerlo o hacerlo bien. Ello implica, por tanto, que los cuestionarios autocumplimentados deberán ser cortos, concisos y lo más sencillos posible.

Una segunda limitación estriba en la autoselección de los sujetos que responden. En el ejemplo anterior de los visitantes de un acontecimiento, del público al que se dirige el investigador sólo una pequeña parte devuelve el cuestionario cumplimentado. Esto no sería un problema si los sujetos que devuelven el cuestionario formaran una muestra aleatoria de la población objeto de estudio (todos los visitantes). Sin embargo, se trata de un segmento particular de esa población, constituido probablemente por las personas motivadas, la que tienen más estudios, quizás las más jóvenes, etc. La consecuencia de ello es que no sabemos en qué medida los datos obtenidos son generalizables a toda la población que queremos estudiar.

Podemos distinguir dos tipos de cuestionarios autocumplimentados: la encuesta de grupo y la encuesta individual. Como ejemplo de la *encuesta de grupo,* podemos pensar en un cuestionario para estudiantes que se reparte en clase en presencia de un encuestador, que entrega los cuestionarios, imparte las instrucciones, resuelve posibles dudas y recoge los cuestionarios cumplimentados. Esta variante permite reducir en gran medida las dos limitaciones antes mencionadas. La presencia del entrevistador durante la cumplimentación hace menos probables los errores más llamativos; además, al controlar que todos los cuestionarios repartidos son cumplimentados se evita el riesgo de autoselección. En general, cuando es posible realizar una encuesta de grupo, el cuestionario autocumplimentado es una técnica recomendable. Sin embargo, en investigación social, al margen del ejemplo escolar, las situaciones con estas características poco frecuentes.

En cuanto a la *encuesta individual,* hay que distinguir entre situaciones con y sin obligación de devolución del cuestionario cumplimentado. Un caso de devolución vinculada es el censo: un encuestador entrega el cuestionario a cada familia y pasa a retirarlo una semana después ya cumplimentado. De esta forma se resuelven dos problemas antes mencionados: los errores de cumplimentación se evitan mediante un control rápido de las respuestas por parte del encuestador en el momento de retirar el cuestionario; la autoselección también queda descartada, dado que la entrega del cuestionario se hace obligatoria en el momento en que el encuestador vuelve a retirarlo.

Esta técnica de recopilación de datos no supone un gran ahorro respecto al proce-

dimiento de la entrevista cara a cara (sigue exigiendo la presencia de un encuestador para cada entrevista, aunque el contacto puede reducirse a pocos minutos), sin embargo, es muy utilizada por los institutos nacionales de estadística y los censos (de población, industria y agricultura), así como para otras investigaciones, sobre todo cuando es necesario recopilar información muy detallada y dispersa en un plazo de tiempo delimitado. Por ejemplo, en los estudios sobre el empleo del tiempo y los hábitos de consumo de las familias, en vez de pedir al entrevistado un esfuerzo de memoria para explicar en qué empleó su tiempo y su dinero la semana pasada, se le deja una ficha-diario en la que él anota durante los días siguientes sus actividades o consumos, lo que permite obtener datos libres de imprecisiones o errores de memoria.

En cuanto a la modalidad de cumplimentación individual sin obligación de devolución del cuestionario, es interesante detenerse en el cuestionario postal. Consiste en enviar por correo un cuestionario a una lista de sujetos representativos de la población estudiada, adjuntando una carta de presentación de la investigación, el cuestionario a cumplimentar y un sobre prefranqueado para su devolución al organismo que realiza la investigación.

A continuación enumeramos muy brevemente las ventajas y desventajas de esta técnica. Las ventajas son:

- Ahorro en los costes, ya que los únicos costes de la investigación son la impresión y los gastos de envío.
- Mayor flexibilidad para responder por parte del entrevistado, que puede cumplimentar el cuestionario cuando mejor le venga.
- Mayor garantía de anonimato en comparación con la entrevista cara a cara, dado que el cuestionario cumplimentado no debe contener elementos que permitan identificar al sujeto que lo ha realizado.
- Ausencia de distorsiones debidas al entrevistador.
- Posibilidad de acceder a sujetos residentes en zonas muy distantes del centro de recogida de datos o en zonas aisladas.

Las desventajas son las siguientes:

- Bajo porcentaje de respuestas, a menudo muy por debajo del 50 %, en parte por falta de una petición explícita del entrevistador para conceder la entrevista.
- Distorsión de la muestra debido a la autoselección: no tenemos garantía de que quienes responden son una muestra aleatoria de la población total, y podría decirse que por lo general no lo son, ya que los sujetos que responden suelen ser distintos de los que no lo hacen (por ejemplo, tienen un nivel de más estudios más alto, etc.).
- El nivel de estudios de la población estudiada debe ser medio-alto, puesto que deben ser personas que estén familiarizadas con la comunicación escrita.
- Falta de control sobre la persona que cumplimenta el cuestionario: por ejemplo, podría responder un familiar, o su secretaria, en lugar del destinatario.
- Imposibilidad de utilizar cuestionarios complejos: las preguntas deben ser

sencillas y la cumplimentación del cuestionario debe ser lineal (hay que evitar, por ejemplo, las preguntas filtro, del tipo «si ha respondido "sí" pase a la siguiente pregunta; si ha respondido "no", pase a la pregunta 25», etc.).
- Extensión del cuestionario: no puede superar ciertos límites, que son muy inferiores a los de la entrevista cara a cara.

El mayor problema de esta técnica reside en la devolución de los cuestionarios. No es fácil superar el 50 % de respuestas. Para Babbie [1979, 335] un 50 % de respuestas es «adecuado», un 60 % es «bueno» y un 70 % o superior «muy bueno», pero advierte que se trata de opiniones limitadas, y recuerda que es mucho más importante tener una muestra no sesgada (aunque como ya hemos apuntado no es fácil asegurar la falta de distorsión), que una muestra numerosa [11].

La cantidad de respuestas recibidas depende de varios factores. Podemos mencionar cuatro:

- La institución que patrocina la investigación, el crédito, el prestigio y el grado de conocimiento de la misma; por ejemplo, es mejor una institución pública y conocida como la universidad que un instituto de investigación desconocido. También es importante quién firma la carta de presentación: si la investigación la ha encargado el ayuntamiento de una gran ciudad y va dirigida a una muestra de ciudadanos, una carta firmada por el alcalde puede ser más eficaz que una carta firmada por el director de la investigación.
- La extensión del cuestionario, la presentación, la facilidad o dificultad de las preguntas: un cuestionario largo, complicado y difícil puede desalentar al destinatario desde el primer momento.
- Las características de los entrevistados: se obtiene un mejor resultado cuando la población estudiada está constituida por un segmento específico con alguna característica en común (miembros de una asociación, licenciados universitarios, pertenecientes a una determinada categoría profesional, etc.).
- La solicitud de respuestas. El procedimiento de solicitud de respuestas más completo consta de cuatro fases: *a*) envío del cuestionario con carta adjunta; *b*) envío de la carta de solicitud (puede ser una postal); *c*) envío de una segunda carta de solicitud acompañada por una segunda copia del cuestionario (por si se hubiera perdido la primera), y *d*) llamada telefónica para solicitar el envío del cuestionario cumplimentado [12].

Para concluir esta parte sobre los cuestionarios telefónicos y autocumplimentados hay que decir que en ambos casos es necesario *evitar en lo posible las preguntas*

[11] Dillman afirma que las tasas de respuesta en las investigaciones analizadas por él oscilan entre el 60 % y el 75 % [1978, 51].

[12] Puede

se puede explicar que los cuestionarios se hacen de forma anónima).

abiertas. En el caso del teléfono, la limitación del tiempo disponible impone a la entrevista un ritmo apremiante que no deja margen para las preguntas abiertas. En el cuestionario autocumplimentado, debido a la ausencia del entrevistador, no se cumple uno de los requisitos previos para la pregunta abierta, que sólo tiene sentido cuando hay una relación interactiva con el entrevistador, que puede guiar, estimular y registrar los datos. En ausencia del entrevistador, la pregunta abierta podría interpretarse de manera errónea y transcribirse de manera incompleta.

5.7.4. Entrevistas informatizadas

Abordaremos muy brevemente las entrevistas que utilizan el ordenador en la fase de recopilación de datos [13]. Ya hemos mencionado el CATI, un sistema que utiliza el ordenador para la realización de entrevistas telefónicas. Otra técnica es el CAPI *(Computer assisted personal interviewing),* que no difiere mucho de la entrevista cara a cara, y en la que el entrevistador, en lugar de tener el cuestionario impreso en papel, lee las preguntas en un ordenador portátil e introduce las respuestas en el mismo. Esta forma de recoger los datos permite eliminar algunos pasos entre la recogida y la elaboración de los mismos, reduciendo el tiempo que transcurre entre ambas operaciones.

Con esta técnica desaparecen las fases de codificación e introducción de los datos (que se producen en el contexto de la entrevista); al final de la jornada laboral, el entrevistador puede enviar por correo electrónico un archivo con las entrevistas al centro de recogida de datos, que puede ir elaborando los datos a medida que son recogidos. A esta ventaja se añade la agilización en el uso de cuestionarios complejos (preguntas condicionadas, preguntas planteadas a subconjuntos de la muestra seleccionados al azar, secuencias de respuestas variadas de forma aleatoria, control de coherencia entre las respuestas, etc.), gracias a las posibilidades que ofrece el ordenador de gestionar la entrevista según una programación previa. La inversión inicial por parte de la empresa de sondeos en ordenadores portátiles para los entrevistadores puede amortizarse rápidamente gracias al ahorro que permiten en las fases de codificación e introducción de los datos.

Otra técnica que utiliza el ordenador para la realización de cuestionarios es la *teleentrevista* (o *entrevista telemática*), en la que no es el entrevistador, sino el propio entrevistado quien lee las preguntas en la pantalla e introduce las respuestas en el ordenador. En este caso, el ordenador ya no es sólo un soporte del texto del cuestionario que sustituye al papel, sino un instrumento de interacción entre el centro de recogida de datos y el sujeto entrevistado. En casa del entrevistado se instala una terminal conectada con el centro de recogida de datos, que envía a los entrevistados las preguntas y recibe sus respuestas por correo electrónico.

La ventaja económica más evidente de esta técnica radica en la eliminación de la figura del entrevistador (se trata de hecho de un cuestionario autocumplimentado). A esta ventaja se le añade otra de gran importancia: la posibilidad de realizar en-

[13] Para estudiar el tema con más profundidad remitimos a Saris [1991].

cuestas longitudinales, es decir, encuestas sucesivas a los mismos sujetos en momentos distintos. La instalación en casa del entrevistado de una terminal del centro de recogida de datos permite realizar entrevistas sucesivas modificando el cuestionario, lo que hace posible una observación permanente que puede estar orientada a hacer un seguimiento de los cambios en la opinión pública, de los hábitos de consumo, etc. (por ejemplo, cómo influyen en la popularidad del gobierno sucesos políticos específicos, o cómo influye en el consumo una campaña publicitaria determinada).

En definitiva, esta técnica constituye un instrumento potente (adecuado sobre todo para el estudio de cambios en un periodo de tiempo breve y que hasta ahora se ha usado sobre todo para sondeos de opinión y estudios de mercado), que, sin embargo, no está exento de problemas: el primero son las limitaciones del cuestionario autocumplimentado (no podemos saber qué miembro de la familia responde al cuestionario) y las encuestas longitudinales (saber que lo están entrevistando durante un periodo de tiempo puede alterar el comportamiento del sujeto estudiado; volveremos sobre esto en el apartado 5.9.2).

5.8. FASES PREVIAS A LA RECOPILACIÓN DE DATOS

En este apartado planteamos algunas observaciones sobre las fases previas a la realización de la encuesta propiamente dicha. Se trata de las entrevistas exploratorias preliminares, la prueba del cuestionario, la preparación de los entrevistadores (y su supervisión una vez iniciado el proceso de realización de las encuestas), y los primeros contactos con los sujetos que van a entrevistarse.

El estudio exploratorio preliminar. Para poder formular las preguntas adecuadas, el investigador debe conocer perfectamente el problema objeto de estudio antes de empezar a redactar dichas preguntas. Esto es se aplica sobre todo al caso del cuestionario estandarizado, que, al estar compuesto de preguntas cerradas, exige que el investigador conozca no sólo las cuestiones, sino también todas las posibles respuestas. Según Schuman y Presser «los investigadores debe empezar con entrevistas de preguntas abiertas realizadas a muestras de la misma población objeto del estudio y utilizar sus respuestas para construir las opciones de respuesta propuestas para las preguntas cerradas, de tal forma que reflejen la esencia y las palabras de lo que los sujetos manifiestan de manera espontánea [...]. Esta consideración es tan obvia que hacerla resulta casi embarazoso, si no fuera porque en la práctica se pasa por alto con frecuencia» [Schuman y Presser, 1981, 108].

En esta fase exploratoria preliminar se utilizan diversas técnicas, que van evolucionando desde los primeros instrumentos poco estructurados y marcadamente cualitativos hacia instrumentos cada vez más estructurados. En esta secuencia podemos mencionar las siguientes técnicas: entrevistas libres a observadores escogidos, entrevistas no dirigidas a miembros de la población objeto de estudio, entrevistas con preguntas abiertas (pero estandarizadas) realizadas a la población objeto de estudio, pruebas de preguntas o partes específicas del cuestionario realizadas a familiares, amigos y compañeros de trabajo, para comparar diferentes formulaciones, etc.

En todos estos casos se tratará por lo general de entrevistas y pruebas realizadas por los investigadores, a los que se les plantea el problema de comprender las distintas formulaciones y facetas de los fenómenos a estudiar.

El pretest. Tras la fase exploratoria preliminar viene la redacción del cuestionario. Cuando se tiene la versión casi definitiva del cuestionario, comienza la *fase de prueba* con lo que se conoce como el pretest.

Conviene recordar que una de las características del cuestionario estandarizado es su rigidez, ya que no se puede hacer nada para modificarlo. Cualquier cambio no sólo sería costoso (por la reimpresión de los cuestionarios), sino que además conllevaría la pérdida de las entrevistas ya realizadas. Por tanto, es imprescindible identificar todos los cambios que sean necesarios antes de empezar con la fase de recopilación de datos propiamente dicha: por ello se prevé una especie de «ensayo general», que tiene todas las características de la encuesta definitiva.

Salvo en el caso de investigaciones muy costosas (para las que es recomendable llevar a cabo una encuesta piloto), será suficiente un pretest con varias decenas de casos [14]. Éste se realizará sobre una muestra de sujetos con las mismas características que los individuos de la población estudiada (procurando cubrir una cierta variedad social en cuanto a género, edad, estudios, etc.), y las entrevistas las llevarán a cabo los mismos entrevistadores y bajo las mismas condiciones que la entrevista propiamente dicha (en casa del entrevistado, etc.).

Uno de los objetivos del pretest es determinar la duración de la entrevista: a menudo este cuestionario previo es más largo que la versión final y contiene demasiadas preguntas, algunas de las cuales se eliminarán en función de los resultados de las pruebas. Hay que añadir que si después de estos cambios se obtiene un cuestionario muy distinto del inicial, éste tendrá que probarse de nuevo: como escribe Sheatsley [1983, 228], «es un error común escribir preguntas completamente nuevas, reformular otras, presumir de haber hecho un recorte de 10 minutos eliminando algunas partes, y mandar la nueva versión a la calle sin haberla probado de nuevo».

Preparación y supervisión de los entrevistadores. Ya hemos apuntado que cuando se prevé la participación de entrevistadores (en entrevistas cara a cara o telefónicas), hay que contar también con la formación de los mismos y la supervisión durante la fase de recopilación de datos. La preparación consistirá en una primera reunión del equipo de investigación y los entrevistadores, con el objetivo de presentar la investigación (quién la encarga y sus fines específicos), describir su diseño (muestreo, identificación de los nombres, modalidades de entrevistas, presencia de otros posibles canales de recogida de información, aparte de las entrevistas), y exponer con detalle el instrumento de investigación (la encuesta). Se intentará prever todos los problemas que pudieran surgir durante la

[14] Sheatsley [1983, 226] opina, por ejemplo, que «no son necesarios más de 12-25 casos para darse cuenta de los puntos débiles del cuestionario» y sugiere asignar estas entrevistas de prueba a tres o más entrevistadores (elegidos entre los más expertos del grupo), encargando a cada uno al menos tres entrevistas, y a continuación celebrar una reunión colectiva para comentar los resultados de las pruebas.

realización de las encuestas, a fin de garantizar una línea de comportamiento común a todos los entrevistadores.

Después de esta primera reunión se realizarán las entrevistas de prueba (el pretest), tras lo cual se celebrará un segundo encuentro para debatir sobre el funcionamiento del instrumento [15]. Una vez redactada la versión definitiva del cuestionario comenzará la encuesta propiamente dicha, durante la cual convendrá celebrar otra reunión con los entrevistadores para hacer un seguimiento común de su evolución (a fin de definir líneas de acción comunes para resolver los problemas que hubieran podido surgir). Durante la realización de la encuesta se contará con supervisores que desempeñarán una función de asesoramiento (resolución de problemas surgidos durante la encuesta) y de control (comprobar que las entrevistas se hayan realizado efectivamente, etc.) [16].

El contacto inicial con los sujetos entrevistados y el problema de los rechazos. Probablemente el momento más delicado de toda la entrevista es el primer contacto, cuando el sujeto seleccionado debe tomar la decisión de aceptar que lo entrevisten o no. No procede detenerse en los sentimientos de sospecha, desconfianza o inseguridad que inevitablemente pueden aparecer en su mente. Al tratarse de una decisión rápida, es determinante la forma en que se presenta el entrevistador, su apariencia externa (atuendo, etc.) y las motivaciones que expone (y las argumentaciones que deberá elaborar en caso de duda o rechazo).

En su presentación, el entrevistador deberá preocuparse por dejar claro: *a*) quién ha encargado la investigación; *b*) cuáles son sus objetivos; *c*) por qué se dirigimos precisamente a él/ella. Deberá, además, *d*) destacar la importancia de su colaboración, así como *e*) asegurarle el anonimato de sus respuestas. Si es posible, resulta muy útil que esta presentación se haga también de forma preliminar con una carta enviada al entrevistado unos días antes del primer contacto.

La argumentación más importante para ganarse la confianza del entrevistado será probablemente la relativa al anonimato. Se ha demostrado, por ejemplo, que es más fácil obtener respuestas sobre el comportamiento electoral cuando el cuestionario es claramente anónimo (simulación de la votación mediante la introducción de una ficha autocumplimentada en una urna, similar a las urnas electorales), que en el caso de una entrevista cara a cara.

Otros dos elementos decisivos son la institución que encarga el estudio y la figura del entrevistado. Si, por ejemplo, una investigación la encarga la administración municipal y va precedida de una carta de presentación y solicitud de colaboración firmada por el alcalde de la ciudad, el porcentaje de rechazos será menor que si la encarga un instituto de investigación desconocido. Por ejemplo, se sabe que en España los sondeos realizados por el Instituto Nacional de Estadística tienen una

---

[15] Después de estas reuniones será útil resumir todas las indicaciones en una serie de instrucciones escritas, que el entrevistador podrá utilizar como referencia para las entrevistas.

[16] El hecho de que se controle su labor también puede tener un efecto motivador para el entrevistador. En general se suelen realizar controles aleatorios a los entrevistados, normalmente telefónicos, con la repetición de algunas preguntas clave para comprobar la congruencia con las informaciones recogidas por el entrevistador.

tasa de rechazos muy inferior al resto de sondeos. Lo cual no es de extrañar, ya que las encuestas las realizan los ayuntamientos, que a menudo emplean como entrevistadores a los funcionarios municipales. Estos funcionarios, sobre todo en los pueblos pequeños, suelen ser conocidos por la población, gracias a lo cual los entrevistados no sienten desconfianza hacia ellos [17].

5.9. ANÁLISIS SECUNDARIO Y ENCUESTAS SUCESIVAS

El proceso de la encuesta por muestreo que hemos presentado hasta el momento no es el único posible. Aquí hemos presentado el caso en el que el mismo investigador: *a*) produce los datos, y *b*) los analiza al mismo tiempo (o en cualquier caso en un plazo de tiempo limitado, por ejemplo, varias semanas). En la actualidad, son cada vez más frecuentes las investigaciones realizadas sobre bases de datos preexistentes, así como las encuestas sucesivas.

5.9.1. Análisis secundario

Cuando el científico social planifica su investigación tiende a concebirla como un proceso autónomo y cerrado en el que él mismo produce y analiza los datos sociales. El investigador parte del supuesto (y de la legítima ambición) de que va a afrontar interrogantes nuevos y originales. La opción de afrontarlos con datos recopilados por otros podría parecer limitada. Nada más lejos de la verdad. Es posible hacer una investigación original con datos ya existentes. La investigación sobre la desviación de Sampson y Laub *(Crime in the Making),* de la que hablamos en el segundo capítulo, es un buen ejemplo de ello.

Análisis secundario. Se trata de *una investigación realizada sobre datos de encuesta por muestreo recopilados previamente y disponibles en forma de matriz de datos.* Es por tanto una forma de reanálisis de bases de datos ya existentes. Nótese que limitamos la aplicación del término «análisis secundario» a datos cuya unidad de análisis sea el individuo, y que sean de tipo individual (por ello hablamos de la «matriz de datos»). Hay que distinguir el caso en que los datos se conocen sólo de forma agregada (por ejemplo, sobre una base territorial, como sucede en las fuentes estadísticas oficiales) del caso en que se reanalizan las tablas o los resultados de investigaciones anteriores.

Durante varias décadas la investigación social ha avanzado según el modelo tradicional del investigador que plantea el problema, formula el diseño de la investigación, recoge los datos y después los analiza. En el curso del tiempo este procedimiento se ha ido encontrando con diversos problemas. En primer lugar, el desarrollo de la disciplina ha ido ampliando cada vez más las exigencias del investigador. Por ejemplo, si para los primeros investigadores que estudiaron empíricamente la relación entre el rendimiento escolar y la clase social podía ser suficiente un estudio

[17] Hablaremos con más detalle sobre el tema de los rechazos en el capítulo 8.

realizado en un solo instituto, para los investigadores posteriores, el objetivo de aportar una contribución original planteaba la exigencia de ampliar la base empírica, la posibilidad de hacer comparaciones entre diversos institutos, entre diferentes grupos sociales, entre la ciudad y el campo, entre el norte y el sur del país (y posteriormente también entre países). De forma análoga, si en los inicios incluso una simple distribución de frecuencia univariada o unas tablas bivariadas podían aportar contribuciones originales, con el tiempo se planteó la exigencia de llevar a cabo análisis más sofisticados, que estudiaran el efecto simultáneo de más variables (análisis multivariado), lo que conlleva disponer de muestras de grandes dimensiones. En definitiva, en poco tiempo la exigencia de una mayor sofisticación ha implicado un rápido aumento de los costes de investigación, lo que ha hecho que quede fuera del alcance de la mayoría de los investigadores.

Actualmente existen dos tendencias. Por una parte, resulta evidente que las investigaciones realizadas en el pasado ofrecen diversas posibilidades para seguir profundizando en ellas de manera original e innovadora, por varios motivos. En primer lugar, porque normalmente el análisis de los datos se limita a una parte de lo que éstos, en su conjunto, pueden decir. En segundo lugar, por los continuos avances teóricos de cada disciplina, que conforme pasa el tiempo va formulando nuevos interrogantes. Además han surgido nuevas técnicas estadísticas que permiten responder hoy, con los nuevos instrumentos de análisis disponibles, a interrogantes antiguos. Todo ello ha llevado a los investigadores a retomar bases de datos ya existentes, para reanalizarlas.

Este hecho se ha visto favorecido por el desarrollo de la informática, que gracias a las redes y a los programas comunes permite acceder y utilizar fácilmente bases de datos recopilados en cualquier parte de un país o del mundo y en cualquier momento. También se ha beneficiado de la aparición de los archivos de datos, instituciones dedicadas a la recogida y sistematización de las bases de datos sobre las investigaciones realizadas, a la documentación de las modalidades de investigación, a la uniformización de los criterios de organización de los datos, y a su distribución (a cambio de una modesta aportación económica) para todo aquel que lo solicite [18].

Al mismo tiempo, la comunidad científica se ha organizado en el plano de la recopilación de datos nuevos, promoviendo la creación de agencias cuyo objetivo es reunir los recursos necesarios para esta tarea tan costosa. Los datos se ponen a disposición de todos los investigadores. Así nacieron las encuestas orientadas específicamente al análisis secundario, que no tienen un tema único: en ciertos casos sí se sitúan en un terreno disciplinar específico (el comportamiento electoral, las escalas de valo-

[18] El más importante de estos archivos es el *Inter-university Consortium for Political and Social Research* (ICPSR), de la universidad de Michigan, fundado en 1962, que recopila datos sobre Estados Unidos y también sobre otros países (actualmente dispone de más de 30000 bases de datos); un equivalente europeo lo constituye el *Economic and Social Research Council Data Archive* (ESRC-DA), activo desde 1967 y ubicado en la Universidad de Essex (Reino Unido). También cabe destacar el *Zentral Archive* para la investigación social empírica de la Universidad de Colonia. En esta institución se archivan datos, cuestionarios y códigos internacionales que se pueden solicitar para realizar investigaciones.

res, la movilidad social, etc.); otras veces se trata de las denominadas encuestas multipropósito, que abarcan un amplio espectro de problemas sociales.

Las ventajas de estos avances son evidentes: ahorro en los costes de investigación, garantía de rigor de la investigación cuando se realiza bajo el control de la comunidad científica, posibilidad de realizar estudios de gran envergadura también para los investigadores con pocos recursos. A ello se añade el hecho de que encargar la labor de la recopilación de datos a instituciones sólidas (por ejemplo, ligadas a fundaciones, a universidades, a la comunidad científica, etc.), abre la posibilidad de repetir la misma investigación durante largos periodos, lo que permite incluir la variable tiempo en las investigaciones sociales (tema que trataremos en el siguiente apartado).

Sin embargo, todo ello también conlleva ciertas limitaciones. Dado que el investigador no tiene control alguno sobre el proceso de recopilación de los datos, le resultará casi imposible valorar la *calidad de estos datos.* Podría ser que la documentación sobre modalidades de recogida de datos, características de la muestra, forma de proceder de los entrevistadores o tasas de ausencia de respuesta de los datos recogidos en el pasado fuera insuficiente; además, es posible que ya no fueran perceptibles ciertos errores cometidos en la fase de memorización o tratamiento de los datos. Los datos recogidos por las nuevas empresas de sondeos que se dedican expresamente al análisis secundario deberían estar exentos de estos errores.

El análisis secundario, en general, plantea otras limitaciones serias. La división del trabajo que se establece entre quien plantea las preguntas de la investigación y analiza los datos, por una parte, y quien recoge estos datos, por otra, puede conllevar ciertas limitaciones, en el sentido de que algunas preguntas importantes podrían quedarse sin responder por falta de datos relativos a ellas. Además, puede generar *investigaciones realizadas a partir de datos disponibles,* sin que se formulen hipótesis teóricas previas: «algunos investigadores obtienen un conjunto de datos, aplican las operaciones estadísticas habituales, y luego deciden cuál es el problema a estudiar [...], pero el enfoque de «datos en busca de análisis» lleva sólo a resultados triviales y previsibles» [Kiecolt y Nathan, 1985, 14].

Metaanálisis. Se trata de una propuesta consolidada en los últimos años en países donde la investigación social está más desarrollada (en particular, en Estados Unidos) y ha producido ya una gran cantidad de logros empíricos que plantea con urgencia el problema de nuevas formas de síntesis. El metaanálisis es un método de selección, integración y síntesis de estudios que tratan sobre el mismo tema, a partir de los *resultados* de los estudios y no a partir de los datos, como sucede con el análisis secundario. Se trata, pues, de una «integración de resultados» más que un «reanálisis de los sujetos». Como escribe Glass, que fue quien introdujo el término metaanálisis:

> El *análisis primario* es el análisis original de los datos de una investigación [...]. El *análisis secundario* es el reanálisis de los datos para responder a las preguntas originales con técnicas estadísticas mejoradas, o bien para responder a nuevos interrogantes con datos antiguos [...]. El metaanálisis, en cambio, es un «análisis de análisis» [...], es el análisis estadístico de una amplia colección de resultados deri-

vados de estudios sobre los individuos, para llegar a una integración de los resultados [Glass, 1976, 3].

El metaanálisis aplica procedimientos estadísticos para integrar los resultados de las investigaciones consideradas. Esta técnica se ha desarrollado paralelamente a la consolidación de las bases de datos informatizadas, que permiten, mediante el uso de palabras clave, localizar con cierta facilidad las investigaciones realizadas sobre un tema específico. Glass y otros colegas, por ejemplo, estudiaron con esta técnica la influencia del tamaño de la clase sobre el aprendizaje de los alumnos, después de encontrar unas 80 investigaciones sobre el tema. El análisis evidenció la fuerte correlación, prácticamente lineal, entre los niveles de aprendizaje y el tamaño de la clase, llegando a la conclusión de que «no existe ninguna duda de que, cuando el resto de condiciones se mantienen invariables, se aprende más en las clases con un menor número de alumnos» [Glass, McGraw y Smith, 1987, 42].

Utilizando la misma técnica, Rotton y Kelly [1985] llevaron a cabo una curiosa investigación con la que quisieron poner a prueba la creencia popular de que el ciclo lunar influye en el comportamiento humano. Para ello recopilaron los resultados de 37 investigaciones empíricas sobre el tema, que relacionaban el ciclo lunar con diversos tipos de comportamientos anómalos desde el punto de vista psíquico y social, como ingresos en hospitales psiquiátricos, suicidios o conductas autodestructivas, llamadas a servicios telefónicos de ayuda psiquiátrica, homicidios y otros actos delictivos, etc. Del metaanálisis de los resultados de estas investigaciones resultó que no existe ninguna relación estadísticamente significativa entre el ciclo lunar y los comportamientos mencionados.

5.9.2. Encuestas sucesivas (diacrónicas)

Al introducir el concepto de variable en el capítulo 3, y al hablar de los experimentos en el capítulo 4, destacamos el hecho de que una variable puede «variar», tanto *entre observaciones*, que suceden en el mismo momento, como *en el tiempo*, sobre una misma observación. Por ejemplo, si queremos estudiar la relación que existe entre las variables conservadurismo político y edad, podemos registrar el conservadurismo político de una muestra de sujetos de distintas edades y ver cómo varía al variar la edad, o podemos estudiar una muestra formada por los mismos sujetos a lo largo de un intervalo de tiempo, repitiendo el registro de sus opiniones políticas en varios momentos distintos, para ver cómo varían estas opiniones con el paso de los años.

El primer ejemplo sería un *estudio transversal*, el segundo sería un *estudio longitudinal*. Hasta ahora en este capítulo hemos hablado de manera implícita de encuestas transversales. Pero todas las técnicas tratadas (en particular, la del cuestionario estandarizado) pueden utilizarse también para estudios longitudinales. Por tanto, podemos hablar de *la encuesta longitudinal, cuando la investigación se realiza en varias ocasiones sucesivas sobre los mismos sujetos*. Obviamente, con esta distinción estamos abordando el problema nada despreciable de la introducción de la variable «tiempo» en la investigación social. Por fundamental que sea para las

ciencias sociales el estudio del cambio, éste es muy difícil de afrontar con los instrumentos de los que disponemos en la actualidad. La encuesta por muestreo, por el hecho de ser una especie de fotografía instantánea tomada en un momento concreto, se presenta como un instrumento particularmente inadecuado bajo este punto de vista. La encuesta longitudinal [19] constituye una posible solución a este problema.

Encuestas longitudinales. La técnica de la encuesta longitudinal, es decir, la entrevista que se repite en varias ocasiones sobre los mismos sujetos, la propuso Lazarsfeld [1948] a los investigadores sociales en los años cuarenta, después de haber experimentado con una muestra de electores de Ohio durante seis meses, en la campaña para las elecciones presidenciales de 1940, para estudiar los efectos de la propaganda electoral. El término inglés con el que se designa este tipo de encuesta es *panel* [20] (por lo que a veces se habla de *encuesta panel* o *panel*).

Las encuestas longitudinales plantean algunos problemas específicos. En primer lugar está la cuestión de la «mortalidad» de la muestra, es decir, la reducción del tamaño de la muestra que se registra inevitablemente en cada nueva «oleada» (en cada nueva fase de recopilación de datos), por los motivos más dispares (rechazos, traslados, muertes, etc.). Además, las encuestas previas pueden influir en las posteriores: ya hemos recordado la influencia de la memoria y del aprendizaje que surge cuando se replantean las mismas preguntas; también podemos recordar el hecho de que el sujeto estudiado, al sentirse objeto de una atención especial, puede modificar su conducta normal, por ejemplo, podría estar más informado sobre la política cuando sabe que le van a hacer preguntas al respecto, etc.).

La repetición de la encuesta sobre los mismos sujetos no es, sin embargo, la única forma de realizar un estudio longitudinal. Otra solución es la *encuesta retrospectiva* (transversal), que consiste en una encuesta transversal normal (que se realiza, por tanto, una sola vez) en la que a los sujetos se les plantea una serie de preguntas sobre su pasado. Las limitaciones de esta solución son evidentes, ya que confía en el recuerdo y en la veracidad del relato de hechos pasados.

Otra posibilidad sería la *conexión de bases de datos censales y administrativos,* que consiste en relacionar datos procedentes de fuentes distintas, como censos, padrones municipales, registros de los centros de salud locales, etc. Esto permite relacionar datos individuales (propiedades socioeconómicas, acontecimientos demográficos, sanitarios, etc.), de distintas fuentes registrados en momentos diferentes, y observar la variación en el tiempo de datos procedentes de la misma fuente, por ejemplo, variaciones de tipo individual entre dos censos sucesivos. La limitación de

este método reside en la pobreza de estos datos (que se reducen fundamentalmente a lo que hemos denominado propiedades sociodemográficas básicas).

Encuestas transversales repetidas. Una manera de incluir el factor tiempo en la investigación, sin entrevistar a las mismas personas, son las *encuestas transversales repetidas* [21], que consisten en registrar la misma información en momentos diferentes, pero no sobre los mismos sujetos, sino sobre muestras de individuos distintas. Obviamente debe cumplirse la condición de que dichas muestras sean comparables entre sí (el diseño de la muestra debe ser idéntico en las distintas encuestas) para garantizar que las posibles diferencias surgidas entre las encuestas se deban a cambios reales en la población y no a variaciones en la composición de la muestra.

Con este tipo de encuestas se pueden captar cambios de tipo agregado (por ejemplo, cómo cambian los valores religiosos o los hábitos de consumo de una población), pero no cambios individuales; por ello sigue siendo muy difícil analizar las causas del cambio, que se pueden identificar más fácilmente cuando la variación se registra de manera individual (con encuestas longitudinales) y, por tanto, puede relacionarse con otras variables individuales. Un ejemplo clásico de encuestas transversales repetidas lo constituyen los *National Election Studies* que se realizan en muchos países en cada cita electoral sobre muestras representativas de electores, utilizando cuestionarios que contienen prácticamente las mismas preguntas: ello permite observar la variación de la opinión pública en el tiempo. En España, encontramos ejemplos de este tipo en algunas de las series de barómetros del CIS, como, por ejemplo, la serie de expectativas económicas, o la serie de opinión pública y política fiscal, o la serie de latinobarómetro, etc. (ver también el siguiente apartado) [22].

El problema general de todos los diseños de investigación que incluyen la variable «tiempo» radica en su coste: para un estudio efectivo sobre el cambio no basta sólo con recopilar datos en dos ocasiones, sino que es indispensable repetir la encuesta durante un periodo de tiempo largo (que podría llegar a durar décadas). De ahí la necesidad de confiar este tipo de estudios a instituciones que puedan garantizar esa continuidad a largo plazo inviable para investigadores individuales o grupos de investigación aislados.

5.10. BASES DE DATOS EN ESPAÑA Y OTROS PAÍSES

En este apartado presentamos algunos de los numerosos *data set* de series de encuestas disponibles en España y en otros países (sobre todo en Estados Unidos y la Unión Europea). Estos estudios se llevan a cabo mediante entrevistas realizadas a los mismos individuos en distintos momentos (encuestas longitudinales), usando el

[21] Lazarsfeld introdujo, en el artículo mencionado [Lazarsfeld, 1948], la distinción entre estudios de *panel* y estudios *trend* (éstos últimos coinciden lo que nosotros llamamos «encuestas transversales repetidas»).

[22] Conviene destacar que en estas series no siempre se mantienen todas las preguntas en el cuestionario, aunque sí una gran mayoría de ellas.

mismo instrumento o encuesta para entrevistar a distintos individuos en distintos momentos (encuestas transversales repetidas), o una combinación de ambas opciones. Normalmente estas series de encuestas tratan aspectos de contenido social, como los cambios económicos y demográficos, las condiciones de salud y las conductas sanitarias, la enseñanza y la movilidad social, la participación política, y otras actitudes u opiniones. Estas bases de datos representan sólo una fracción de todas las series disponibles en ciencias sociales, por lo que una exploración más detallada de los catálogos de los bancos de datos y las agencias oficiales revelarían muchas más. Las consideraciones que se presentan a continuación pretenden servir como guía genérica al lector.

5.10.1. Estudios transversales repetidos

En España, el *Centro de Investigaciones Sociológicas* (CIS) es uno de los principales organismos públicos dedicado a la realización de estudios sobre la realidad social y política española. Desde 1963 hasta la actualidad ha realiza una media de entre 30 y 40 estudios anuales, la mayoría de los cuales proceden de encuestas y están disponibles en su banco de datos. El CIS cuenta con un catálogo que se puede consultar en la página web de la institución, www.cis.es. Entre los numerosos estudios, existen diversas series de encuestas transversales electorales, como las encuestas preelectorales y postelectorales, de cultura política, o las de coyuntura política y social, los llamados barómetros de opinión pública.

Los barómetros se realizan con una periodicidad mensual o bimensual y recogen información sobre aspectos políticos, económicos y coyunturales. Las preguntas de los cuestionarios de los barómetros no son siempre las mismas, pero en casi todos ellos se suelen incluir la intención de voto, la evaluación de líderes políticos o de la situación económica, así como un buen número de características sociodemográficas. Además, desde 1984 hasta la actualidad, el CIS ha realizado una serie de encuestas anuales sobre opinión pública y política fiscal. En el ámbito sanitario, se han realizado también diversas encuestas de ámbito nacional sobre la salud o la satisfacción del usuario con los servicios sanitarios. Otro ámbito de reciente aplicación, y del que se dispone de datos individuales transversales, es el relacionado con las cuestiones de defensa y el ejército. Desde 1997 y con una periodicidad anual o bianual, el CIS ha elaborado otra serie de estudios bajo el título «Defensa nacional y profesionalización del ejército». En general, cabe puntualizar que en todas estas series se repiten algunas preguntas, pero también se van incorporando preguntas adicionales, que suelen estar más relacionadas con la coyuntura del momento en que se realiza cada encuesta.

La *Fundación Foessa,* que fue creada por Cáritas en 1965, ha publicado también cinco informes sociológicos sobre la situación social en España para los años 1967, 1970, 1975, 1980/83 y 1994 y varias series de estudios monográficos sobre pobreza, género o desigualdad territorial. Los temas principales de estos informes sociólogicos son la situación económica y política, las reformas sociales, los cambios de valores, la desigualdad y la marginación social. Otra serie de encuestas ransversales sobre la realidad social en España fue la producida entre 1990 y 1996

por el *Centro de Investigaciones sobre la Realidad Social* (CIRES). Esta organización privada realizó durante el periodo mencionado encuestas de opinión sobre diversos temas sociopolíticos y para muestras representativas de todo el territorio español.

El *Instituto Nacional de Estadística* (INE) es otra institución pública que dispone de datos individuales, en este caso microdatos, aunque su finalidad principal es la actividad estadística pública y la difusión de datos estadísticos (en el capítulo 7 se explica con más detalle el contenido del banco de datos del INE, que se puede consultar en la página web de la institución, www.ine.es). Los microdatos son los ficheros que contienen los datos individuales de una estadística. Se trata de ficheros en formato ASCII estructurados por campos que contienen, para cada registro individual de la encuesta, los valores de cada variable. Las operaciones estadísticas para las que existen ficheros de microdatos con series disponibles a lo largo del tiempo son muy variadas, por ejemplo, las referentes al movimiento natural de la población, la encuesta de fecundidad, la encuesta de morbilidad hospitalaria, etc.

Entre las encuestas mencionadas, cabe destacar la *Encuesta de población activa* (EPA), que es trimestral y está centrada en las unidades familiares. Se trata de una encuesta que el INE viene realizando desde 1964. Su finalidad principal es obtener datos sobre la población activa y sus diversas categorías (ocupados, desempleados), así como de la población que no está integrada en el mercado laboral (inactivos). En esta encuesta se seleccionan 3484 secciones censales, de entre las más de 30000 en las que está dividida España. En cada una de las secciones se entrevista una media de entre 18 y 22 hogares, dependiendo de la provincia. La muestra inicial es de 64000 familias por trimestre, si bien en la práctica queda reducida a unas 60000 familias entrevistadas de manera efectiva, lo que equivale a unas 200000 personas. Los resultados individuales en formato de microdatos de la encuesta están disponibles desde el tercer trimestre de 1976.

En el contexto latinoamericano, el *Latinobarómetro,* que se realiza cada año en 17 países de América Latina, es una de las encuestas transversales repetidas de referencia internacional, cuyos contenidos se pueden consultar en la pagina web del proyecto, www.latinobarometro.org. El estudio del *Latinobarómetro* comenzó en 1995 en ocho países (Argentina, Brasil, Chile, México, Paraguay, Perú, Uruguay y Venezuela) y se extendió a partir de 1996 a 17 países de América Latina, exceptuando Cuba y la República Dominicana. Gracias a la colaboración con el CIS, esta encuesta también se realiza en España. El cuestionario recoge preguntas de temas variados, tales como democracia e instituciones políticas, economía, integración y acuerdos comerciales, políticas sociales y distribución de la riqueza, desigualdad de género, cultura y participación política, y temas relacionados con el medio ambiente. Además, cada año se incluye un bloque temático, así como algunas preguntas sobre coyuntura. Entre los temas tratados hasta el presente cabe mencionar el comercio y la inversión extranjera, la democracia, los partidos políticos y las instituciones, el fraude y la corrupción, el capital social, la pobreza o el fenómeno de la inmigración.

Los censos que se realizan cada 10 años en Estados Unidos, aunque se suelen incluir en la categoría de las estadísticas oficiales (que se discutirán en el capítulo

7), son también una rica fuente de datos individuales. Un ejemplo de ello son las muestras de microdatos integrados de uso público (*Integrated Public Use Microdata Samples,* IPUMS), cuya pagina web es www.ipums.umn.edu/usa. A diferencia de las tablas tradicionales de estadísticas agregadas sobre diferentes áreas geográficas, estas muestras de microdatos han sido seleccionadas a partir de los registros individuales de personas y hogares inscritos en los censos entre 1850 y 1990. Desde 1960, la Oficina del Censo *(Census Bureau)* ha producido las IPUMS en consonancia con sus esfuerzos de difusión de los datos. Como cabe esperar en series transversales que abarcan 150 años, las variables que aparecen en el censo de cada año van cambiando, pero en las IPUMS se registran de forma invariable características como la edad, el género, el origen étnico, los estudios, la participación en el mercado laboral y la composición del hogar.

Otra serie de microdatos producida por la Oficina del Censo de Estados Unidos es la *Current Population Survey* (CPS) cuya página web es www.bls.gov/cps. Se trata de una serie de encuestas longitudinales que se usa mucho para el estudio de tendencias. La CPS utiliza un diseño que se caracteriza por una muestra de hogares que son entrevistados durante cuatro meses consecutivos y, después de un paréntesis de ocho meses, se los vuelve a entrevistar durante cuatro meses más. Los datos de esta encuestas se suelen tratar como estudios transversales repetidos. Aunque en esencia la CPS es menos extensa que el censo que se realiza cada diez años, esta serie de datos incluye muestras de entre 55000 y 60000 unidades de hogares, suficientes para proporcionar estimaciones nacionales y estatales de las características de la fuerza laboral. Además, incluye otras preguntas relacionadas con el desempleo, los horarios de trabajo, los salarios y otros temas procedentes de cuestionarios complementarios. La CPS se ha venido realizando desde 1943 pero las series con microdatos sólo están disponibles desde 1964.

Las series del *American National Election Study* (ANES), que se inició en 1948, ofrecen datos individuales sobre participación política, valores y actitudes sociales y políticas, percepciones y opiniones sobre líderes y grupos políticos o de evaluación de políticas públicas. El tamaño de las muestras de estos estudios oscila entre 1000 y 2500 entrevistados, y también incluyen características sociodemográficas como la edad, el género, el origen étnico, la religión, la ocupación y el nivel de estudios. Para facilitar la realización de análisis comparados en el tiempo se ha creado también un base de datos acumulativa que contiene variables básicas de las encuestas electorales bianuales desde 1952. Los datos están disponibles en la página web www.electionstudies.org.

Posiblemente una de las encuestas sociales más conocidas es la *General Social Survey* (GSS), que se realiza en Estados Unidos desde 1972 con una periodicidad anual o bianual. El amplio cuestionario de la GSS incluye temas como el empleo, el trabajo, la educación, la familia, las relaciones interraciales, las relaciones sexuales, la religión o la ética. Cada año, la GSS también recoge información sobre un tema específico. Estos módulos exploran más en profundidad cuestiones como la orientación en el trabajo, la religión, el género, la salud mental, el papel del gobierno o las redes sociales. Los datos están disponibles a través del archivo de la Universidad de Michigan, en la página web http://webapp.icpsr.umich.edu/GSS/

Desde 1985 los módulos temáticos de la GSS han adquirido un componente de comparación gracias al *International Social Survey Program* (ISSP), que realiza el mismo cuestionario en varios países. Entre los temas monográficos de los cuestionarios destacan las desigualdades sociales, la distribución de los roles familiares, el papel del gobierno y su intervención en la economía, los valores sociales (nacionalismo, religión, etc.) Para disponer de una descripción más detallada del contenido de los cuestionarios en cada recopilación de datos y las características de las muestras en cada país, se puede consultar la pagina web del proyecto www.issp.org.

Con la intención de medir el nivel de apoyo y la actitud de los ciudadanos con relación al Mercado Común Europeo y las instituciones comunitarias, en 1970 se lanzó una serie de encuestas que con el tiempo se han ido ampliando tanto por sus objetivos temáticos como por la cantidad de países incluidos en el proyecto. Las series del *Eurobarómetro* han ido incluyendo un número cada vez mayor de países hasta cubrir todos los Estados de la Unión Europea (www.europa.eu.int/comm/public_opinion). El *Eurobarómetro* trata temas tan variados como las actitudes hacia las instituciones europeas y el proceso de integración, la energía nuclear, la religiosidad, el género, los roles en el entorno laboral, el racismo, las tecnologías de la información, los problemas medioambientales, las condiciones laborales, etc. Las preguntas del cuestionario común son la edad, el género, el estado civil, el nivel de estudios, la ocupación, los ingresos, el origen étnico, las orientaciones e identidades políticas, la clase social subjetiva, la religión o la composición del hogar.

Otro proyecto europeo de estudios transversales que está en la actualidad en su etapa inicial es la *European Social Survey* (ESS, www.europeansocialsurvey.org). La ESS es un proyecto comparado que se lleva a cabo cada dos años en 24 países europeos, entre ellos España, y que analiza el cambio y la continuidad de las actitudes, atributos y comportamientos sociales y políticos de los ciudadanos europeos.

La *World Values Survey* (www.worldvaluessurvey.org/services/index.html) es otra encuesta transversal que se repite en el tiempo y que incluye a España (y más de 50 países a partir de las fases segunda y tercera). En sus cuestionarios se abordan temas como el significado de la vida y el grado de satisfacción con ésta, los problemas y asuntos sociales cotidianos, la confianza interpersonal, las creencias religiosas, el trabajo, el tiempo libre, etc. Esta encuesta se realizó en diez países de Europa Occidental y se repitió en otros doce países en la primera fase entre 1981 y 1983. La encuesta se repitió y fue ampliando el número de países en los periodos de 1990-1993 y 1995-1997. La cuarta fase se realizó entre 1990 y 2000. Los datos de esta encuesta están disponibles para las entidades miembro del ICPSR de la Universidad de Michigan, o en el Zentral Archive de Colonia, y en el CIS español a través del Archivo de Estudios Sociales (ARCES) [23].

Finalmente cabe destacar el Estudio Comparado de Sistemas Electorales (Com-

[23] Otras páginas web que pueden ser de utilidad para el lector: www.icpsr.umich.edu; www.gesis.org/ZA/; http://arces.cis.es/.

parative Study of Electoral Systems). Se trata de un programa de investigación desarrollado por equipos que estudian las elecciones y sus resultados en todo el mundo. Los países que participan en el estudio incluyen un módulo común de preguntas en sus encuestas post-electorales, y los datos que se producen se depositan en un archivo común. Éste incluye no sólo los datos individuales de encuestas sino también datos objetivos que contienen información sobre las campañas electorales, los sistemas electorales, e información demográfica, regional y nacional de cada país. Todos estos datos se pueden consultar y descargar de forma gratuita en la página web del proyecto: www.cses.org

5.10.2. Estudios longitudinales

La experiencia de estudios longitudinales en España ha sido más bien escasa hasta el momento, debido al elevado coste económico y organizativo de los estudios de este tipo. Entre los estudios del CIS existen algunas series de encuestas longitudinales, centradas únicamente en el ámbito electoral y con un seguimiento limitado. Concretamente, el CIS ha realizado estudios longitudinales de dos oleadas para las siguientes elecciones celebradas en España: las elecciones generales de 1977, las autonómicas catalanas de 1984, las generales y andaluzas de 1986, las europeas de 1999 y las generales y andaluzas del año 2000. También ha realizado un estudio longitudinal de cuatro oleadas para las elecciones europeas de 1994.

En el ámbito autonómico, la Fundación Jaume Bofill, una entidad privada, está realizando un estudio longitudinal sobre las desigualdades sociales en Cataluña que toma de modelo la conocida encuesta longitudinal británica *British Household Panel Survey*. Aunque el tema principal del estudio es la desigualdad, también recoge información sobre otros aspectos de interés social. La primera oleada del panel se realizó en 2001-2002 y el proyecto prevé al menos la realización de otros tres estudios anuales.

El *European Community Panel* es otro ejemplo de estudio longitudinal que empezó en 1994. Esta encuesta la realiza Eurostat (el instituto de estadística de la Unión Europea) en doce países de la UE, entre ellos España. Las entrevistas personales se realizan cada año a cerca de 6000 familias. Los temas que trata son variados, por ejemplo, la salud, las condiciones del mercado laboral, la composición del hogar, la educación y el nivel de bienestar. Los datos de España de este estudio se encuentran disponibles en la base de datos del INE.

En Estados Unidos, las encuestas nacionales longitudinales *(National Longitudinal Surveys),* que representan cinco grupos numerosos de hombres, mujeres y niños, se empezaron a realizar en 1966 con la recopilación de datos de un grupo de 5020 hombres de edades comprendidas entre 45 y 59 años. En el mismo año se seleccionó una muestra de hombres de edades entre 14 y 24 años, en 1967, un grupo de mujeres entre 30 y 44 años, y en 1968 otra de mujeres entre 14 y 24 años, todas de tamaño similar. Estos grupos originales se fueron estudiando cada dos años, finalizando en 1981 para el grupo de jóvenes y en 1990 para el grupo de hombres de más edad. En los grupos de mujeres se ha continuado realizando el seguimiento. En 1979 se seleccionó un nuevo grupo (el NLSY79) de 12686 hom-

bres y mujeres jóvenes, de edades comprendidas entre 14 y 22 años, del cual se viene haciendo un seguimiento anual.

Los datos de estas encuestas reflejan la fecha del principio y el final de la trayectoria laboral de cada individuo, con información sobre cada uno de los trabajos desempeñados. Otros datos, como posibles cambios de estado civil, fertilidad, acceso a ayudas de la asistencia pública o a la prestación por desempleo, se recopilan también con el mismo formato de historial. Las encuestas dirigidas a mujeres recogen información sobre el cuidado de los hijos, el reparto de las tareas domésticas y actitudes hacia el trabajo de las mujeres. Las encuestas a los grupos que se están aproximando a la edad de jubilación incluyen preguntas sobre planes de jubilación y pensiones.

Otra serie de estudios longitudinales realizados en Estados Unidos es el *Panel Study of Income Dynamics*, que se inició en 1968 entrevistando a una muestra de 4800 unidades familiares. Esta muestra, que se caracteriza por una sobrerrepresentación de las familias con bajos ingresos, se sigue a lo largo de toda su vida hasta que exista un cambio en la composición del hogar (por ejemplo, cuando uno de los hijos se casa y forma otro hogar). Hasta 1997, el seguimiento de la muestra era anual, y desde entonces se realizan entrevistas cada dos años. Esta encuesta fue creada para estudiar los factores determinantes de los ingresos familiares, e incluye preguntas sobre actitudes y comportamientos, trayectorias laborales e ingresos, ayuda de familiares o amigos, y características sociodemográficas.

En Estados Unidos también existen series de estudios con datos longitudinales sobre enseñanza, como el *National Longitudinal Study of the Class of 1972* (NLSC72), el *High School and Beyond* (HS&B) o el *National Education Longitudinal Study* (NELS).

5.11. NUEVAS TENDENCIAS Y NUEVOS RIESGOS

El motivo por el que hemos dedicado tanto espacio a la encuesta por muestreo es que ésta ha sido siempre y continúa siendo hoy día la técnica de investigación más difundida en sociología y probablemente en la investigación social en general. Sus procedimientos se han puesto a punto y se han probado mediante miles de investigaciones en Estados Unidos y en otros países occidentales en los años cincuenta y sesenta, hasta consolidarse un modelo estándar que sigue siendo válido en la actualidad [24].

Sin embargo, en los últimos años están surgiendo algunas tendencias que despiertan una cierta preocupación. Se trata de que, en el campo de la investigación mediante encuesta por muestreo, estamos asistiendo, tanto a escala nacional como internacional, a un creciente y grave distanciamiento entre el momento de la elaboración teórica, el de análisis de los datos, y el de recopilación de los mismos. Las causas pueden ser varias: la disponibilidad de archivos de datos ya recogidos y listos

[24] Muchos de los manuales escritos en Estados Unidos a finales de los años sesenta y principios de la década siguiente siguen siendo útiles y se siguen reeditando.

para su elaboración sobre fenómenos sociales muy diversos, o la existencia de empresas especializadas en la realización de encuestas, que ofrecen servicios en poco tiempo y a precios reducidos liberando al investigador de la tarea de organizar la recogida de datos de campo, o la difusión de la investigación social que impone muestras cada vez más numerosas, que sólo puede ser gestionadas por organismos especializados, lo que margina a las investigaciones de menor envergadura.

La encuesta por muestreo clásica, tal y como la hemos descrito aquí, pasaba por fases controladas y siempre *directamente gestionadas* por el investigador (o el equipo de investigación). El investigador definía las hipótesis, y a partir de ellas construía una primera versión del cuestionario; él mismo se encargaba personalmente de supervisar el estudio exploratorio. Para ello, elegía a los entrevistadores, los formaba, mantenía con ellos una relación directa durante la recopilación de datos, organizaba y controlaba la fase de codificación, y por último procedía al análisis de dichos datos. Es cierto que había una separación entre la función del investigador y la función del entrevistador (debido a la imposibilidad física del investigador de realizar centenares de entrevistas), pero el proceso de recogida de la información estaba dirigido y controlado paso a paso por el investigador.

Actualmente, los modelos más comunes son otros. El que más se aleja del modelo clásico es aquel en el que el investigador se limita a analizar datos ya recopilados por otros, lo que hemos denominado análisis secundario. Otro modelo muy frecuente es aquel en el que el investigador prepara el cuestionario, pero encarga la recopilación de los datos a una empresa de sondeos.

En el manual de los años setenta del *Institute for Social Research* de la Universidad de Michigan, que era partidario de un enfoque conductual, se recomendaba que los entrevistadores estuvieran al corriente de los objetivos de cada pregunta y que el director de la investigación aclarara con ellos cualquier posible fuente de dudas [Isr 1976, 4]. En otras palabras, se consideraba importante que, aunque hubiera una clara separación entre las figuras del investigador y el entrevistador, se estableciera un contacto sólido entre ambos y que el investigador pudiera comunicar al entrevistador las motivaciones específicas de cada pregunta. Hoy día esto es menos frecuente, o incluso inexistente cuando el investigador trabaja sobre datos ya recogidos y codificados, o cuando subcontrata la recogida de datos a una empresa de sondeos. En ese caso no tiene contacto directo con los entrevistadores, sino con un empleado de la empresa de sondeos. Y papel del investigador se reduce al de un mero instrumento que registra, una prolongación del cuestionario.

Los entrevistadores profesionales hoy hacen una entrevista sobre los jóvenes y la droga, ayer realizaron una sobre el consumo de cosméticos y mañana llevarán a cabo otra sobre la popularidad de un presentador de televisión. Todas estas entrevistas se realizan con el mismo estilo profesional e impersonal, y el sociólogo que dirige el estudio no tiene la posibilidad de reunir a los entrevistadores, explicarles cuáles son los objetivos de la investigación y discutir con ellos las reacciones de los entrevistados, es decir, formar un grupo y aprovechar su experiencia para entrar en contacto, aunque no directo, con los sujetos estudiados.

De este modo se crea una distinción entre las partes nobles de la investigación (desarrollo teórico, análisis estadístico, interpretación de los resultados) y las partes

vulgares (la producción o recopilación de información), que se subcontratan, y que Gilli [1971, 128] llamaba los «peones de la investigación». El investigador no sólo delega en otros estas tareas, sino que renuncia también a su control. Esta automaticidad en la fase de recopilación de los datos y esta falta de control del investigador provocan una disminución de la calidad de la producción de la evidencia empírica.

Esta reducción de la calidad resulta desconcertante, sobre todo si pensamos que en el modelo actual de investigación social se dedica una gran atención al análisis de los datos. Por un lado nos preocupamos cada vez más por crear modelos de análisis estadístico multivariados con un alto grado de sofisticación, y por otro nos preocupamos cada vez menos por saber qué registran efectivamente esas variables manipuladas con tanta perfección. Debemos denunciar la escasa sensibilidad hacia la «fidelidad» de los datos [Maradi; 1990] que conlleva un grave riesgo: que realicemos operaciones muy sofisticadas con datos inconsistentes, y como consecuencia acabemos generando investigaciones desechables.

5.12. SÍNTESIS DEL CAPÍTULO 5

1. La encuesta por muestreo se basa en la recogida de información preguntando a los individuos que son objeto de la investigación, que forman una muestra representativa, mediante un procedimiento de cuestionario estandarizado, con el fin de estudiar las relaciones entre las variables. El instrumento de investigación está formado por dos partes fundamentales, las preguntas y las respuestas, que pueden formularse de forma estandarizada o de manera libre. Cuando ambas partes son estandarizadas, estamos ante el llamado cuestionario, que es el instrumento más difundido para la producción de datos en investigación social.
2. El investigador que decide explorar la realidad social preguntando directamente a los actores sociales se encuentra con dos dilemas. El primero está relacionado con la oposición entre la idea de que la realidad social se puede conocer de forma objetiva (posición *objetivista*), y la concepción de que el conocimiento sólo puede nacer de la interacción entre el sujeto que estudia y el sujeto estudiado (posición *constructivista*). El segundo dilema surge de la contraposición entre quienes piensan que en el mundo social existen regularidades que se registran en la actuación individual (posición *uniformista*) y quienes afirman que todo acto social es un hecho único (posición *individualista*). El investigador que decide utilizar la técnica de la encuesta por muestreo para producir los datos sociales ha resuelto los dos dilemas eligiendo la primera alternativa de cada uno de ellos.
3. Otra limitación de esta técnica tiene que ver con la fiabilidad del comportamiento verbal, y plantea dos problemas distintos. El primero se refiere a las *conductas socialmente deseables:* si en una cultura una actitud o un comportamiento determinados tienen una connotación negativa, el entrevistado puede estar tentado de atribuirse la actitud o el comportamiento socialmente adecuados, en lugar de los reales. La segunda cuestión es la de la *falta de opinión*

(o seudoopiniones): en el curso de una entrevista puede suceder que al preguntar a los sujetos por temas complejos sobre los que no tienen una opinión, éstos contesten al azar, porque no se atreven a confesar que no tienen una opinión al respecto.

4. En función del *contenido*, las preguntas pueden clasificarse en tres tipos, según se refieran a *propiedades sociodemográficas* básicas, *actitudes* o *comportamientos*. En función de la *forma*, las preguntas pueden ser *abiertas* y *cerradas*.
5. El modo en que se formula una pregunta puede influir enormemente en la respuesta. Para evitar errores en la formulación, proponemos una serie de sugerencias sobre la sintaxis, el lenguaje, el contenido, la focalización en el tiempo, los efectos de la memoria, el orden de las preguntas, el número de opciones de respuesta, y los modos de afrontar los problemas de las conductas socialmente deseables y la ausencia de opinión.
6. En algunas ocasiones se usa la llamada batería de preguntas, que consiste en una serie de preguntas de contenido bastante similar que se formulan del mismo modo y se presentan al entrevistado en bloque. De esta forma se ahorra tiempo y espacio (en el cuestionario), la comprensión del mecanismo de respuesta resulta más fácil para el entrevistado, se mejora la validez de las respuestas y se facilita la construcción de índices de síntesis de las preguntas de la batería.
7. Existen tres formas fundamentales de realización de las encuestas: la entrevista cara a cara, la entrevista telefónica y el cuestionario autocumplimentado. En la entrevista cara a cara, el entrevistador desempeña un papel muy importante, por lo que es fundamental minimizar los efectos de su presencia. La entrevista telefónica permite reducir el tiempo y los costes de la investigación y hace más sencilla la formación y el control de los entrevistadores. Sin embargo el entrevistado se implica menos, no es posible utilizar elementos de apoyo visual, las personas que no tienen teléfono no pueden incluirse en la muestra, y resulta muy difícil acceder a ciertas categorías sociales. El cuestionario autocumplimentado puede realizarse de manera individual o en grupo; en el primer caso la situación más frecuente es la del cuestionario postal.
8. La recopilación de información suele ir precedida de las siguientes operaciones: entrevistas exploratorias preliminares, pretest de prueba del cuestionario, preparación (y supervisión una vez iniciado el trabajo de campo) de los entrevistadores y contacto con las personas a entrevistar.
9. El análisis secundario consiste en un nuevo análisis de los datos de una encuesta por muestreo realizada con anterioridad. En la actualidad este tipo de análisis se ha visto muy favorecido por el desarrollo de la informática, que facilita enormemente el intercambio de bases de datos y archivos.

Podemos introducir la variable «tiempo» en la encuesta por muestreo mediante la realización de encuestas longitudinales *(panel studies)* o la repetición de encuestas transversales *(repeated cross-sectional surveys)*. En el primer caso se trata de repetir la encuesta sobre los mismos sujetos en momentos distintos, en el segundo se repite la encuesta, pero sobre una muestra repre-

sentativa (y equivalente) de sujetos distintos.

10. En los países donde la investigación social está más desarrollada se han consolidado en los últimos años una serie de programas de investigación con un alcance temporal amplio y cuyo objetivo es el estudio del cambio social, en diversos campos, como las transformaciones demográficas, la salud, la enseñanza, la participación política, la movilidad social, los hábitos de consumo o la opinión pública.

5.13. LECTURAS COMPLEMENTARIAS

En la obra de C. Marsh *The Survey Method: The Contribution of Surveys to Sociological Explanation* (Allen y Unwin, 1982, 180 pp.), puede encontrarse una reconstrucción de la historia de la encuesta por muestreo y una contribución al debate metodológico sobre esta técnica.

Para una disertación sobre el problema de la estandarización y sobre la comparación entre el cuestionario predefinido e invariable y las entrevistas flexibles, véase la obra de H. Houtkoop-Steenstra *Interaction and the Standardized Survey Interview: The Living Questionnaire* (Cambridge U.P. 2000, 209 pp.).

Sobre el diseño del cuestionario sugerimos tres libros: el de J. M. Converse y S. Presser *Survey Questions: Handcrafting the Standardized Questionnaire* (Sage, 1986, 80 pp.), un resumen escueto pero completo de los conocimientos relevantes; el de F. J. Fowler Jr. *Improving Survey Questions* (Sage, 1995, 191 pp.), que contiene una gran cantidad de sugerencias prácticas; y el de S. Sudman, N. M. Bradburn y N. Schwarz *Thinking About Answer: The Application of Congnitive Processes to Survey Methodology* (Jossey-Bass, 1996, 304 pp.), un exhaustivo tratado sobre los efectos del diseño del cuestionario en los resultados de la investigación.

Para una breve reflexión sobre la entrevista telefónica, sugerimos el artículo de R. M. Groves «Theories and Methods of Telephone Surveys», en *Annual Review of Sociology* 1990, y para una descripción más completa y práctica, el libro de P. J. Lavrakas *Telephone Survey methods: Sampling, Selection and Supervision* (Sage, 1993, 181 pp.). Para los cuestionarios postales indicamos un ensayo breve, el de D. A. Dillman «The Design and Administration of Mail Surveys», en *Annual Review of Sociology 1991,* y uno práctico, el de T. W. Mangione *Mail surveys: Improving the Quality* (Sage, 1995, 129 pp.).

En lo que respecta a los estudios longitudinales, la obra de S. Menard, *Longitudinal Research* (Sage, 1991, 81 pp.) ofrece una presentación sintética. El libro de D. Kasprzyk, G. Duncan, G. Kalton y M. P. Singh (editores), *Panel Surveys* (Wiley, 1989, 592 pp.), contiene una selección de ensayos que tratan sobre todos los aspectos del diseño y el análisis de las encuestas por muestreo longitudinales. Para una reflexión sobre el diseño y las distintas estrategias de análisis de las encuestas transversales repetidas en el tiempo, remitimos al libro de G. Firebaugh *Analyzing Repeated Surveys* (Sage, 1997, 71 pp.). Por último, sobre el análisis secundario sugerimos el título de K. J. Kiecolt y L. E. Nathan *Secondary Analysis of Survey Data* (Sage, 1985, 87 pp.).

En cuanto a la bibliografía disponible en castellano, destacamos los tres libros siguientes aparecidos recientemente en la colección Monografías de la editorial del CIS. Por un lado, tenemos el libro de F. Alvira Martín, *La encuesta: una perspectiva general metodológica,* y la obra de V. Díaz de Rada, *Manual de trabajo de campo en la encuesta*, donde se reflexiona sobre la utilidad de la encuesta como instrumento de producción de datos y los pormenores de la etapa de realización de entrevistas, respectivamente. Por su parte, F. Bernardi, en su obra *Análisis de la historia de acontecimientos*, reflexiona sobre la naturaleza de los datos longitudinales y ofrece una guía útil para analizarlos.

CAPÍTULO 6

LA TÉCNICA DE LAS ESCALAS

6.1. LA OPERACIONALIZACIÓN EMPÍRICA DE CONCEPTOS COMPLEJOS 210

6.2. PREGUNTAS DE RESPUESTA GRADUADA: LA AUTONOMÍA SEMÁNTICA DE LAS RESPUESTAS 213

6.3. LA ESCALA DE LIKERT 219

6.4. EL ESCALOGRAMA DE GUTTMAN 227

6.5. EL DIFERENCIAL SEMÁNTICO 233

6.6. EL TEST SOCIOMÉTRICO 238

6.7. ESCALAS UNIDIMENSIONALES Y ESCALAS MULTIDIMENSIONALES 243

6.8. SÍNTESIS DEL CAPÍTULO 6 245

6.9. LECTURAS COMPLEMENTARIAS 246

En este capítulo abordamos uno de los problemas más difíciles de la investigación social. El problema de la «medición», y en particular la manera de transformar en operaciones empíricas los conceptos complejos de las ciencias sociales, como los estados emocionales, los rasgos psicológicos, las actitudes políticas o las escalas de valores. Tras un examen teórico inicial y un breve paréntesis histórico, presentamos algunas de las técnicas de escalas más sencillas y conocidas; por último, en la parte final se presenta una técnica para estudiar las interacciones entre individuos en grupos pequeños.

6.1. LA OPERACIONALIZACIÓN EMPÍRICA DE CONCEPTOS COMPLEJOS

La expresión «técnica de las escalas» (en inglés *scaling*), hace referencia a un conjunto de procedimientos creados por la investigación social para «medir» determinadas variables en el hombre y en la sociedad. Ya hemos hablado sobre la complejidad de los conceptos utilizados en las ciencias sociales y sobre la posibilidad o imposibilidad de observar directamente muchos de ellos. Optimismo, depresión, prejuicios raciales, autoritarismo, religiosidad, inteligencia, integración social, conflictividad, etc., son conceptos típicos de las ciencias sociales, que caracterizan al ser humano, pero que no son fácilmente traducibles al lenguaje de la investigación empírica. Podemos pensar que muchos de los comportamientos y las opiniones de un individuo dependen de su conservadurismo político o su religiosidad, sin embargo, estas propiedades no se pueden observar directamente, lo que podemos observar son sus comportamientos o sus opiniones, que se derivan de la propiedad subyacente y que están influidos por ésta, pero que no coinciden con ella.

Ya tratamos esta cuestión en el capítulo 3, cuando hablamos de conceptos e indicadores. Entonces comentamos que un concepto general como la religiosidad puede operacionalizarse mediante un concepto específico que se denomina indicador, y que está ligado a él por una relación de significado (de este modo podemos operacionalizar la religiosidad con la práctica religiosa, y el conservadurismo político con el partido al que se ha votado).

La técnica de las escalas sirve para conseguir este objetivo de forma más sistemática y formalizada. La diferencia respecto a lo que ya hemos dicho sobre el uso de los indicadores radica en el hecho de que ahora no se trata sólo de sustituir un concepto con uno o más indicadores, sino con un conjunto coherente y orgánico de indicadores, elaborando también criterios intersubjetivos para controlar la coherencia efectiva entre indicadores y concepto y comprobar que el procedimiento se ha completado. Podemos decir, por tanto, que una *escala es un conjunto coherente de elementos que se consideran indicadores de un concepto más general.*

El *elemento* es el componente individual (afirmación, pregunta, comportamiento, respuesta a un test, cualidad), y la *escala* es el conjunto de los elementos. El concepto subyacente al elemento tiene distintas denominaciones, según la disciplina: los psicólogos hablan de «rasgo» o «constructo», los sociólogos emplean fre-

cuentemente el término «variable latente».

Una prueba de habilidad matemática es un ejemplo de escala: se establece la hipótesis de que existe un concepto general, como la «habilidad matemática», que se puede registrar mediante una serie de pruebas específicas (resolución de ecuaciones, problemas, pruebas lógicas, etc.), cuyos resultados se sintetizan en una puntuación. La tabla 5.1 (pregunta 1) del capítulo anterior es otro ejemplo de escala, que contiene una serie de trastornos psicofísicos extraídos del cuadro de depresión de Carroll: en este caso los síntomas son los elementos de la escala, la depresión es el concepto general, y el resultado final de la escala es una puntuación, que se basa en las respuestas individuales y se asigna a cada uno de los individuos que realizan la encuesta. Esta puntuación se considera una medida del grado de depresión.

La aplicación más frecuente de la técnica de las escalas en el campo de la sociología y la psicología social es la denominada *medida de las actitudes,* en la que la unidad de análisis es el individuo, el concepto general es una actitud y los conceptos específicos son opiniones. Por «actitud» entendemos «el conjunto de tendencias y sentimientos, prejuicios e ideas preconcebidas, nociones, temores, inquietudes y convicciones de una persona hacia una determinada cuestión» [Thurstone y Chave, 1929, 67]. La actitud es, pues, una creencia que no se puede registrar directamente, mientras que la opinión es una de las formas en que se manifiesta la actitud, es decir, una expresión de ella que se puede registrar empíricamente.

Podríamos decir que la actitud es el concepto general y las opiniones son sus indicadores. El procedimiento para registrar las actitudes consiste en presentar a los sujetos estudiados una serie de afirmaciones y pedirles que expresen su opinión al respecto. Si combinamos adecuadamente las respuestas podemos llegar a una puntuación individual que estima la posición de cada sujeto sobre la actitud en cuestión.

En la tabla 6.4 presentamos una de estas escalas, en la que la actitud es el «sentimiento de eficacia política», un aspecto de la participación política que se registra mediante la opinión expresada en afirmaciones individuales sobre la política, los diputados, etc. Si puntuamos cada una de las respuestas según un baremo que va de 1 a 4, (1 para el mínimo de eficacia y 4 para el máximo) y sumamos los puntos, podemos atribuir a cada entrevistado una puntuación total de eficacia política, que va desde el valor mínimo, 9 (el total de preguntas del ejemplo es nueve), hasta el valor máximo, 36. La escala está formada por la suma de las puntuaciones obtenidas en cada una de las afirmaciones, que son sus «elementos».

Desde las primeras aplicaciones, en los años treinta, los sociólogos y los psicólogos sociales han elaborado centenares de escalas para «medir» las actitudes y, más en general, para registrar varias dimensiones de la personalidad humana: estados emocionales (ansiedad, depresión, resentimiento), rasgos psicológicos (autoestima, introversión), necesidades (autorrealización, poder), relaciones sociales (estatus social, integración familiar), tendencias políticas (izquierda-derecha, alienación política)[1].

[1] Algunas de las escalas de actitud más utilizadas se pueden encontrar en Miller [1991], que contiene las escalas utilizadas en la «American Sociological Review» *(Inventory of Mesures Utilized in the American Sociological Review, 1981-87),* y en las dos obras de Robinson, Shaver y Wrightsman [1991] sobre actitudes psicológicas y [1999] sobre actitudes políticas.

En el ámbito de las ciencias de la educación se han elaborado escalas para medir habilidades y capacidades (manuales, mentales), o la facilidad o dificultad para el aprendizaje (test de evaluación escolar, profesional).

Aunque la «medida de las actitudes» constituye el campo de aplicación más importante de la técnica de las escalas, no es el único. En primer lugar, la técnica se puede utilizar no sólo para registrar las propiedades de los individuos a través de sus respuestas a una serie de estímulos (los elementos de la escala), sino también para atribuir una puntuación a los estímulos en función de las respuestas de los individuos. Por ejemplo, podemos atribuir a las profesiones una puntuación de prestigio social, o bien establecer una jerarquía de popularidad entre los políticos, a partir de las respuestas y las opiniones expresadas por los entrevistados.

Además, la técnica de las escalas puede emplearse para registrar las propiedades de otras unidades que no sean individuos, por ejemplo, para juzgar la eficiencia de las instituciones (gobiernos, empresas, organismos públicos, etc.), para asignar puntuaciones de cohesión social a las comunidades, para determinar el poder de una serie de cargos profesionales a partir de las relaciones de autoridad-sumisión que mantienen en la interacción cotidiana en el trabajo, etc.

¿Qué podemos decir sobre el tipo de variables que genera la técnica de las escalas? ¿Son trata de variables nominales, ordinales o cardinales? La dimensión subyacente se suele concebir como una propiedad continua: el grado de prejuicio racial de una persona, o la religiosidad, varían de forma gradual en los individuos, o al menos así se cree. Que estas propiedades continuas no se traduzcan en variables cardinales se debe sólo a nuestra incapacidad para medirlas, es decir, para establecer una unidad de medida de las mismas. La técnica de las escalas representa la forma en que las ciencias sociales han hecho frente a este problema, proponiendo procedimientos que conduzcan a puntuaciones cardinales para evaluar las propiedades en cuestión.

¿Se ha alcanzado o se puede alcanzar este objetivo ? La técnica de las escalas tradicionales ha producido variables que hemos denominado «cuasicardinales» (ver el capítulo 3, apartado 6), y ese «cuasi» indica precisamente la imposibilidad de cumplir el objetivo de atribuir un significado numérico pleno a las puntuaciones de las escalas [2]. Sin embargo, los avances teóricos, junto con los nuevos recursos de cálculo derivados de la informática, han permitido recientemente elaborar modelos estadísticos capaces de producir variables para las que se realiza esa «igualdad de intervalos» que equivale a la existencia de una unidad de medida. Se trata de aplicaciones complejas y de difusión limitada (las escalas de Rasch), por lo que podemos decir que el cuerpo central de las técnicas de investigación social no va más allá de las escalas cuasicardinales.

Recordemos brevemente que los primeros intentos de registrar las actitudes mediante el uso de escalas se remontan a mediados de los años veinte, con las

[2] Ni siquiera la escala más estudiada y perfeccionada, el cociente intelectual, tiene una verdadera unidad de medida, y sigue siendo dudoso que los intervalos numéricos se correspondan realmente con la propiedad latente (no sabemos si la diferencia entre las puntuaciones de 95 y 100 es efectivamente igual a la diferencia entre las puntuaciones de 100 y 105).

propuestas de Allport y Hartman, Bogardus y Thurstone. Fue Thurstone [1927; 1928; 1931] quien hizo una primera sistematización en este campo, al formular tres propuestas diferentes *(Paired Comparison, Rank Order, Equal Appearing Intervals),* de notable interés metodológico (sobre todo la tercera), pero que hoy se consideran superadas, sobre todo porque su aplicación es muy laboriosa. Por lo contrario, es decir, por la gran sencillez de realización, tuvo un gran éxito (que aún perdura) la propuesta expuesta por Likert en 1923. Más adelante, tendría una gran relevancia la aportación de Guttman [1944; 1950]. Los manuales de metodología y la práctica de la investigación social se han centrado durante años en los tres modelos de las escalas de Thurstone, Likert y Guttman. Recientemente, gracias a las nuevas posibilidades de cálculo que ofrece la informática, ha surgido una nueva aproximación de tipo probabilístico para el registro de propiedades continuas, que ha encontrado vías de aplicación adecuadas; esta aproximación fue presentada inicialmente por Lazarsfeld [1950], y vio sus primeras aplicaciones en las propuestas de Rasch [1960] y Mokken [1971]. No entraremos en los detalles de la misma, ya que su complejidad supera los límites de este texto.

6.2. PREGUNTAS DE RESPUESTA GRADUADA: LA AUTONOMÍA SEMÁNTICA DE LAS RESPUESTAS

Una escala está formada por varios elementos, que en la mayoría de los casos constan de varias preguntas; usando una noción que presentamos en el capítulo 3, podemos decir que, por lo general, una *escala está constituida por una batería de preguntas.*

Antes de estudiar en detalle la cuestión de la construcción de las escalas, prestaremos atención al formato de las preguntas. En una pregunta cerrada, cuando las opciones de respuesta que se presentan al entrevistado están ordenadas (de manera creciente o decreciente), se pueden presentar de tres formas.

La primera consiste en presentar unas posibilidades de respuesta que son semánticamente autónomas, es decir, cada una tiene un significado intrínseco pleno que, para ser comprendido, no necesita relacionarse con el significado de las otras opciones presentes en la escala. En las preguntas y las correspondientes respuestas de la tabla 6.1 tenemos un ejemplo de esta forma: para responder que su nivel de estudios es universitario un entrevistado no necesita conocer las otras opciones de respuesta.

En la segunda forma las opciones de respuesta tienen una *autonomía semántica parcial.* El caso más común son las respuestas del tipo «mucho», «bastante», «poco», «nada», o similares (véanse los ejemplos de la tabla 6.2). En este caso el significado de cada opción de respuesta es sólo parcialmente autónomo con respecto a las otras opciones. No queda claro qué quiere decir exactamente que la política nos interese «bastante». En el contexto de la serie que va desde «mucho» hasta «nada», donde la opción «bastante» está situada después de «mucho» y antes de «poco», resulta más fácil atribuirle un significado.

Por último están las denominadas escalas de intervalos de autoposicionamiento [Cantrill y Free, 1962], en las que sólo hay dos categorías extremas dotadas de

Tabla 6.1. Preguntas con opciones de respuesta semánticamente autónomas

Pregunta 1. ¿Ha ido usted a la iglesia en el último año?	**Pregunta 2. ¿Cuál es su nivel de estudios?**
1. Nunca	1. Sin estudios
2. Dos-tres veces al año	2. Graduado escolar
3. Una vez al mes	3. Secundaria
4. Dos-tres veces al mes	4. Diplomatura
5. Una o más veces a la semana	5. Licenciatura

Pregunta 3. *Una pareja puede gestionar su dinero de diversas maneras. ¿Puede indicar, entre las siguientes, cuál se parece más a la que se aplica en su familia?*

1. El marido se queda con su sueldo y la esposa le pide dinero cuando lo necesita.
2. El marido se queda con su sueldo, pero le da a su esposa una cantidad fija de dinero para los gastos de la casa.
3. El marido ingresa su sueldo en una cuenta común de la que los dos cónyuges retiran dinero cada vez que lo necesitan.
4. El marido le da su sueldo a su mujer, que después le devuelve a su marido un porcentaje de ese dinero para sus gastos personales.
5. Otra (especificar).

significado, y entre ellas se coloca una escala, representada por casillas, cifras o un segmento, en el que la entrevistado marca su posición. En la tabla 6.3 reproducimos cuatro ejemplos de ello. El primero refleja la dimensión de orientación ideológica izquierda-derecha, el segundo emplea el llamado «termómetro de los sentimientos», y los dos últimos contienen preguntas sin especificidad propia, para que se comprenda que la técnica se puede aplicar a cualquier pregunta cuyas respuestas se puedan imaginar como un continuo delimitado por dos alternativas contrarias.

Tabla 6.2. Preguntas con respuesta de autonomía semántica parcial

¿Está usted de acuerdo con la siguiente afirmación: «Hoy en España los trabajadores hacen demasiadas huelgas, lo que perjudica al progreso del país»?	**¿Le interesa la política?**	**Indique el tipo de lugar de residencia**
1. Totalmente de acuerdo	1. Mucho	1. Urbano
2. Bastante de acuerdo	2. Bastante	2. Cuasiurbano
3. Más bien en desacuerdo	3. Poco	3. Semiurbano
4. Totalmente en desacuerdo	4. Nada	4. Semirrural
		5. Cuasirrural
		6. Rural

En cuanto al tipo de variable que se obtiene con estos tres procedimientos, podemos decir que en el primero de los casos obtendremos variables ordinales: sólo se asegura el orden de las modalidades, pero se desconoce totalmente la distancia entre ellas. La autonomía semántica de las categorías hace que el entrevistado las escoja por su contenido, al margen de su posición con respecto a las otras.

Tabla 6.3. Escalas de intervalos de autoposicionamiento

Pregunta 1. *En política suele hablarse de izquierda y de derecha. Sirviéndose de esta ficha, ¿dónde situaría su orientación política?*

Izquierda									Derecha	No sé
1	2	3	4	5	6	7	8	9	10	99

Pregunta 2. *Ahora queremos pedirle su opinión sobre algunas instituciones, hechos o partidos cuyo nombre está escrito en estas fichas (dar las fichas y dejar que las estudien con calma; después presentar la figura con el termómetro). Esta figura representa un termómetro, graduado de 0 a 100 grados. Puede usar este termómetro para darnos una idea de su simpatía o antipatía hacia las instituciones que aparecen en las fichas: si siente mucha simpatía o tiene usted una opinión muy positiva sobre ella, coloque la ficha a 100, o a 95-90 grados; si siente mucha antipatía o tiene una opinión muy negativa sobre ella, coloque la ficha a 0 o a 5-10 grados; si tiene usted una opinión intermedia, coloque la ficha entorno a los 50 o los 40-60 grados. Coloque con calma todas las fichas y desplácelas todas las veces que quiera.*

Pregunta 3. *Hay quien sostiene que los desempleados deberían aceptar cualquier trabajo disponible o perder la prestación por desempleo. En cambio otros piensan que los desempleados deberían tener derecho a rechazar un empleo que no desean. Ayudándose de esta ficha, ¿dónde situaría su posición al respecto?*

Aceptar cualquier empleo o perder la prestación								Tener derecho a rechazar un empleo que no desea		No sé
1	2	3	4	5	6	7	8	9	10	99

Pregunta 4. *¿Cuál es su grado de satisfacción con respecto a los siguientes aspectos de su vida? (Marque una cruz en la barra horizontal: si está totalmente insatisfecho, ponga la cruz cerca del extremo izquierdo; si está muy satisfecho póngala cerca del extremo derecho, de manera que su grado de satisfacción sea proporcional al segmento que queda a la izquierda del signo trazado).*

	Insatisfecho	Satisfecho	No sé
Situación económica			
Relación con su mujer			
Relación con sus hijos, etc.			

Tampoco en el caso de las respuestas con autonomía semántica parcial estamos en condiciones de afirmar que las distintas modalidades de respuesta son equidistantes entre sí: para un entrevistado, «bastante» podría tener una connotación muy positiva y estar más cercano a «mucho» y más alejado de «poco», mientras que para otro podría suceder lo contrario. Sin embargo, dado que las respuestas individuales no tienen significado autónomo, el entrevistado toma como referencia las otras opciones para interpretarlas, y es probable que en este proceso intervenga un mecanismo de comparación cuantitativa, especialmente cuando el número de opciones que se ofrecen es bastante elevado.

Este proceso de autoevaluación de la distancia entre las distintas opciones de respuesta es más probable que se produzca con las respuestas de intervalos de autoposicionamiento. Al atribuir un significado a las categorías intermedias, el entrevistado pone en marcha de manera automática un proceso mental de subdivisión graduada del espacio semántico entre los dos extremos, estableciendo correspondencias entre los estados sobre la propiedad y las posiciones en dicho espacio semántico. De este modo, si está muy satisfecho con una determinada situación o está totalmente de acuerdo con una determinada afirmación, se situará en un extremo, si está un poco menos de acuerdo se dirigirá hacia la zona central, si está equidistante entre los dos extremos se colocará exactamente en el centro, y pasará a la otra mitad de la escala si su opinión se orienta hacia el polo opuesto. Es probable que la subdivisión que el sujeto realiza mentalmente sobre la escala consista en una subdivisión de la misma en espacios iguales, por lo que esta técnica debe garantizar la equidistancia entre las categorías.

El hecho de que sea el mismo sujeto quien implícitamente establezca la unidad de medida de la escala, hace que este procedimiento sea subjetivo: en otras palabras, no disponemos de una unidad de medida externa, válida para todos los sujetos estudiados y para el investigador. Precisamente por ello no hablamos de verdaderas variables cardinales, sino de variables *cuasicardinales*.

Todos los investigadores están de acuerdo en que, a la hora de realizar el análisis estadístico, estas variables se traten como si fueran cardinales. En la práctica de la investigación social es bastante común tratar como variables cardinales también las variables del segundo y del primer tipo, y por lo general se asigna a las categorías valores numéricos en secuencia simple (1, 2, 3...). Sin embargo, esta práctica es controvertida, y ha suscitado un amplio debate entre los investigadores sociales.

Para concluir este punto, trataremos algunas cuestiones de carácter técnico relativas a la formulación de las preguntas. Los resúmenes 6.1 y 6.2 tratan sobre los temas de la respuesta neutra y el número de opciones de respuesta.

La expansión de la entrevista telefónica, que no permite utilizar soportes visuales (por ejemplo, fichas donde están escritas las siete opciones de respuesta que abarcan desde «totalmente de acuerdo» hasta «totalmente en desacuerdo»), cuya finalidad es facilitar la comprensión de las opciones de respuesta y la elección entre las distintas posibilidades, ha contribuido al uso de preguntas formuladas mediante una sucesión de elecciones binarias. Se le puede preguntar al entrevistado si está de acuerdo o en desacuerdo con una determinada afirmación; después, en función de lo que haya contestado, se le pregunta si «mucho» o sólo «bastante»; y así sucesiva-

Resumen 6.1. LA OPCIÓN CENTRAL NEUTRA

Cuando la variable es «de autonomía semántica parcial» se plantea el problema de si es conveniente ofrecer de manera explícita al entrevistado una opción central neutra (por ejemplo, «totalmente de acuerdo, bastante de acuerdo, *ni de acuerdo ni en desacuerdo,* bastante en desacuerdo, totalmente en desacuerdo»), o si es mejor no hacerlo, obligando el entrevistado a adoptar una postura clara. Los expertos no se ponen de acuerdo sobre cuál de las dos formulaciones es más adecuada: la presencia de una opción neutra puede favorecer la fidelidad del registro en los casos en que existan efectivamente posiciones intermedias, pero también puede ofrecer una escapatoria a quienes prefieren no exponer claramente lo que piensan [3]. En general, parece oportuno ofrecer una opción neutra, a menos que se quiera forzar al individuo a que elija una opción. Pero es muy recomendable, como ya se ha dicho en el capítulo 5, apartado 4, ofrecer siempre la posibilidad de la opción «no sé», para evitar lo que hemos llamado «seudoopiniones» [4].

mente [5]. Parece que esta formulación lleva a resultados similares a los que se obtienen con una única pregunta que contenga desde el principio todas las alternativas graduadas [Groves, 1989, 467].

En cuanto a las gradaciones de los intervalos de autoposicionamiento, en los ejemplos antes citados hemos presentado tres soluciones distintas para la secuencia de las respuestas: la de casillas vacías, la de la secuencia con cifras (a menudo de 1 a 7, o de 1 a 10; o bien, con el termómetro de los sentimientos, de 1 a 100), y la que presenta una línea continua. Entre la primera y la segunda técnica no hay grandes diferencias. La formulación de la línea continua ha sido propuesta por las ciencias sociales hace poco, adoptando así modalidades de encuesta que se han aplicado con éxito durante los últimos treinta años en el campo de las sensaciones físicas

[3] Para una reseña de las investigaciones que se han realizado sobre la cuestión de la opción neutra remitimos a la obra de Schuman y Presser [1981, capítulo 6]. Más en general, sobre la formulación de las preguntas para las escalas, véase De Vellis [1991].

[4] El tratamiento de las respuestas «no sé» en la fase de análisis de los datos es una cuestión bastante delicada. Existe una costumbre bastante común que consiste en atribuirles una puntuación intermedia entre los dos extremos, tratándolas como si fueran neutras. Pero esto es un error. Por ejemplo, ante la alternativa entre mayores impuestos y menores servicios sociales, quien elige una posición intermedia, interpretable como una posición de equidistancia entre los dos extremos, no es lo mismo que quien elige «no sé» porque no entiende la pregunta o porque nunca se ha planteado el problema.

[5] Este procedimiento se utiliza por ejemplo para la escala de posicionamiento político izquierda-derecha. Se le pregunta al entrevistado si es «de izquierdas, de centro o de derechas». Si dice que «de izquierdas (derechas)», se le pregunta: «¿Pero es usted exactamente de izquierdas (derechas), o bien de centro-izquierda (derecha)?». Si dice que «de centro», se le pregunta: «¿Pero es usted exactamente de centro-centro, o bien es un poco más de centro-izquierda o de centro-derecha?». De este modo, al final se obtienen siete posiciones.

Resumen 6.2. EL NÚMERO DE OPCIONES DE RESPUESTA

Otra cuestión controvertida es la del número de opciones de respuesta que se ofrecen al entrevistado: cinco, como en el ejemplo anterior, o siete (muy de acuerdo, más bien de acuerdo, en parte de acuerdo, etc.), o tres, etc. En general, podemos afirmar que si la escala contiene pocas preguntas (o directamente se basa en una sola pregunta), conviene ofrecer al entrevistado una amplia gama de posibilidades de respuesta (por ejemplo, cinco o siete opciones). Y si la escala contiene muchas preguntas, éstas podrían plantearse como una elección binaria (Sí/No). El objetivo de la escala es captar la diversidad entre los entrevistados con relación a la propiedad que se registra. Si una escala está compuesta por una sola pregunta y ofrece diez opciones posibles, podemos distribuir a los sujetos en diez posiciones. Si una escala está compuesta por diez preguntas, todas binarias, también obtendremos (sumando las respuestas de las diez preguntas) una puntuación final a diez posiciones.

Téngase en cuenta, por otra parte, que cuanto mayor sea el número de opciones, más difícil resultará la elección (es más fácil responder con un «de acuerdo/en desacuerdo» que establecer el propio grado de acuerdo). En cualquier caso, el problema ha encontrado soluciones específicas dentro de las diversas técnicas (por ejemplo, en las escalas de Likert las preguntas suelen presentar cinco o siete opciones, en las de Guttman las respuestas son dicotómicas, etc.). Por ello volveremos sobre el tema en la presentación de estas técnicas.

(luminosidad, intensidad de los sonidos, sabores, olores, estimaciones de peso, etc. [Lodge, 1981]).

Normalmente se le pide al entrevistado que trace un segmento de longitud proporcional al grado de aprobación de la afirmación propuesta, de simpatía hacia un determinado personaje, de confianza en una cierta institución, etc. La longitud se mide después en la fase de codificación y se transforma en un número. Este procedimiento de tipo gráfico se defiende argumentando que las convenciones relativas a las distancias físicas tienen un carácter de intersubjetividad mayor que las relativas a las distancias semánticas.

En este apartado hemos estudiado la estructura de las preguntas. Pero una escala suele estar compuesta por varias preguntas o, para ser más exactos, varios elementos. Nunnally [1978, 66-68] cita tres razones por las que las escalas de varios elementos son preferibles a las de un único elemento. En primer lugar, la complejidad de los conceptos que se pretende registrar hace difícil abarcarlos con un único indicador (recordemos lo dicho en el capítulo 3, apartado 7, sobre las distintas dimen-

siones de los conceptos complejos). En segundo lugar, un único registro no ofrece la precisión necesaria, puesto que no consigue discriminar entre las distintas posiciones de los sujetos con respecto a la propiedad considerada. Por ejemplo, en el caso muy frecuente de las preguntas dicotómicas (por ejemplo, a favor/en contra de una determinada afirmación), se consigue dividir a los sujetos en dos grupos, y no es posible hacer distinciones más sutiles. Por último, cada pregunta está más expuesta a los errores accidentales, que tienen menor peso cuando la escala se basa en un número mayor de respuestas.

En conclusión: las escalas compuestas por un único elemento son menos válidas, menos precisas y menos fiables. Por este motivo, en los apartados que siguen presentaremos las técnicas de escalas de varios elementos más difundidas.

6.3. LA ESCALA DE LIKERT

El nombre de esta técnica proviene del psicométrico Renis Likert, que la propuso por primera vez a principios de los años treinta [Likert, 1932]. La denominación de «escala de Likert» engloba una amplia variedad de escalas, también llamadas *summated rating scales* en inglés, término que se puede traducir como *escalas aditivas*. Esta técnica sigue siendo actualmente el procedimiento más utilizado para el estudio de las actitudes y se ha desarrollado y enriquecido notablemente en comparación con las propuestas iniciales de Likert. Al ilustrarlas introduciremos también algunos elementos generales (sobre la redacción de las preguntas, la validación de la escala, etc.), que se aplican a todas las técnicas de construcción de las escalas.

El procedimiento que subyace a las escalas de Likert es simple e intuitivo. Es la primera operación que se le ocurre a cualquiera para atribuir una puntuación total basándose en las puntuaciones de pruebas individuales: hacer la suma de los puntos de las pruebas individuales. Así, en un test escolar compuesto por 30 preguntas, podemos atribuir un punto a cada pregunta exacta y sumar los distintos puntos: el estudiante que no ha cometido errores obtendrá un 30. De manera análoga, en el campeonato de fútbol la puntuación de cada equipo en la clasificación general se obtiene sumando los puntos conseguidos en cada partido.

Tradicionalmente el formato de las preguntas individuales de las escalas de Likert está representado por una *serie de afirmaciones* para cada una de las cuales el entrevistado debe decir si está de acuerdo y en qué medida. En la versión inicial del mismo Likert se proponían siete alternativas: totalmente de acuerdo, de acuerdo, en parte de acuerdo, dudoso, en parte en desacuerdo, en desacuerdo, totalmente en desacuerdo. Después se han reducido a cinco, a veces incluso a cuatro, con la eliminación de la categoría intermedia. Remontándonos a la clasificación del apartado anterior, se trata de preguntas con respuestas de autonomía semántica parcial.

La construcción de la escala consta de cuatro fases: *a)* formulación de las preguntas, *b)* aplicación de las preguntas a una muestra de sujetos, *c)* selección de las preguntas y determinación del grado de coherencia de la escala, y *d)* control de la validez y la unidimensionalidad de la escala. A continuación pasamos a comentar

cada una de estas fases.

En la *primera fase* se determinan las dimensiones de la actitud estudiada y se formulan afirmaciones que reflejen los diversos aspectos del concepto general que se quiere destacar. Por ejemplo, Adorno y sus colegas identificaron uno de los componentes del autoritarismo en una «una actitud de impaciencia y oposición hacia los individuos de carácter cariñoso; los individuos autoritarios tienen miedo de los sentimientos verdaderos, pues las emociones podrían escapar a su control» [Adorno *et al.*, 1950, 235]. A partir de esta consideración incluyeron en su escala F de fascismo-autoritarismo algunas preguntas destinadas expresamente al reconocimiento de este componente particular de la personalidad.

Hay que subrayar la importancia de esta fase teórica. Aunque la teoría se puede aplicar a un conjunto cualquiera de afirmaciones centradas en un único tema, confiando luego a operaciones estadístico-matemáticas la tarea de seleccionar las afirmaciones pertinentes y descartar las poco congruentes, la escala resultante será más válida, es decir, más adecuada para registrar efectivamente la propiedad subyacente para la que ha sido diseñada, si las afirmaciones que contiene se formulan después de identificar en la teoría los distintos aspectos de los conceptos que se quiere estudiar.

Esos conceptos suelen ser complejos y contienen múltiples *dimensiones* que la escala debe cubrir (recordemos la relación entre conceptos e indicadores comentada en capítulo 3, apartado 7). Volviendo al que es uno de los ejemplos más conocidos de aplicación de esta escala, *La personalidad autoritaria,* de Adorno y varios autores más, éstos identificaron, a partir de investigaciones anteriores, de estudios psicológicos, de la literatura general sobre el antisemitismo y el fascismo, nueve dimensiones de la personalidad autoritaria, en torno a las cuales construyeron cada una de las preguntas-afirmaciones de la escala.

Para concluir la reflexión sobre este punto con una afirmación sintética (que no sólo se aplica a las escalas de Likert, sino también al resto de escalas), podemos decir que es muy recomendable adoptar un enfoque deductivo, más que inductivo.

A continuación, en la *segunda fase*, la escala se aplica a una muestra de entrevistados. Lo único que merece la pena subrayar de esta fase es el hecho de que en *general esta técnica exige un alto nivel de formación de los entrevistados.* Afirmaciones del tipo «sería preferible aumentar los impuestos antes que reducir las prestaciones sociales que ofrece el Estado», o bien «es preferible renunciar a la utilización de las centrales nucleares para la producción de energía eléctrica, dado que las ventajas económicas que ofrecen quedan anuladas por los riesgos que conllevan para la población», pueden ser difíciles de comprender para un número importante de entrevistados.

Es necesario recordar también lo que ya hemos dicho sobre las baterías de preguntas: que son particularmente sensibles a los errores provocados por las respuestas emitidas al azar (pseudo-opiniones) o de manera mecánica y sistemática para todas las preguntas de la batería. Por ello es importante que se incluyan las respuestas «no sé», y que no todas las afirmaciones tengan el mismo sentido, a fin de obligar a los entrevistados a reflexionar, puesto que para que sus respuestas sean coherentes deberán expresar acuerdo en unas ocasiones y desacuerdo en otras.

En la tabla 6.4 hemos reproducido una escala del «sentimiento de eficacia política». Este concepto fue introducido en 1954 por el equipo de investigación del Survey Research Center de la Universidad de Michigan para indicar la impresión de del ciudadano de que está integrado en el sistema político, en el sentido de que percibe que puede ejercer influencia en los procesos políticos. En la investigación de la que se ha extraído este ejemplo, la escala contenía nueve afirmaciones, cinco sobre política general y cuatro sobre política municipal.

Una vez que se ha aplicado la escala a los sujetos, corresponde valorar la capacidad efectiva de la escala para cumplir el objetivo para el que ha sido creada. Como sabemos, la premisa para ello es que todos los elementos que componen la escala estén correlacionados con un mismo concepto subyacente, ya que en la primera fase los elementos se han elegido partiendo de esta perspectiva. Lo que no sabemos es si esta elección, que el investigador ha realizado partiendo de una reflexión sobre el concepto a estudiar, ha sido correcta.

Por lo tanto, tras la aplicación de la escala a los sujetos, es necesario definir algún criterio empírico para asegurarnos de que esta dimensión común a todos los elementos sea efectivamente reconocible incluso para los entrevistados. De hecho, es posible que algunos de los elementos de la escala no sean coherentes con el resto, es decir, que estén semánticamente relacionados con otros conceptos distintos, y que, por tanto, se deban eliminar. Dicho de manera más rigurosa: debemos asegurarnos de que la escala es *unidimensional.*

Esta comprobación se realiza en la *tercera fase.* Para ello se utilizan dos instrumentos básicos: la correlación *elemento-escala* (o correlación «parte-todo»), que sirve para identificar los elementos de la escala que no son coherentes con los demás, y el *coeficiente alfa,* que sirve para juzgar el grado total de coherencia interna de la escala (ver también los resúmenes 6.3 y 6.4).

Al eliminar los elementos inadecuados de la escala, el investigador tendrá en cuenta tanto las correlaciones elemento-escala como el coeficiente alfa. Eliminará los elementos con correlación elemento-escala demasiado baja, y de este modo hará que aumente la correlación media entre los elementos. Es decir, se eliminarán los elementos que presenten una correlación elemento-escala menor, siempre y cuando esta operación produzca un aumento en el valor de alfa.

En la tabla 6.4 hemos reproducido el coeficiente alfa de la escala, los coeficientes de correlación elemento-escala y, para cada elemento, el valor de alfa que quedaría si ese elemento se eliminara. Vemos que la escala presenta un valor de alfa muy elevado ($\alpha = 0{,}88$), por lo que en principio podríamos decir que es satisfactorio. Los valores de las correlaciones elemento-escala son elevados para todos los elementos. El que presenta la correlación más baja es el número 2 ($r = 0{,}47$), pero su eliminación no produce un aumento de alfa (que sigue siendo igual a 0,88); por otro lado, era la única afirmación de carácter cognitivo de la escala, por lo que los investigadores decidieron mantenerla.

A continuación pasamos a la *cuarta fase,* en la que corresponde controlar la validez de la escala y su carácter unidimensional. Para controlar la validez debemos considerar la diferencia entre la aplicación clásica de la técnica de las escalas la aceptabilidad de la escala (siguiendo el procedimiento descrito en la fase ante-

Tabla 6.4. Escala de eficacia política

Pregunta: A continuación voy a leer algunas afirmaciones que se suelen hacer con frecuencia. Querría saber si está usted de acuerdo o no con cada una de estas afirmaciones. Le ruego que a medida que yo lea estas afirmaciones, me vaya diciendo si está totalmente de acuerdo, de acuerdo, en desacuerdo *o* totalmente en desacuerdo *con cada una de ellas. Diga exactamente lo que piensa; no hay respuestas correctas ni incorrectas. Si no sabe qué responder a alguna de estas preguntas o no tiene las ideas claras, diga «no sé».*

	% totalmente de acuerdo + de acuerdo	correlación elemento-escala [a]	valor
1. La gente como yo no tiene ninguna influencia en lo que hace el gobierno	62,9	0,58	0,87
2. A veces la política parece tan complicada que no es posible entender lo que está pasando	79,6	0,47	0,88
3. No creo que los políticos se interesen demasiado por lo que piensa la gente como yo	69,6	0,62	0,87
4. Normalmente las personas elegidas para el Parlamento pierden pronto el contacto con los electores	82,5	0,60	0,87
5. A los partidos sólo les interesan los votos de la gente, no sus opiniones	67,7	0,64	0,87
6. La gente como yo no tiene ninguna influencia sobre lo que hace el gobierno de los Ayuntamientos	50,1	0,70	0,86
7. Normalmente las personas elegidas para el gobierno municipal de mi ciudad pierden muy pronto el contacto con los electores	47,4	0,71	0,86
8. La gente como yo no tiene ninguna influencia sobre lo que hace un concejal de distrito	46,8	0,67	0,87
9. Normalmente las personas elegidas como concejales de distrito pierden pronto el contacto con los electores	41,5	0,68	0,87
		Alfa = 0,88	

[a] Correlación «correcta»: correlación entre las puntuaciones sobre ese elemento y las puntuaciones sobre el índice de los elementos restantes.

Resumen 6.3. LA CORRELACIÓN ELEMENTO-ESCALA

Para calcular la correlación elemento-escala, se calcula la puntuación de cada sujeto sobre toda la escala y el coeficiente de correlación entre esta puntuación y la puntuación sobre cada elemento individual [6]. El coeficiente de correlación es una medida que cuantifica el grado de relación existente entre dos variables cardinales. Si las dos variables covarían (es decir, al variar una, varía también la otra), el coeficiente de correlación toma un valor elevado (hasta el máximo de 1 si la correlación es positiva y de –1 si la correlación es negativa) [7]; si las variables no están correlacionadas, el coeficiente tendrá un valor bajo, que en caso de ausencia total de correlación será igual a cero. En nuestro caso, el coeficiente de correlación nos dice si la puntuación de cada elemento va en la misma dirección que la puntuación global que tiene en cuenta el resto de los elementos.

Por ejemplo, si estamos midiendo el autoritarismo, tendrá que ocurrir que si a una persona se le atribuye una puntuación global elevada sumando las respuestas a todas las preguntas (es decir, resulta «autoritaria»), esa persona tendrá que tener una puntuación bastante elevada (respuestas de tipo «autoritario») también sobre cada uno de los elementos de la escala. Es posible encontrar excepciones individuales, pero si hay una pregunta que de manera sistemática (es decir, para un número muy elevado de individuos) presenta valores contradictorios con la puntuación global de cada uno de los individuos, habrá que deducir que esa pregunta específica presenta algún problema. Puede ser que sea ambigua, que sea malinterpretada debido a una formulación inadecuada, o que haga hincapié en un aspecto distinto del autoritarismo. En todo caso, está claro que esa pregunta no es congruente con el resto de la escala y por tanto se deberá eliminar.

y una aplicación simplificada. En el planteamiento tradicional, la construcción de la escala es una operación con autonomía propia, es decir, cuando se construye una escala para medir variables como la autoestima, la ansiedad, la religiosidad, etc., se busca crear un instrumento que sea aplicable también a poblaciones distintas de aquella para la que se creó inicialmente la escala.

Se parte de un elevado número de elementos (que puede llegar hasta los cincuenta), que se aplican a una muestra limitada de sujetos (por lo general 100-200); a partir de esta primera prueba se eliminan los elementos incoherentes y se establece

[6] El procedimiento más correcto consiste en calcular para cada elemento la correlación entre su puntuación y la que resulta de sumar (o calcular la media) de todos los elementos *restantes* (correlación «correcta» elemento-escala).

[7] Correlación positiva: al aumentar una variable también aumenta la otra; correlación o negativa: al aumentar una variable disminuye la otra.

Resumen 6.4. LA COHERENCIA INTERNA DE LA ESCALA: EL ALFA DE CRONBACH

Se han elaborado algunos criterios para valorar *la coherencia interna global de una escala.* Se trata de índices que resumen esta coherencia, y entre los más conocidos se encuentra el alfa de *Cronbach,* que se basa en la matriz de relación entre todos los elementos y su número. Su fórmula es:

$$\alpha = \frac{nr}{1 + r\,(n - 1)}$$

donde n es el número de elementos de la escala, y r es su correlación media.

Aunque lo parezca, alfa no es un coeficiente de correlación; suele tomar un valor positivo [8], entre 0 y 1; cuanto más alto es su valor, mayor es la coherencia interna de la escala. Nunnally [1978, 245] sugiere un valor mínimo de 0,70 como criterio de aceptabilidad de la escala. Un alfa inferior a este valor significa que los elementos de la escala tienen poco en común, o que su número es demasiado bajo. Por ejemplo, si los elementos de la escala son 10 y la media de las 45 correlaciones entre los elementos es 0,30, el valor de α será 0,81 (suficiente). Si los elementos son cinco y la correlación media es la misma, el valor de α será igual a 0,68 (insuficiente). Alfa aumenta a medida que aumentan el número de elementos de la escala y su correlación media [9].

rior). La escala se aplica a diversas investigaciones, y en el curso de su utilización se van acumulando pruebas de validez sucesivas (validez predictiva, validez concomitante, validez por grupos conocidos, validez por frase), que permiten comprobar que efectivamente registra la propiedad para la cual ha sido creada [10].

Este modo de proceder es común en la investigación psicológica, cuando las propiedades que son objeto del estudio están representadas por rasgos de la personalidad, conceptos psicológicos muy delimitados y situados en lo profundo de la psi-

[8] Si el valor de alfa es negativo, quiere decir que hay elementos que están correlacionados negativamente, debido a polaridades no coherentes (por ejemplo, en un caso la puntuación máxima indica un autoritarismo alto y en otro caso indica un autoritarismo bajo). Si todas las puntuaciones van en la misma dirección, alfa tendrá sólo valores positivos.

[9] Esta particularidad matemática lleva a muchos psicométricos a construir y proveer escalas con demasiados elementos para aumentar la probabilidad de que alfa tenga un valor alto. Sin embargo esto aumenta también la probabilidad de que el entrevistado se canse y acabe respondiendo al azar.

[10] Recordemos la definición de «validez» que presentamos en el apartado 9 del capítulo 3 («la validez es la medida en que una determinada forma de traducción de un concepto en variable registra efectivamente el concepto que se pretende registrar»); ver dicho apartado para consultar las formas de control de la validez.

que, que requieren escalas bastante elaboradas, generalmente constituidas por varias decenas de elementos [11]. Los sociólogos y politólogos han adoptado un punto de vista más simplificado en el uso de la técnica de las escalas, que en su caso suelen tener una decena de elementos aproximadamente; además, en el mismo instrumento de registro (que suele ser un cuestionario) pueden coexistir varias escalas; la escala es sólo un aspecto de la investigación.

Un ejemplo de ello sería una investigación sobre actitudes hacia la política compuesta por una escala de «sentimiento de eficacia política», otra de «participación política» y otra de posicionamiento «izquierda-derecha», junto con otras muchas preguntas sobre socialización familiar, educación, profesión, religiosidad, etc. En este caso la escala se construye con una sola operación, sin separación entre «construcción» y «aplicación»; los controles de validez son insuficientes y no sistemáticos, y el ámbito de aplicación de la escala queda limitado a la población específica estudiada.

Pasamos ahora al problema del *control del carácter unidimensional* de la escala. Aunque el procedimiento de *análisis de los elementos* que hemos presentado en la tercera fase tenga como fin establecer que todos los elementos son indicadores de la misma propiedad, éste no es suficiente para garantizar el carácter unidimensional de la escala. Es posible, por ejemplo, que los elementos impliquen dos propiedades distintas, incluso muy distintas (por ejemplo, autoritarismo e ideología de clases), pero correlacionadas entre sí, lo que haría que los valores de las correlaciones elemento-escala y el coeficiente alfa fueran satisfactorios, aunque la escala no tuviera carácter unidimensional.

Una técnica muy eficaz para el control del carácter unidimensional de la escala es el *análisis factorial.* Este procedimiento es demasiado complejo para abordarlo en este manual, pero podemos entender su aplicación para el control de la unidimensionalidad de las escalas sin necesidad de conocerlo en profundidad (véase el resumen 6.5, en la página siguiente).

Para concluir este apartado, diremos que las escalas de Likert son la técnica más difundida en las ciencias sociales para el registro de las propiedades continuas, en particular las actitudes. Su popularidad se atribuye a la simplicidad de su estructura teórica y su facilidad de aplicación, que ni siquiera requiere, en principio, los recursos de cálculo de los ordenadores modernos, lo que explica su enorme éxito en los años cincuenta-setenta.

Esta técnica presenta, sin embargo, algunos inconvenientes. El primero guarda relación con la puntuación que se asigna a cada uno de los elementos. Cada elemento es una variable ordinal, por lo general «con autonomía semántica parcial», que suele presentar cinco opciones de respuesta (desde «totalmente de acuerdo» hasta «totalmente en desacuerdo», y similares), a las que se asignan de manera arbitraria, las puntuaciones en escala simple 1, 2, ..., 5, que se tratan luego como escalas

[11] Los psicométricos trabajan con escalas individuales muy largas (para elevar el valor de alfa) y en el laboratorio; los sociólogos y politólogos incluyen escalas (necesariamente más breves) en cuestionarios que aplican a estudios de campo.

cardinales (por ejemplo, en la correlación elemento-escala). El mismo Likert, consciente de este problema, propuso inicialmente una técnica más elaborada en la que las puntuaciones de las respuestas se asignaban a partir de los resultados de la encuesta («método sigma», Likert, 1932), pero la abandonó, ya que las puntuaciones más simples basadas en la secuencia de los números naturales eran muy parecidas a las calculadas mediante el criterio más sofisticado.

El segundo inconveniente guarda relación con la imposibilidad de *reproducir* la escala, es decir, el hecho de que la puntuación final de la escala es global y no hay manera de conocer a qué combinación de respuestas corresponde una puntuación determinada. Por este motivo puede suceder que dos puntuaciones idénticas corres-

Resumen 6.5. EL ANÁLISIS FACTORIAL

El objetivo del análisis factorial es reducir una serie de variables relacionadas entre sí a un número inferior de variables hipotéticas (factores o variables latentes) independientes entre sí [12]. El punto de partida para el análisis es la matriz de correlación entre las variables observadas, y el objetivo es explicar estas correlaciones mediante la existencia de factores subyacentes. Por ejemplo, la correlación entre las notas en álgebra y las notas en geometría de un determinado grupo de estudiantes (es decir, el hecho de que, por lo general, si un estudiante obtiene buenas notas en álgebra, también obtiene buenas notas en geometría, y si es mediocre en una de las dos materias, también lo es en la otra), se podría reinterpretar deduciendo que existe un factor subyacente que influye en ambas notas, que en este caso sería la capacidad de abstracción matemática del estudiante.

En la tabla 6.5 hemos reproducido los resultados del análisis factorial realizado sobre los elementos de la escala de la tabla 6.4. Las «saturaciones factoriales» indicadas en la tabla se pueden interpretar como coeficientes de correlación entre el elemento de la escala y el factor subyacente. El análisis, realizado según el procedimiento estándar, muestra la existencia de dos factores: el primero saturado por las cinco primeras preguntas y el segundo por las cuatro últimas. Por tanto, la escala es bidimensional, ya que las propiedades subyacentes son dos. La interpretación de los dos factores es bastante inmediata. El primero está relacionado con las preguntas relativas a la política nacional, el segundo con las relativas a los representantes políticos del Ayuntamiento y del distrito municipal. Por ello, los investigadores decidieron dividir las preguntas en dos escalas de «sentimiento de eficacia política»: uno referido a la política nacional y otro a la política local.

[12] Los factores también pueden estar relacionados entre sí («rotación oblicua»), pero aquí nos ceñiremos a la situación más sencilla.

Tabla 6.5. Análisis factorial sobre la serie de indicadores de eficacia política

Elemento nº	Saturaciones	
	Factor nº 1	Factor nº 2
1	**0,661**	0,295
2	**0,702**	0,104
3	**0,764**	0,245
4	**0,752**	0,224
5	**0,679**	0,360
6	0,346	**0,761**
7	0,300	**0,811**
8	0,210	**0,858**
9	0,204	**0,875**

pondan a respuestas muy distintas.

Por último, la puntuación final de la escala no representa una variable cardinal: no disponemos de ningún elemento para poder afirmar, por ejemplo, que la distancia entre las puntuaciones 14 y 18 sea la misma que entre las puntuaciones 8 y 12.

6.4. EL ESCALOGRAMA DE GUTTMAN

La propuesta de Guttman [1944; 1950] nace con el objetivo de ofrecer una solución para el problema del carácter unidimensional de la escala, que es uno de los puntos débiles de la técnica de Likert. La escala de Guttman consiste en una sucesión de elementos que presentan una dificultad creciente, de manera que si se ha respondido afirmativamente a una pregunta, se debe haber respondido también afirmativamente a todas las que la preceden en grado de dificultad. De ahí el nombre de escalograma o escala acumulativa (por contraste con el nombre de «escala aditiva», atribuido a la técnica de Likert, en la que todos los elementos de la escala tienen el mismo grado de dificultad).

Para ilustrar este carácter acumulativo de los elementos utilizaremos como ejemplo la «escala de distancia social» propuesta por Bogardus [1925] en los años veinte. En esta escala, para determinar los prejuicios de los entrevistados con respecto a ciertas minorías étnicas, se planteaba una secuencia de preguntas del tipo: ¿Estaría dispuesto a aceptar a un individuo de color como visitante de su país?, ¿estaría dispuesto a aceptar a un individuo de color como vecino de su casa?, ¿estaría dispuesto a aceptar a un individuo de color como amigo?, ¿estaría dispuesto a casarse con un individuo de color?

Es evidente que quien está dispuesto a casarse con un individuo de color, probablemente también estará dispuesto a tenerlo como amigo, como vecino, etc., mientras que quien no esté dispuesto a aceptarlo como amigo, probablemente tampoco esté dispuesto a casarse con él. Si los elementos de la escala están perfectamente

escalonados, sólo son posibles ciertas secuencias de respuestas. Si asignamos la puntuación 1 a la respuesta positiva y 0 a la negativa, en principio sólo deberían ser posibles las secuencias: 1111, 1110, 1100, 1000, 0000 (ver la tabla 6.6), mientras que, por ejemplo, la secuencia 1011 no debería producirse (sería una persona que está dispuesta a casarse con un individuo de color y a tenerlo como amigo, pero no como vecino). Como puede verse en la tabla 6.6, las respuestas posibles producen una matriz dividida en dos triángulos, uno de los cuales contiene sólo respuestas positivas y el otro sólo respuestas negativas.

Asignando la puntuación 1 a la respuesta afirmativa y 0 a la respuesta negativa y sumando las puntuaciones de cada individuo sobre todos los elementos, obtenemos la puntuación total sobre la escala para cada sujeto. Nótese que a partir de esta puntuación individual se pueden deducir las respuestas dadas por ese individuo a cada uno de los elementos de la escala: por ejemplo, quien haya obtenido una puntuación de 2 en la escala de tabla 6.6 habrá dado la secuencia de respuestas 1100. Por tanto podemos saber, no sólo a cuántas preguntas, sino también a qué preguntas concretas ha dado una respuesta afirmativa. Esta posibilidad de deducir las respuestas a cada una de las preguntas a partir de la puntuación de la escala es una característica típica del escalograma de Guttman, que también se designa con el término *reproducibilidad* (a partir de la puntuación se pueden «reproducir» las respuestas a cada pregunta). Nótese también que esta técnica sólo utiliza elementos dicotómicos, es decir, preguntas con respuestas del tipo «sí/no», «de acuerdo/en desacuerdo», «a favor de/en contra de» (por contraste con la técnica de Likert, en la que la respuesta se graduaba sobre una escala de intensidad).

Es obvio que sólo una escala ideal proporciona las secuencias de respuesta que aparecen *en la tabla 6.6. En cada aplicación real habrá también* respuestas que no estén contempladas en las secuencias previstas en el modelo, y que se considerarán «errores». Se trata de establecer cuántos errores puede tolerar una escala para ser aceptable y cuáles son los procedimientos a seguir para minimizar dichos errores.

Igual que en el caso de la escala de Likert, en la construcción de la escala de Guttman podemos distinguir tres o cuatro fases: la formulación de las preguntas, su aplicación a una muestra de sujetos, el análisis de los resultados con la eliminación de los elementos que indican demasiados errores, la definición de un índice global de aceptación de la escala, y una eventual fase de validación de la misma.

Para la fase de formulación de las preguntas (dicho de forma más general, la identificación de los elementos de la escala), recordamos lo ya dicho al hablar de la escala de Likert sobre la importancia de la reflexión conceptual. Para la escala de Guttman podemos añadir dos observaciones específicas al respecto. En primer lugar, el hecho de que las respuestas deben darse de forma binaria. En segundo lugar, el hecho de que las preguntas se deben proyectar teniendo en cuenta la forma final de la escala, un conjunto de elementos de intensidad creciente.

A diferencia de la escala de Likert, en la que todas preguntas podían tener más o menos la misma intensidad, en el caso de las escalas de Guttman éstas deben cubrir toda la extensión del continuo de la actitud subyacente, a fin de obtener una batería de preguntas con dificultad creciente. Por ejemplo, para una escala sobre la actitud progresista-conservadora en política, el investigador deberá formular

Tabla 6.6. Respuestas a la pregunta: ¿Estaría dispuesto a aceptar a un individuo de color como visitante de su país, como vecino de casa, como amigo o como cónyuge?

Como visitante	Como vecino	Como amigo	Me casaría con él	Puntuación
1	1	1	1	4
1	1	1	0	3
1	1	0	0	2
1	0	0	0	1
0	0	0	0	0

una serie de afirmaciones variando el campo de aplicación (de la política fiscal, políticas de empleo, valoración de hechos históricos como el fascismo, etc.), y también proponer afirmaciones graduales desde el extremo revolucionario al extremo reaccionario, pasando por todas las posiciones intermedias.

En cuanto a la fase de aplicación a una muestra de sujetos, la escala de Guttman no plantea problemas muy distintos de los que planteaba la escala de Likert. La diferencia más notable reside en el carácter binario de las opciones de respuesta. La verdadera especificidad de las escalas de Guttman se manifiesta en el momento de hacer el análisis de los resultados. Igual que para las escalas de Likert, el objetivo de esta fase es evaluar la escalabilidad de todos los elementos, descartar los menos coherentes con el modelo, establecer un índice de escalabilidad de la escala y establecer su aceptación o no.

El procedimiento consiste, en primer lugar, en identificar los llamados «errores» de la escala, o sea, las respuestas que no se integran en las secuencias previstas en el modelo. Existen dos técnicas distintas para su identificación: una es la que adoptó originalmente Guttman, y la otra se propuso después [Edwards, 1957]. Presentaremos sólo esta última, puesto que se considera la más adecuada.

El procedimiento se desarrolla como sigue. Supongamos que la matriz de los datos inicial sea la que representamos en la tabla 6.7a (para simplificar imaginemos que sólo hemos aplicado la escala a ocho sujetos). Las columnas y las filas de la matriz se reordenan, de manera que las preguntas de la escala estén ordenadas de izquierda a derecha según el número de respuestas positivas recibidas, y los casos estén ordenados de arriba a abajo según la puntuación total recibida. De esta forma se visualiza la estructura de dos triángulos de la matriz (uno compuesto predominantemente por 1 y el otro por 0, como se muestra en la tabla 6.7b).

La tabla 6.7b sugiere que hay dos secuencias de error: el caso D y el caso G. Los errores se identifican comparando la secuencia observada en cada caso con la secuencia teórica correcta que debería obtenerse *con la puntuación total de ese caso concreto.* Consideremos el caso D: la secuencia observada es 10110, con una puntuación total de tres. La secuencia correcta para esa puntuación total debería ser 11100, por lo tanto, hay dos errores (la segunda posición debería ser un 1 y la

cuarta debería ser un 0).

Una vez identificados de este modo los «errores» (que hemos marcado con un asterisco en la tabla 6.7b), se trata de decidir si son lo bastante irrelevantes como para garantizar la «escalabilidad» de las preguntas de la escala. Guttman propuso un índice, que llamó coeficiente de reproducibilidad, que mide el grado de desviación de la escala observada respecto a la escala perfecta. Su fórmula es la siguiente:

$$C_r = 1 - \frac{\text{Número de errores}}{\text{Número total de respuestas}} = 1 - \frac{\text{Número de errores}}{\text{Número de elementos} \times \text{número de casos}} = \frac{\text{Número de respuestas correctas}}{\text{Número total de respuestas}}$$

En nuestro ejemplo:

$$C_r = 1 - \frac{4}{5,8} = 0,90$$

Este índice también se puede interpretar como la proporción de respuestas «correctas» (correspondientes a las secuencias teóricas) sobre el número total de respuestas. Guttman sugirió un valor de $C_r = 0,90$ (es decir, errores iguales o inferiores al 10 % de las respuestas) para la aceptación de la escala [13]. Si la escala presenta un coeficiente de reproducibilidad inferior a 0,90, se van eliminando los elementos con mayor número de errores, y se va recalculando el índice C_r.

El coeficiente de reproducibilidad de la escala es la media de los coeficientes de reproducibilidad de cada uno de los elementos (proporción de respuestas correctas / respuestas totales de cada elemento). A partir de ahí se puede demostrar que el coeficiente de reproducibilidad de cada elemento no puede ser inferior a su proporción de respuestas en la categoría modal [14]. Esto implica que el coeficiente puede estar inflado debido a una distribución de las respuestas muy asimétrica.

Por ejemplo, la respuesta modal del elemento 1 de la tabla 6.7 es «no» y su proporción es 0,75 (seis «no» de un total de ocho respuestas), por lo que el coeficiente de reproducibilidad del elemento número 1 no puede descender por debajo de 0,75. Por lo tanto, una escala de Guttman compuesta íntegramente por afirmaciones extremas que provocan respuestas como un 90 % de «sí» frente a un 10 % de «no» (o a la inversa) obtiene automáticamente un C_r en torno al 90 %.

Una vez eliminados los elementos no escalables y establecida la aceptación de la

[13] El número de errores depende de la forma en que se calculan éstos, por lo que el umbral de aceptación de $C_r = 0,90$ resulta ser un criterio más severo si el para calcular estos errores utilizamos el procedimiento de Goodenough-Edwards.

[14] La categoría modal (o «moda») de una variable nominal es la que presenta mayor frecuencia. Si una variable presenta el 40 % de respuestas «sí» y el 60 % de respuestas «no», la moda es la categoría «no» y su proporción es de 0,60; si las respuestas han sido del 90% y el 10%, la proporción de la categoría modal es 0,90.

Tabla 6.7. Identificación de los errores de una escala de Guttman

Casos	Elementos n 1	n 2	n 3	n 4	n 5	Puntuación
a) matriz originaria						
A	1	1	1	1	1	5
B	0	0	0	0	1	1
C	0	0	1	0	1	2
D	0	1	0	1	1	3
E	0	1	1	0	1	3
F	0	0	0	0	0	0
G	1	1	1	0	0	2
H	0	1	1	1	1	4
N.° de respuestas afirmativas	2	4	5	3	6	
	n 5	n 3	n 2	n 4	n 1	
b) matriz ordenada						
A	1	1	1	1	1	5
H	1	1	1	1	0	4
E	1	1	1	0	0	3
D	1	0*	1	1*	0	3
C	1	1	0	0	0	2
G	0*	1	0	0	1*	2
B	1	0	0	0	0	1
F	0	0	0	0	0	0
N.° de errores	1	1	0	1	1	

* Error

escala, se trata de atribuir puntuaciones a los sujetos. Para los sujetos que presentan una secuencia de respuestas acorde con el modelo no hay ningún problema: la puntuación total de la escala para cada sujeto es la suma de las puntuaciones 0/1 obtenidas en cada elemento. Aunque el sujeto presente una secuencia con errores (por ejemplo, 01010), se respeta el mismo criterio si se ha seguido la técnica de determi-

nación de los errores aquí presentada: se considera correcta la secuencia aceptada por el modelo que tenga la misma puntuación final (en el ejemplo: 11000), y por lo tanto la puntuación del sujeto se mantiene igual a la suma de las respuestas positivas que ha dado, aunque presente errores. En este punto, el proceso de construcción de la escala se considera concluido.

Se han realizado diversas sugerencias para mejorar el rendimiento de esta técnica [15]. Aquí vamos a presentar tres de ellas, recomendadas por el mismo Guttman. La primera es evitar preguntas con tasas demasiado elevadas (superiores al 80 %) de aceptación o rechazo, ya que son poco discriminantes y elevan de manera artificial el coeficiente de reproducibilidad de la escala. En segundo lugar, Guttman sugirió

Resumen 6.6. MEDIDAS DE REPRODUCIBILIDAD

Edwards [1957] propuso calcular un *índice de reproducibilidad marginal mínima (MMR, minimal marginal reproducibility)* con la siguiente fórmula:

$$MMR = \frac{\Sigma \text{ proporción de respuestas en la categoría modal}}{N}$$

donde N es el número de elementos de la escala

En nuestro ejemplo de la tabla 6.7:

$$MMR = \frac{0{,}75 + 0{,}50 + 0{,}63 + 0{,}63 + 0{,}75}{5} = 0{,}65$$

Este índice indica el valor mínimo que puede adoptar el coeficiente de reproducibilidad de la escala, cualesquiera que sean las secuencias de las respuestas. Se sugiere, por tanto, comparar el C_r con el *MMR*: sólo si el primero es superior a 0,90 y también claramente superior al segundo se puede afirmar que la adecuada reproducibilidad de la escala se debe a una escalabilidad efectiva de sus elementos y no a la distribución marginal de las respuestas. En nuestro ejemplo se cumple este requisito.

[15] Señalemos que la técnica de Guttman también puede emplearse con respuestas politómicas, transformándolas en dicotómicas mediante la agregación de las categorías (el punto de corte puede elegirse de tal modo que se maximice la escalabilidad de la escala; esto se consigue, por ejemplo, unificando las respuestas «muy de acuerdo» y «bastante de acuerdo» para ciertos elementos, y contrastando los «muy de acuerdo» con el resto de las respuestas para otros elementos).

emplear un número lo bastante elevado de preguntas, mostrando que una escala con sólo cuatro preguntas puede producir valores elevados de C_r aunque las preguntas sean estadísticamente independientes entre sí. El mismo Guttman recomendó estudiar atentamente las secuencias erróneas, pues la presencia repetida de una secuencia no prevista en el modelo puede ser síntoma de otra dimensión subyacente.

Podemos concluir diciendo que el escalograma de Guttman ha supuesto un avance muy importante en el desarrollo de la técnica de las escalas. En el debate en torno a éste que tuvo lugar durante al menos treinta años, se propusieron numerosas sugerencias para su mejora [cfr. McIver y Carmines, 1981]. No obstante, sigue presentando algunos aspectos problemáticos. En primer lugar, la puntuación final que se obtiene sobre la variable latente sigue siendo fundamentalmente, como sucedía en las escalas de Likert, una variable ordinal, y tampoco en este caso tenemos ningún elemento para afirmar, como sucedía con la participación política, que la distancia entre las puntuaciones 2 y 4 es igual que la distancia entre las puntuaciones 5 y 7.

En segundo lugar, la técnica de Guttman es aplicable cuando estamos ante actitudes (comportamientos, acontecimientos...) bien definidas y escalables. Cuando la actitud se hace compleja resulta difícil escalonarla en secuencias acumulativas, las categorías se pueden superponer, por lo que al final hay demasiados errores.

Por último, el modelo es determinista e inflexible, pero la realidad social sólo puede interpretarse correctamente mediante modelos probabilísticos que contemplen la posibilidad de error y el paso gradual entre las distintas posiciones. Ésta es la crítica más importante al modelo de Guttman. De hecho, su escala se considera ya superada por las escalas de tipo probabilístico, como las escalas de Rasch.

6.5. DIFERENCIAL SEMÁNTICO

La técnica del diferencial semántico fue desarrollada en los años cincuenta por el psicólogo estadounidense Charles Osgood y sus colegas [Osgood, 1952; Osgood, Suci y Tannenbaum, 1957]. El objetivo de estos investigadores era crear un instrumento para determinar, con la máxima estandarización, el significado que los conceptos adquieren para los individuos, un instrumento capaz responder a la pregunta: «¿qué significa esto (un concepto, como madre, patria, guerra, etc.) para usted?».

El problema del «significado» que un término, una situación social, un objeto cognitivo en general adquiere para un individuo es complejo, está cargado de elementos filosóficos y psicológicos y tiene una enorme importancia para el estudio de los comportamientos. Como escribe Nunnally, «el comportamiento humano está determinado por el significado de los acontecimientos más que por sus propiedades intrínsecas; el niño reacciona positivamente a la voz de su madre porque para él significa alimento, calor, protección» [1959, 383].

Osgood y sus colegas lo plantearon como sigue: «Cuando queremos descubrir lo que significa para cierta persona una cosa determinada, le pedimos que nos lo diga [...]. [Podemos preguntar]: ¿qué significa para usted la palabra «sofisticado»? La respuesta podría ser: sé lo que significa, aunque me resulte difícil decirlo [...]; significa ser inteligente y conocer las cosas y las personas [...], saber mucho. Es una

forma de ser relajada y refinada, llena de gracia y ligereza [...], equilibrada [...], quiero decir, tener *savoir faire*» [Osgood, Suci y Tannenbaum, 1957, 18].

Cuando los sujetos estudiados son personas inteligentes, con formación y con excelentes dotes de expresión verbal, esta forma de conocer los significados a partir de la descripción que ellos mismos nos ofrecen, puede ser válida y lo bastante sensible: estas personas son capaces de expresar, mediante una descripción verbal, los elementos distintivos y determinantes de lo que ese concepto significa para ellos. Pero con individuos que no tienen esta capacidad de expresión verbal, este sistema puede ser muy limitado: las descripciones emitidas de manera espontánea pueden resultar demasiado pobres. Además, resulta difícil comparar entre sí todas las descripciones obtenidas con estos procedimientos tan subjetivos.

Con el diferencial semántico Osgood propone una técnica para extraer los significados que determinados conceptos adquieren para los entrevistados, que, en lugar de basarse en la descripción subjetiva y directa del significado por parte del entrevistado, se basa en las asociaciones que establece el entrevistado entre este concepto y otros propuestos de forma estandarizada a todos los entrevistados. En lugar de preguntar: ¿qué significa para usted el término «sofisticado»?, se plantea una serie de preguntas como: ¿«sofisticado» es duro o blando?, ¿agradable o desagradable?, ¿rápido o lento? Y así sucesivamente. Y para mejorar la sensibilidad del instrumento se pide que se clasifique la intensidad de cada opinión en una escala de siete puntos. «De esta forma el diferencial semántico es básicamente una combinación de asociaciones controladas y *escalas de intensidad*» [*ibídem,* 20].

Para explicar el formato de la pregunta, consideremos el ejemplo presentado en la figura 6.1, que procede de una investigación realizada en Italia sobre los valores de los jóvenes, en la que se pidió a los entrevistados que evaluaran a cuatro partidos políticos [Sciolla y Ricolfi, 1989]. Esta técnica emplea una serie de escalas de intervalos de autoposicionamiento típicas, en las que las categorías extremas tienen significado autónomo, mientras que el significado (gradual) de las categorías intermedias lo establece el entrevistado por comparación. En la tabla 6.8 hemos reproducido 50 parejas de atributos polares, empleadas por Osgood y sus colegas en sus investigaciones. Ellos aplicaron estas escalas a personas (extranjero, madre, yo mismo, políticos, etc.), objetos concretos (cuchillo, nieve, motor, etc.), objetos abstractos (arte moderno, pecado, liderazgo, etc.), acontecimientos (debate, nacimiento, etc.), e instituciones (Naciones Unidas, hospital, familia, etc.).

Osgood sostenía que la lista de los atributos bipolares no debe tener relación con el objeto evaluado, y por tanto debe ser siempre la misma. Incluso ciertos adjetivos que no tienen nada que ver con el objeto en cuestión resultan útiles para extraer significados más profundos o inconscientes que tiene el objeto para el entrevistado[16].

El test se puede realizar fácilmente y suele ser bien aceptado por los entrevistados, incluso cuando las parejas de atributos parecen poco pertinentes en relación

[16] Sin embargo el propio Osgood modificó la lista para adaptarla a cada investigación. Además, la técnica se ha aplicado muchas veces utilizando atributos relacionados específicamente con el objeto a evaluar, junto con las parejas tradicionales.

con el objeto al que se aplican; en general, se recomienda al entrevistado que conteste de forma instintiva, sin reflexionar demasiado sobre cada pareja.

En cuanto al número de adjetivos bipolares propuestos para un mismo objeto, la técnica es muy flexible. En los dos ejemplos que hemos presentado abarca desde los 12 en la investigación sobre partidos políticos hasta los 50 propuestos inicialmente por Osgood. De forma análoga, los objetos sometidos a evaluación pueden ser muchos (por ejemplo, una serie de 30 personajes públicos del mundo del espectáculo, de la política, de la economía, etc.), o pocos (por ejemplo, los cuatro partidos políticos que hemos mencionado). La elección dependerá del objetivo del estudio y la relevancia que tenga la técnica en el marco general de la investigación.

En la aplicación a los partidos políticos, el diferencial semántico se incluía dentro de un cuestionario que utilizaba otros instrumentos para revelar las actitudes, desde preguntas tradicionales hasta escalas Likert. De ahí su uso limitado, en cuanto al número de objetos evaluados y el número de parejas de atributos empleadas. Osgood y sus colegas citan otras aplicaciones, en el campo del estudio de la personalidad, basadas totalmente en la técnica del diferencial semántico, y en las que el número de juicios solicitados (que la cifra que resulta de multiplicar el número de atributos bipolares y el número de objetos a evaluar) puede llegar hasta 400 (por ejemplo, 40 parejas polares y diez objetos a evaluar).

Las respuestas obtenidas se pueden someter a diversas operaciones, que aquí sólo citaremos brevemente. Para describir de manera sintética la imagen del objeto

Pregunta: *«Hablemos ahora de los partidos italianos. Debe decirnos su opinión de cada uno de ellos, tratando de situarlo en cada una de las escalas siguientes» (pregunta aplicada a los partidos Socialista, Comunista, Democracia Cristiana y Verdes [Sciolla y Ricolfi, 1989]).*

Partido Socialista

	1	2	3	4	5	6	7	
Viejo								Joven
Fuerte								Débil
Limpio								Sucio
Rápido								Lento
Cercano								Lejano
Intolerante								Tolerante
Abierto								Cerrado
Intransigente								Indulgente
Ganador								Perdedor
De izquierdas								De derechas
Alegre								Triste
Ineficiente								Eficiente

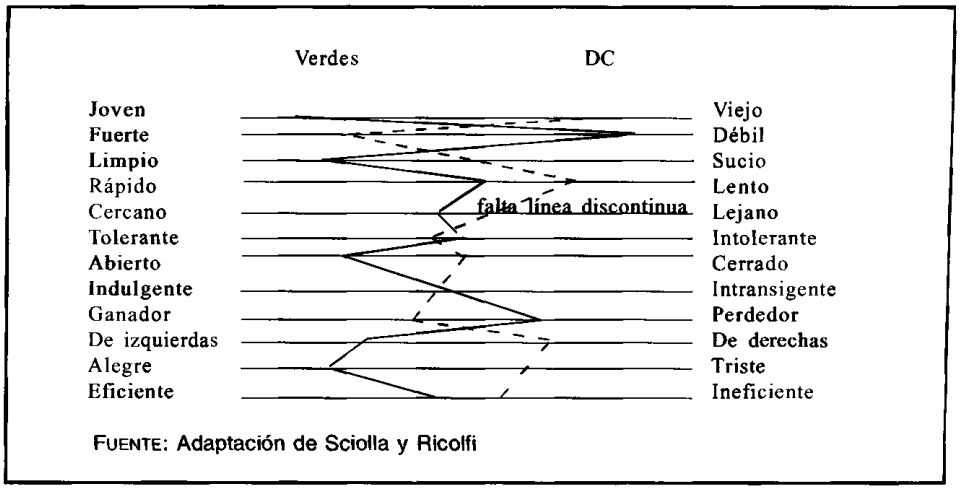
Notas para el entrevistador:
· mostrar la ficha al entrevistado y pedirle que ponga una cruz sobre la posición que se corresponda mejor con su imagen del partido para cada una de las escalas indicadas;
· es mejor que las respuestas se den instintivamente, sin reflexionar demasiado tiempo en cada escala.

Figura 6.1. Ejemplo de aplicación del diferencial semántico

sometido a valoración, se puede emplear el «perfil». Este perfil es una representación gráfica que se construye uniendo con una línea la puntuación media obtenida por ese objeto en cada pareja de atributos. En la figura 6.2 hemos reproducido los perfiles de los Verdes y la Democracia Cristiana de la aplicación del diferencial

Tabla 6.8. Diferencial semántico. Parejas de atributos empleadas por Osgood, Suci y Tannenbaum [1957]

1. Bueno-malo	18. Agradable-desagradable	35.Simpático-antipático
2. Grande-pequeño	19. Negro-blanco	36. Brillante-apagado
3. Bonito-feo	20. Amargo-dulce	37. Grave-agudo
4. Amarillo-azul	21. Feliz-triste	38. Anguloso-circular
5. Duro-blando	22. Gracioso-soso	39. Perfumado-hediondo
6. Dulce-amargo	23. Vacío-lleno	40. Honesto-deshonesto
7. Fuerte-débil	24. Agresivo-pacífico	41. Activo-pasivo
8. Limpio-sucio	25. Pesado-ligero	42. Rugoso-liso
9. Alto-bajo	26. Mojado-seco	43. Fresco-rancio
10. Tranquilo-inquieto	27. Sagrado-profano	44. Rápido-lento
11. Sabroso-insípido	28. Relajado-tenso	45. Justo-injusto
12. Valioso-sin valor	29. Valiente-cobarde	46. Áspero-suave
13. Rojo-verde	30. Alto-bajo	47. Cercano-lejano
14. Joven-viejo	31. Rico-pobre	48. Agitado-apacible
15. Amable-grosero	32. Luminoso-oscuro	49. Sano-enfermo
16. En voz alta-baja	33. Caliente-frío	50. Ancho-estrecho
17. Profundo-superficial	34. Espeso-ligero	



FUENTE: Adaptación de Sciolla y Ricolfi

Figura 6.2. Ejemplos de «perfiles» en el diferencial semántico

semántico a los partidos políticos mencionada antes; para facilitar la lectura del gráfico hemos indicado a la izquierda la parte «positiva» de cada pareja de atributos. La imagen de la DC que obtenemos de estos perfiles es la de un partido «viejo, de derechas, ineficiente, sucio, cerrado, lento, triste», mientras que los Verdes son «jóvenes, rápidos, alegres, cercanos, limpios, abiertos, tolerantes, débiles y perdedores»; en la misma investigación el PCI resultó ser «eficiente, de izquierdas, severo y poco tolerante», mientras que el partido socialista se percibía como un partido «indulgente y lejano» [Sciolla y Ricolfi, 1989, 109-110].

En este caso, el perfil está centrado en el concepto y consiste en las valoraciones medias de los entrevistados, pero también se pueden obtener perfiles individuales, es decir, de los sujetos entrevistados, basados en las respuestas de cada uno de los individuos. Por ejemplo, la técnica se ha empleado como instrumento para conocer el nivel de autoestima del entrevistado, pidiéndole que coloque en una batería de atributos bipolares primero el objeto cognitivo «yo, tal como soy» y luego «yo, tal como me gustaría ser». La distancia entre ambos perfiles permite conocer el nivel de autoestima del sujeto y, para citar una aplicación de la técnica, la evolución en el tiempo de esta distancia puede emplearse para seguir el avance de un paciente durante un tratamiento con psicoterapia.

Sin embargo, la aplicación más importante del diferencial semántico es la exploración de la dimensión de los significados. Si pedimos a un entrevistado que valore un determinado objeto cognitivo a partir de 30 parejas de adjetivos, obviamente las distintas valoraciones estarán guiadas por la visión de conjunto (el «significado») que el entrevistado tiene del objeto en cuestión. La condición es que esta visión no sea unidimensional, sino que presente diversos aspectos que contribuyan al significado global de ese objeto. Mediante el análisis factorial se puede determinar cuáles son las dimensiones básicas que subyacen a las valoraciones de una determinada muestra de individuos entrevistados.

Utilizando la serie de 50 adjetivos bipolares (los que aparecen en la tabla 6.8) para una serie desigual de objetos cognitivos y repitiendo la encuesta sobre muestras distintas de entrevistados, Osgood y sus colegas constataron que, en líneas generales, pueden identificarse tres dimensiones básicas que subyacen a las diversas opiniones: la *evaluación* (relacionada con las parejas bueno-malo, positivo-negativo, bello-feo, justo-injusto, valioso-sin valor, etc.), la *potencia* (duro-blando, fuerte-débil, pesado-ligero, etc.), y la *actividad* (activo-pasivo, rápido-lento, caliente-frío, etc.). De estas tres dimensiones la primera es la más relevante, y parece representar lo que en las técnicas de escalas tradicionales se solía denominar la «actitud» (favorable o desfavorable) hacia un determinado objeto cognitivo.

A partir de esta clasificación de las dimensiones del espacio semántico, pueden sumarse [17] los valores para cada atributo bipolar (según la dimensión a la que pertenece el atributo), obteniendo así para cada individuo unas puntuaciones globales sobre las tres dimensiones mencionadas, como sucedería con una escala aditiva de

[17] Se tratará de una suma ponderada, utilizando las puntuaciones factoriales obtenidas del análisis factorial.

tipo Likert [18]. La puntuación sobre la dimensión de «evaluación» es la que mejor se presta a su comparación con las puntuaciones obtenidas con otras técnicas de escalas: Osgood y sus colegas encontraron importantes correlaciones entre el factor «evaluación» del diferencial semántico y las puntuaciones obtenidas con las escalas de Thurstone (ambas técnicas se utilizaron para estudiar las actitudes en relación con la Iglesia, la pena capital y las personas de color), y con las escalas de Guttman.

Osgood y sus colegas afirmaban que las tres dimensiones de evaluación, potencia y actividad rigen el espacio semántico de los individuos en la gran mayoría de los casos, independientemente de la cultura de los sujetos (realizaron estudios en varios países) y de los conceptos sometidos a evaluación. Muchos investigadores no comparten esta afirmación generalizadora, y algunas investigaciones han encontrado dimensiones que no coinciden con ninguna de las tres propuestas por Osgood.

Aunque estas objeciones puedan contradecir una cierta teorización sobre el espacio semántico que Osgood consideraba tan importante, no restan interés a su técnica como instrumento para revelar la *estructura* de las actitudes. Probablemente la contribución más original del diferencial semántico al reconocimiento de las actitudes sea haber desvelado las múltiples dimensiones de los significados, al introducir la riqueza del punto de vista multidimensional en un campo que hasta entonces había aplicado una perspectiva unidimensional (volveremos sobre esto al final del capítulo).

6.6. EL TEST SOCIOMÉTRICO

La técnica que presentamos en este apartado tiene un objetivo bien distinto del que tienen en común las técnicas presentadas hasta ahora. De hecho, no tiene como finalidad revelar variables latentes como la religiosidad o la depresión, ni actitudes con respecto a determinados objetos cognitivos como los partidos políticos o los inmigrantes. Las técnicas sociométricas nacieron con el fin de revelar las relaciones interpersonales existentes dentro de un grupo de individuos. Tal vez convenga recordar que el objeto de este capítulo es presentar las principales técnicas de las escalas, es decir, los procedimientos elaborados por las ciencias del comportamiento, sobre todo entre los años treinta y sesenta, con el objetivo de «medir al hombre y a la sociedad». Y aunque no han cumplido ese objetivo de «medición», han producido técnicas que siguen siendo válidas hoy día para el estudio del hombre y sus comportamientos.

El test sociométrico fue propuesto entre 1940 y 1950 por Jacob Moreno [1953], un académico multidisciplinar, médico, terapeuta y sociólogo, de origen rumano, pero estadounidense de adopción, y nace con el objetivo de estudiar de manea sistemática y con procedimientos estandarizados, las relaciones interpersonales que se

[18] Con la diferencia de que en la escala de Likert las respuestas tienen autonomía semántica parcial (como: totalmente, bastante, poco, en absoluto de acuerdo), mientras que en este caso las respuestas están representadas por gradaciones de intervalos (de autoposicionamiento).

establecen dentro de grupos pequeños (o relativamente pequeños). Su campo de aplicación ideal es una clase escolar [19]. George Bastin comienza su libro sobre esta técnica con la descripción de un experimento muy sencillo. Escogió doce clases de primero de enseñanza primaria de un centro escolar, y tres meses después del inicio del curso entrevistó a cuatro profesores por grupo, de los que daban clase todos los días, pidiéndoles que describieran las relaciones que se habían establecido entre los alumnos: las amistades, los subgrupos, los estudiantes aislados, los líderes, etc. Los profesores no fueron capaces de desentrañar las redes de relaciones existentes entre sus alumnos. Simultáneamente Bastin aplicó un cuestionario sociométrico a las clases y éste reveló con nitidez el mapa de las relaciones entre los chicos, mostrando la existencia de dinámicas que los educadores no habían percibido.

En su formulación más sencilla, el test sociométrico consiste en un cuestionario compuesto por unas cuantas preguntas, que giran en torno al tema de la preferencia o rechazo por el resto de los individuos de un grupo. Siguiendo con el ejemplo escolar, las cuatro preguntas siguientes planteadas a los alumnos de la clase son suficientes para realizar un test sociométrico completo [Bastin, 1961]:

1. ¿De tus compañeros actuales, cuáles te gustaría volver a tener en tu clase el año que viene? Indica todos los que quieras, en orden de preferencia.
2. ¿Y cuáles son los compañeros que no te gustaría volver a tener en tu clase el año que viene? Indica todos los que quieras, ordenados empezando por el que menos te gustaría tener de compañero el año próximo.
3. Adivina cuáles de tus compañeros han dicho que les gustaría estar contigo en clase el año que viene.
4. Adivina cuáles de tus compañeros han dicho que no les gustaría estar contigo en clase el año que viene.

En este ejemplo la encuesta se realizó con motivo del fin del curso escolar y el paso a un nuevo ciclo que da lugar a la formación de nuevas clases, pero las preguntas pueden girar en torno a temas similares: los compañeros con los que te gustaría (no te gustaría) jugar, estudiar, formar un grupo de trabajo, ir al cine, a los que querrías (no querrías) invitar a tu casa, a tu fiesta de cumpleaños, etc.

La elaboración de los datos es tan sencilla como la aplicación del cuestionario. Se construye una matriz cuadrada, en cuyas partes izquierda y superior se sitúa la lista de los componentes del grupo, y se registran las preferencias o rechazos en las casillas correspondientes a cada pareja.

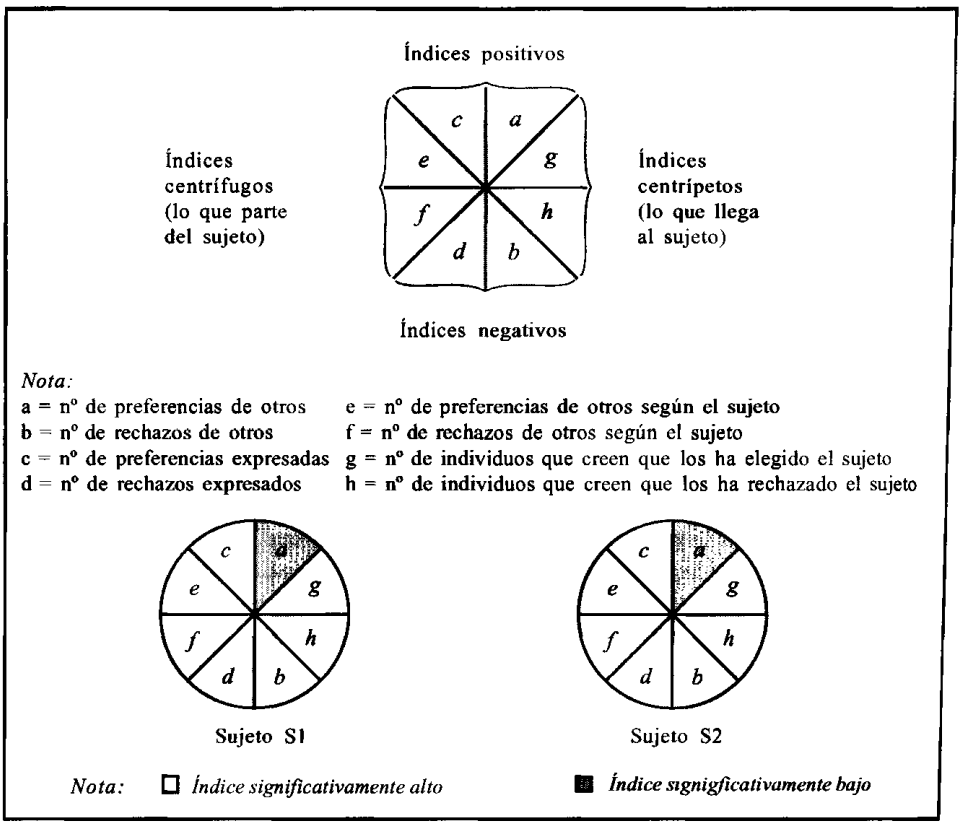
La primera operación que realiza con los datos es el análisis del *estatus socio-*

[19] La figura de Moreno es bastante singular: unió su vocación de investigador a la de profeta y cautivador para fundar un movimiento orientado a la construcción de una sociedad utópica basada en la cooperación, la solidaridad y la libertad. Con el nombre de «sociometría» se suele designar la totalidad de propuestas de Moreno, entre ellas, las técnicas del psicodrama y el sociodrama, orientadas a la liberación de la creatividad y la espontaneidad, con objetivos fundamentalmente terapéuticos. Aquí nos limitaremos a estudiar la técnica propuesta por Moreno para el estudio de las relaciones existentes dentro de los grupos, conocida como «test sociométrico».

métrico individual, que consiste en analizar para cada individuo el número de preferencias emitidas y el número de rechazos recibidos. En función de las preguntas arriba mencionadas pueden elaborarse otros índices básicos, que nos ofrecerán un retrato psicosocial del sujeto.

En la figura 6.3 hemos reproducido estos índices y el gráfico para dos sujetos. Hemos tomado estos dos casos aislados, porque hubo muy pocos compañeros que los eligieron (el índice *a* fue muy bajo) y muchos que los rechazaron (el índice *b* fue muy alto); además, también hubo muchos compañeros que se consideraban rechazados por ellos (índice *h* alto). En esta situación, el sujeto S_1 percibe su posición de excluido (índice *f* de rechazos que cree haber recibido alto) y reacciona expresando a su vez muchos rechazos (índice *d* alto), mientras que el sujeto S_2 parece no ser consciente de esta situación de aislamiento (índice *f* no significativo) y no reacciona ante ello (índice *d* de rechazos no significativo).

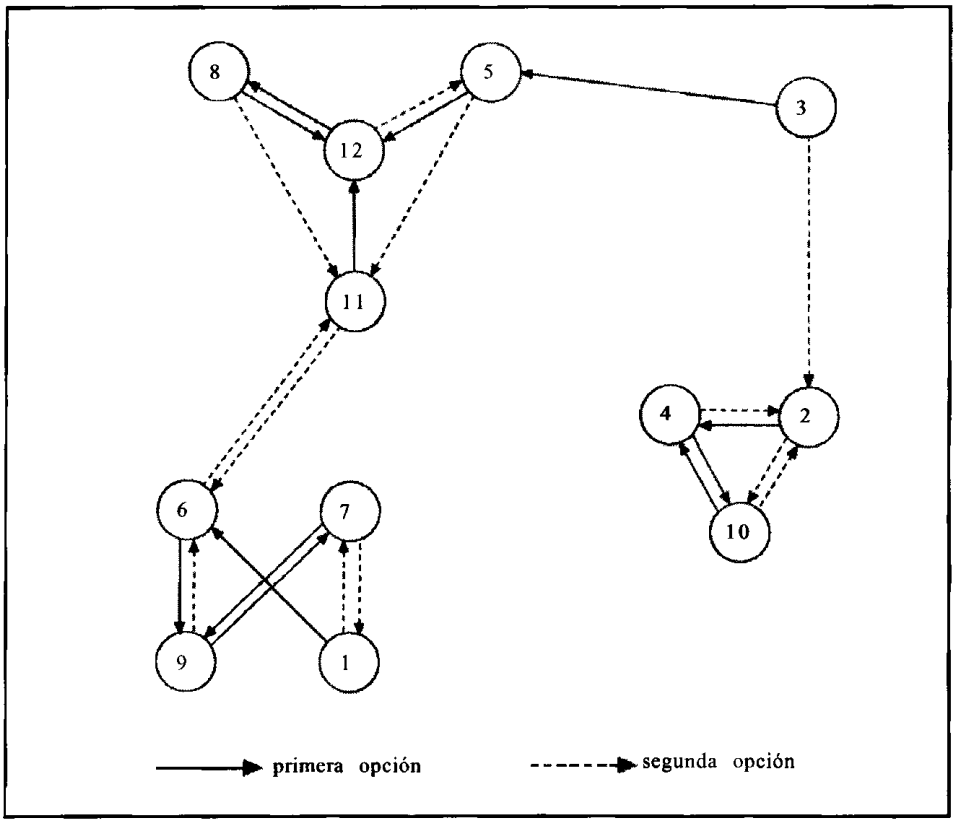
Además de los análisis para definir la situación de los individuos en el aspecto



FUENTE: Bastin [1961, 44]

Figura 6.3. Índices sociométricos individuales.

de las relaciones con los demás, se realizan operaciones para ilustrar la estructura sociométrica del grupo. La figura 6.4 muestra una sencilla representación o sociograma de las dos primeras preferencias en un grupo de doce estudiantes [Nunnaly, 1959, 392]. Como puede verse en el gráfico, el grupo se organiza en tres subgrupos (todas las primeras opciones de los sujetos y casi todas las segundas se agotan en cada subgrupo), más una persona aislada (3). El grupo formado por los sujetos 5, 8, 11 y 12 presenta un líder (el sujeto 12, que es la primera opción de todos los demás), mientras que los otros dos grupos presentan una estructura más paritaria (aunque las diferencias entre los sujetos son claras: por ejemplo, en el grupo 1, 6, 7 y 9 el sujeto 1 está más aislado que los otros, ya que sólo es una segunda opción). Dos de los tres grupos están conectados entre sí por los sujetos 6 y 11, que funcionan como intermediarios, mientras que el tercero (2, 4 y 10) es un grupo cerrado, ya que todas las preferencias quedan dentro del mismo. El estado de aislamiento y desorientación del sujeto 3 es aun más evidente por el hecho de que, además de no



FUENTE: Nunnaly [1959, 392].

Figura 6.4. Sociograma de las dos primeras preferencias en un grupo de 12 estudiantes.

ser escogido por ningún otro sujeto, él escoge a sujetos pertenecientes a grupos diferentes.

Moreno y los investigadores que han seguido el planteamiento sociométrico propusieron otras muchas formas de representar gráficamente la estructura del grupo y la condición de los individuos, como diagramas de círculos concéntricos en cuyo centro se encuentran los individuos más populares, y hacia el exterior los individuos más aislados. También se pueden construir sociogramas distintos según las preguntas que se consideren (en la figura 6.4 hemos visualizado sólo las preguntas sobre elecciones de sujetos, pero también están los rechazos, las elecciones y rechazos supuestos, las elecciones y los rechazos recíprocos, etc.).

Para un estudio más detallado de este tema se recomiendo consultar la literatura especializada. Aquí nos limitaremos a decir que la técnica resulta útil como instrumento de *diagnóstico individual,* con el fin de identificar las relaciones de dominación-dependencia, de aislamiento, y de afinidad (simpatía, amistad) y conflicto (antipatías, enemistades), como instrumento para revelar la *estructura relacional del grupo,* las redes de comunicación, la organización jerárquica informal, las vías por las cuales discurre la información, los comentarios personales, las órdenes, etc., y como herramienta para *estudiar la psicología de los grupos,* identificar los puntos de tensión, las estratificaciones sociales, las barreras étnicas, religiosas, lingüísticas, de género, de edad, etc.

En general, el test sociométrico es adecuado para el estudio de grupos estructurados y menos apropiado para el análisis de grupos informales (los amigos del bar o los jóvenes que juegan juntos en la playa), puesto que el radio de elección del sujeto debe estar limitado. Ya hemos citado su gran utilidad para el estudio de las dinámicas en las clases escolares. Otro campo de aplicación ha sido el laboral (departamentos de fábricas, organización empresarial); también podemos mencionar estudios sobre la organización militar, donde podría analizarse, por ejemplo, si la jerarquía formal se corresponde con la real, si las órdenes siguen el canal de comunicación jerárquico o uno alternativo, etc.

Con respecto a estas últimas aplicaciones, originalmente la técnica se creó para estudiar las relaciones afectivas existentes dentro de los grupos, es decir, criterios de elección de tipo «afectivo», basados en la simpatía, amistad, admiración, con preguntas como: ¿con quién te gustaría jugar?, ¿a quién invitarías a cenar? Posteriormente la técnica se ha aplicado también al estudio de las relaciones funcionales, es decir, criterios de elección de tipo «funcional», basados en la estimación de las capacidades, las competencias y la utilidad, con preguntas como: ¿con quién querrías estudiar?, ¿a quién elegirías para tu equipo de fútbol?, ¿a quién querrías tener en tu equipo de trabajo? Por último, esta técnica también se ha utilizado, aunque con menos aplicaciones, para estudiar los procesos comunicativos o «diagramas de interacción» dentro de las organizaciones formales.

La técnica propuesta por Moreno fue muy popular en los años cincuenta y sesenta, periodo en el que se aplicaba en múltiples campos. Poco a poco su utilización para la investigación se fue reduciendo, en parte por la disgregación de la construcción teórica de la «sociometría», elaborada por Moreno con un fin de renovación social más que de investigación científica. Más recientemente, el estudio sistemáti-

co de las relaciones sociales ha cobrado vigor en el marco del análisis de redes sociales, que recibió un fuerte impulso gracias al desarrollo de la informática y las nuevas posibilidades que ofrecen los ordenadores para la gestión y el análisis de datos relacionales [cfr. Knoke y Kulinski 1982; Wasserman y Faust 1994].

6.7. ESCALAS UNIDIMENSIONALES Y ESCALAS MULTIDIMENSIONALES

En los dos últimos apartados nos hemos alejado en parte del tema central de este capítulo, la técnica de las escalas para la operacionalización empírica de conceptos complejos. En los primeros apartados hemos esbozado el camino que recorrieron las ciencias sociales en sus intentos de «medir» al hombre y su mente, pero hemos hablado sólo de las escalas denominadas «unidimensionales», es decir, de las situaciones de investigación en las que se presume que existe una sola dimensión fundamental subyacente a un conjunto de observaciones.

El razonamiento utilizado ha sido el siguiente: queremos registrar un concepto complejo o latente como la alienación o el autoritarismo, ¿cómo podemos hacerlo? A partir de consideraciones puramente teóricas, los investigadores elegimos unas expresiones explícitas (aunque parciales) de este concepto latente, que formulamos como opiniones, las presentamos a los sujetos entrevistados, verificamos que estas variables también estén relacionadas entre sí en las respuestas de los entrevistados (para comprobar que están determinadas por una única propiedad subyacente), y utilizamos las respuestas de cada individuo para operacionalizar el concepto latente.

El punto de partida en esta secuencia es el concepto general que se quiere registrar, y para operacionalizarlo se identifican unos indicadores. Pero hay otra forma de abordar el problema de los conceptos complejos subyacentes a un conjunto de observaciones, que consiste en partir directamente de las observaciones, preguntándonos *cuántas y qué* dimensiones latentes subyacen a ellas. Ésta es la técnica de las denominadas *escalas multidimensionales.*

Consideremos, por ejemplo, la investigación realizada en 1968 por Weisberg y Rusk [1970] en torno a la opinión de los electores sobre los candidatos a las elecciones presidenciales estadounidenses de ese año. Las preferencias de los entrevistados se registraron con el «termómetro de los sentimientos», aplicado a doce candidatos a las elecciones primarias. Los autores utilizaron la técnica de las escalas multidimensionales para identificar las dimensiones que subyacían a la opinión expresada.

Estas técnicas de análisis interpretan las correlaciones entre los estímulos (o elementos, en nuestra terminología), como medidas de *proximidad.* Si, por ejemplo, existe una alta correlación entre las puntuaciones obtenidas por los candidatos Humphrey y Johnson, eso significa que los dos candidatos están cercanos en el espacio mental de los entrevistados: si un sujeto valora positivamente a uno, también valorará positivamente al otro. Este espacio conceptual puede imaginarse como un espacio físico definido por dimensiones (así como en geometría el plano se define con dos dimensiones, el espacio con tres dimensiones, y se pueden imaginar espacios de n dimensiones).

Kruskal y Wish [1978, 7] ofrecen un interesante ejemplo de ello. Imaginemos que tenemos un mapa geográfico y que nos piden que construyamos una tabla con las distancias en kilómetros entre las ciudades. La tarea es fácil de realizar midiendo con una regla las distancias sobre el mapa y convirtiéndolas luego en kilómetros. Supongamos que estamos ante el problema contrario: tenemos la tabla con las distancias en kilómetros y queremos reconstruir las posiciones de las ciudades sobre el mapa. Esta tarea es más difícil que la anterior, pero existen procedimientos geométricos para realizarla.

El problema de la técnica de las escalas multidimensionales es similar al proceso de construcción del mapa a partir de las distancias. A partir de las proximidades y distancias entre determinados objetos (en nuestro ejemplo, los candidatos) debemos reconstruir el espacio conceptual en el que se sitúan estos objetos (en la mente de los entrevistados, si se trata de datos de entrevistas).

En primer lugar, se trata de comprender si existe un espacio común a todos los sujetos que han respondido; a continuación, hay que identificar cuántas dimensiones tiene este espacio; por último, es necesario dar un nombre a estas dimensiones (en el caso de las ciudades la tarea era más sencilla, pues se sabía desde el principio que las dimensiones eran dos, y se conocían sus nombres: latitud y longitud).

En la investigación antes citada sobre los candidatos a las presidenciales estado-

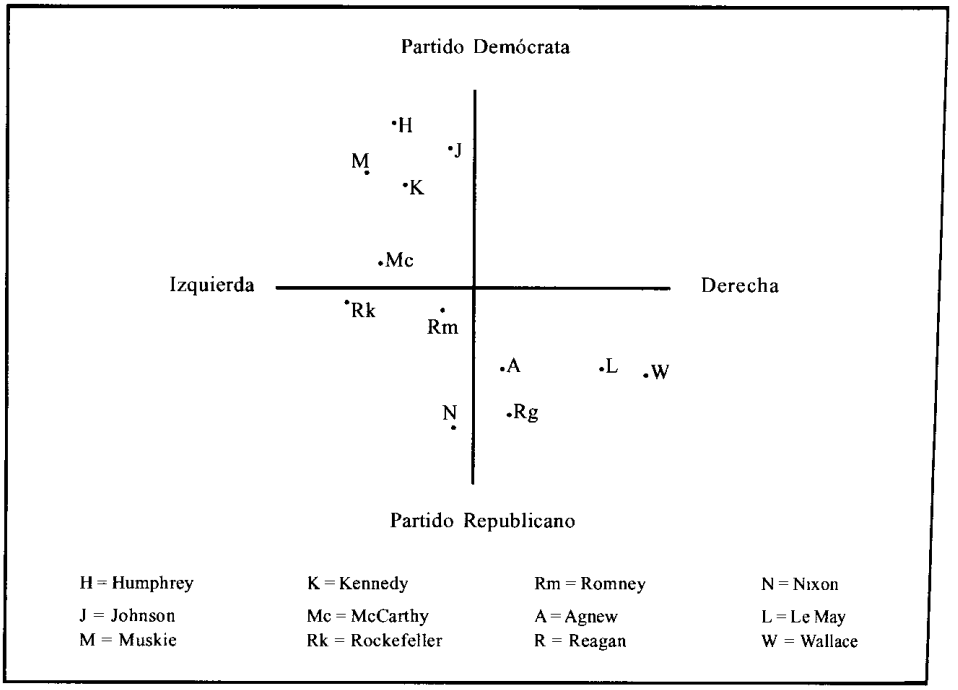


Figura 6.5. Situación de los candidatos a las elecciones presidenciales de EEUU en 1968 en el espacio definido por las dimensiones «identificación con el partido» e «ideología» [Weisberg y Rusk, 1970].

unidenses, se encontraron dos dimensiones (la figura 6.5 muestra la representación de los doce candidatos en estas dimensiones). Los autores interpretaron estas dimensiones como el eje de identificación con el partido (republicano / demócrata) y el eje de ideología (izquierda / derecha). En otras palabras, al expresar sus preferencias los entrevistados aplicaron de forma más o menos inconsciente dos criterios de evaluación: valoraron a cada uno de los candidatos en función de su pertenencia al partido y su posición progresista-conservadora. Por ejemplo, Humphrey se situó cerca del centro en la dimensión de ideología, y fue identificado claramente con el partido demócrata. Rockefeller, en cambio, se situó en una posición equidistante entre el partido republicano y el demócrata (en realidad era republicano, pero con tendencias liberales), pero con un claro posicionamiento ideológico de izquierdas.

No profundizaremos en las escalas multidimensionales, dado que son técnicas muy complejas desde el punto de vista conceptual y matemático y su estudio supera los objetivos de este manual.

7.8. SÍNTESIS DEL CAPÍTULO 6

1. Con el término «técnica de las escalas» designamos una serie de procedimientos elaborados por los investigadores sociales para operacionalizar conceptos complejos. Una escala es un conjunto coherente de elementos que se consideran indicadores de un concepto más general. El *elemento* es una afirmación, pregunta, comportamiento, respuesta a un test o atributo; la *escala* es el conjunto de una serie de elementos. En sociología y en psicología social la aplicación más común de las escalas son las denominadas *escalas de actitudes.* La técnica de las escalas genera variables «cuasicardinales», que presentan propiedades muy similares a las de las variables ordinales.
2. En una pregunta cerrada con respuestas ordenadas las categorías de respuesta pueden ser *semánticamente autónomas,* tener *autonomía semántica parcial,* o ser del tipo denominado *escalas de intervalos* de autoposicionamiento. En los dos primeros casos los resultados son variables ordinales, en el tercero, una variable cuasicardinal.
3. El concepto *escala de Likert* abarca una amplia variedad de escalas, que también se conocen con el nombre de escalas *aditivas,* ya que la puntuación de la escala se deriva de la suma de las puntuaciones de los elementos que la componen. La construcción de una escala de Likert consta de cuatro fases: formulación de las preguntas, aplicación a una muestra de sujetos, análisis de los elementos y del grado de coherencia interna, y control de la validez y el carácter unidimensional de la misma.
4. El *escalograma de Guttman* es una escala *acumulativa,* porque está constituida por elementos de dificultad creciente, y se supone que el sujeto que da una respuesta afirmativa a una determinada pregunta también ha dado una respuesta afirmativa a todas las preguntas que preceden a aquélla en grado de dificultad. Esta característica de escalabilidad se mide con el coeficiente de reproducibilidad.

5. El *diferencial semántico* de Osgood fue creado para estudiar el significado de los conceptos, no a través de la descripción subjetiva de los mismos emitida por los entrevistados, sino mediante las asociaciones entre estos conceptos y otros presentados de forma estandarizada a los sujetos. Con esta técnica, Osgood identificó tres dimensiones fundamentales que subyacen a todas las opiniones de los sujetos sobre los distintos conceptos: *valoración, potencia* y *actividad.*
6. El *test sociométrico,* creado por Jacob Moreno, tiene por objetivo identificar la red de relaciones interpersonales existentes entre los individuos de un grupo. En su versión básica el test sociométrico consiste en un cuestionario que contiene unas cuantas preguntas para plantear a los sujetos del grupo sobre las preferencias y rechazos que sienten hacia los demás componentes del grupo.

6.9. LECTURAS COMPLEMENTARIAS

Una recopilación interesante de las aportaciones clásicas de la tradición empírica estadounidense a las técnicas de medición y las escalas, es la obra clásica de G. M. Maranell *Scaling. A Sourcebook for Behavioral Scientists* (Aldine, 1974, 436 pp.). En D. Krebs P. y Schmidt (eds.) *New Directions in Attitude Measurement* (de Gruyter, 1993, 378 pp.) puede encontrarse una versión actualizada con aportaciones más recientes.

La obra de R. F. DeVellis, *Scale Development: Theory and Applications* (Sage, 1991, 121 pp.) contiene un texto introductorio sobre la medición, que también puede servir como guía práctica para la construcción de las escalas más sencillas.

Una guía muy útil para la construcción de escalas de actitud se encuentra en las recopilaciones elaboradas por J. P. Robinson, P. R. Shaver y L. S. Wrightsman (eds.) *Measures of Personality and Social Psychological Attitudes* (Academic Press, 1991, 753 pp.) y *Measures of Political Attitudes* (Academic Press, 1998, 801 pp.), que contienen una completa exposición sobre los instrumentos disponibles en la actualidad.

CAPÍTULO 7

LAS FUENTES ESTADÍSTICAS OFICIALES

UNA INVESTIGACIÓN BASADA EN ESTADÍSTICAS OFICIALES 248

LAS ESTADÍSTICAS OFICIALES 251

7.2.1. La producción 252
7.2.2. La unidad de análisis 254
7.2.3. El contenido 256
7.2.4. El tamaño del registro 257

LAS ESTADÍSTICAS OFICIALES EN ESPAÑA Y EN OTROS PAÍSES 258

7.3.1. Agencias oficiales 258
7.3.2. Distribución a través de organismos oficiales 261
7.3.3. Metadatos 263
7.3.4. Unidades geográficas 263
7.3.5. Normas y clasificaciones 264
7.3.6. Otros canales de distribución 265
7.3.7. Datos agregados y microdatos 266

EL ESTUDIO DE LAS ESTRUCTURAS SOCIALES 267

SÍNTESIS DEL CAPÍTULO 7 269

LECTURAS COMPLEMENTARIAS 270

Las sociedades actuales producen una gran cantidad de datos sobre sí mismas, que son el resultado de iniciativas emprendidas expresamente por las administraciones para conocer la sociedad (el ejemplo más típico de ello son los censos), o de procedimientos de la burocracia pública (por ejemplo, estadísticas demográficas que proceden del registro de los nacimientos, fallecimientos, cambios de domicilio, etc.). Estos datos constituyen una inestimable fuente de información sobre la sociedad y pueden servir como material empírico para la investigación social. En este capítulo estudiaremos estos datos, dedicando especial atención a la estructura y disponibilidad de las estadísticas oficiales en España y en el mundo anglosajón.

7.1. UNA INVESTIGACIÓN BASADAA EN ESTADÍSTICAS OFICIALES

El gran sociólogo francés Émile Durkheim, en lo que se puede considerar la primera investigación en sociología y que sigue siendo un brillante ejemplo de combinación de preguntas teóricas y evidencia empírica, estudió un fenómeno muy discutido en su época, el suicidio. Durkheim se propuso aplicar técnicas de análisis cuantitativas para el estudio de un asunto sobre el que hasta entonces sólo se habían realizado especulaciones filosóficas, y lo hizo utilizando como material empírico las estadísticas oficiales.

A continuación veremos el proceso que siguió Durkheim para el estudio de uno de los aspectos del problema, que él denominó «suicidio egoísta». En primer lugar, Durkheim se pregunta qué influencia tienen las distintas confesiones religiosas en el suicidio, y para responder a esta pregunta comienza «observando el mapa de los suicidios en Europa» [Durkheim 1897; trad. esp. 1982, 141]. Este mapa muestra que el suicidio está mucho más difundido en los países protestantes que en los católicos[1]. Ante la posibilidad de que esta relación sea «espuria» (como diríamos hoy día), es decir, que la diferencia clave entre los países católicos y los protestantes[2] no sea el factor religioso, sino otros factores (por ejemplo, culturales o de desarrollo económico), «es necesario comparar las dos religiones en el contexto de una misma sociedad», según Durkheim. Para ello, el autor analiza las tasas de suicidio en las distintas provincias del reino de Baviera, que presentan grandes diferencias en cuanto al número de católicos, mientras que al ser provincias de un mismo Estado, la variable «sociedad» se mantiene constante en todas ellas. Durkheim constata que la tasa de suicidios vuelve a ser «directamente proporcional al número de protestantes e inversamente proporcional al número de católicos» [*ibí-*

[1] En los países católicos (como España, Portugal o Italia) las estadísticas oficiales de la época reflejaban una media de 58 suicidios por cada millón de habitantes, mientras que en los países protestantes (como Prusia, Sajonia o Dinamarca) la cifra era de 190.

[2] En palabras de Durkheim: «La civilización de España y Portugal están muy por debajo de la civilización de Alemania, y esta inferioridad podría ser la causa de lo que acabamos de constatar en la evolución del suicidio. Si queremos eludir esta posible fuente de error [...], es preciso comparar ambas religiones en el seno de una misma sociedad» [1897, trad. esp. 1982, 142].

Tabla 7.1. Media de suicidios por millón de habitantes en las provincias bávaras (1867-75)

Provincias con minoría católica (menos del 50%)		Provincias con mayoría católica (del 50 al 90%)		Provincias totalmente católicas (más del 90%)	
Palatinado del Rin	167	Baja Franconia	157	Alto Palatinado	64
Franconia Central	207	Suavia	118	Alta Baviera	114
Alta Franconia	204			Baja Baviera	49
Media	192	*Media*	135	*Media*	75

FUENTE: Durkheim [1897, trad. esp., 1982: 142].

dem, 142] (ver tabla 7.1).

Según Durkheim, Suiza también constituye un caso interesante para este tipo de análisis, pues dado que «se encuentran en ella poblaciones francesas y alemanas, se puede observar por separado la influencia del culto sobre los dos grupos, y esta influencia es la misma en uno u otro». Durkheim advierte que «en los cantones católicos se producen cuatro o cinco veces menos suicidios que en los cantones protestantes, sea cual sea la lengua del cantón», por lo que se puede afirmar que «la influencia del culto es tan fuerte que supera a todas las demás» [*ibídem,* 144]. Y esta influencia debería atribuirse al hecho de que la religión protestante fomenta un mayor grado de individualismo: «la única diferencia fundamental entre el catolicismo y el protestantismo consiste en que el segundo deja más libertad de interpretación de la religión al creyente [...]. El protestante determina en mayor medida sus creencias [...], [mientras que] el católico recibe la fe ya hecha, y no tiene ningún espacio para la interpretación personal» [*ibídem,* 148-149].

Una vez que ha demostrado con diversos datos la influencia del factor religioso en la tasa de suicidios de una determinada sociedad, Durkheim pasa a analizar otros factores, en particular el factor familiar y el factor político. «Si sólo tenemos en cuenta cifras absolutas, parece que los solteros se suicidan menos que los casados, hasta tal punto que algunos autores afirman que el matrimonio y la vida familiar hacen aumentar las probabilidades de suicidio» [*ibídem,* 149]. Pero esta interpretación es errónea, porque no tiene en cuenta ni el número absoluto de solteros y casados ni el hecho de que entre los solteros se cuentan también los niños y los jóvenes, entre los cuales las cifras de suicidio son muy bajas.

Durkheim afirma que si consideramos sólo a las personas mayores de 16 años y si además tenemos en cuenta la edad media de los solteros y los casados [3], el hecho de estar casado reduce el peligro de suicidio aproximadamente a la mitad, reducción que depende también del género y la edad. Como se puede apreciar en la tabla 7.2,

[3] Durkheim constata que la edad influye en la tasa de suicidios (que aumenta progresivamente a medida que aumenta la edad media de las clases de población consideradas), y la edad media de los solteros es inferior a la edad media de los casados.

Tabla 7.2. Media de suicidios anuales por millón de habitantes por grupos de edad, sexo y estado civil (Francia 1889-91)

	Solteros (a)	Casados (b)	Coeficiente de preservación (a/b)	Solteras (a)	Casadas (b)	Coeficiente de preservación (a/b)
20-25	237	97	2,40	106	53	2,00
25-30	394	122	3,20	151	68	2,22
30-40	627	226	2,77	126	82	1,53
60-70	1.768	635	2,78	204	151	1,35
70-80	1.938	704	2,81	206	209	0,98

FUENTE: Durkheim [1897, trad. esp., 1982: 175].

el «coeficiente de preservación» del matrimonio, que Durkheim calcula estableciendo la relación entre la tasa de suicidios de los solteros y la de los casados, es siempre superior a uno[4], lo que significa que la tendencia al suicidio es menor entre los casados.

Durkheim se pregunta si esta protección contra el suicidio que ofrece la familia se deriva de la vida en pareja o de la presencia de los hijos. Para responder a esta pregunta compara, manteniendo constantes otras condiciones como el género y la edad, las tasas de suicidio de personas casadas con y sin hijos, y constata que «la inmunidad [al suicidio] que presentan los casados en general no se debe [...] a la sociedad conyugal, sino a la sociedad familiar» [*ibídem,* 189]; además «la familia, es decir, el grupo completo formado por los padres y los hijos, es el factor esencial de la inmunidad de los casados» [*ibídem,* 200].

Por último, Durkheim estudia la influencia de lo que él llama la «sociedad política», y también en este aspecto su estudio produce resultados de gran interés. Analizando las variaciones en la tasa de suicidios en diversas naciones y en momentos históricos diferentes, constata que en todos los periodos de inestabilidad política, como revoluciones, guerras, disturbios sociales, o simples crisis electorales, ha disminuido la tasa de suicidios en la sociedad. «La única explicación posible para estos hechos es que las grandes conmociones sociales, como las guerras, avivan los sentimientos colectivos [...], y provocan [...] una mayor integración de los individuos» [*ibídem,* 213].

Durkheim concluye esta parte de su investigación con las siguientes consideraciones:

> Por lo tanto, hemos establecido las tres afirmaciones siguientes: *a*) el suicidio varía en sentido inverso al grado de integración de la sociedad religiosa; *b*) el

[4] Si el coeficiente de preservación es igual a uno, la tasa de suicidios entre las dos categorías es la misma; si el coeficiente de preservación es igual a 2, significa que la tasa de suicidios en los solteros es el doble que en los casados.

Tabla 7.3. Número de suicidios por país y año

	1865	1866	1867		1869	1870	1871	1872
Italia	678	588	657	Prusia	3.186	2.963	2.732	2.950
Austria	1.464	1.265	1.407	Suavia	710	657	653	687
				Francia	*5.114*	*4.157*	*4.490*	*5.275*

FUENTE: Durkheim [1897, trad. esp., 1982: 209-210].

> suicidio varía en sentido inverso al grado de integración de la sociedad familiar; y, *c*) el suicidio varía en sentido inverso al grado de integración de la sociedad política. Llegamos así a la siguiente conclusión general: el suicidio varía en sentido inverso al grado de integración de los grupos sociales a los que pertenece el individuo.
>
> [*Ibidem,* 209-210].

7.2. LAS ESTADÍSTICAS OFICIALES

Durkheim analizó las fuentes estadísticas disponibles en su época, publicadas por los organismos gubernamentales existentes entonces. Como hemos visto en el apartado anterior, las fuentes estadísticas oficiales pueden constituir una base empírica muy importante y generar resultados muy convincentes.

Pero, ¿qué entendemos por «estadísticas oficiales»? La palabra «estadística» deriva de la palabra «estado», y su significado original es «ciencia que describe los estados». El adjetivo «estadística» para referirse a esta ciencia fue introducido por

Resumen 7.1. LAS ESTADÍSTICAS OFICIALES EN LA INVESTIGACIÓN SOCIAL

La sociología empírica nació en Europa precisamente a partir de este tipo de datos. La investigación de Durkheim es sólo entre muchos otros. Podríamos citar los trabajos de sociología política de André Siegfried (*Tableau politique de la France del l'Ouest,* 1913) y de Herbert Tingsten (*Political Behaviour,* 1937), que, aunque disponían de un número limitado de variables, realizaron estudios histórico-sociales admirables analizando variaciones territoriales y temporales. No sería hasta finales de los años treinta, con el desarrollo de la investigación social en Estados Unidos, que entonces no tenía bases estadísticas oficiales, cuando comenzó a utilizarse la encuesta por muestreo, que en la posguerra se convirtió en el modelo dominante.

el italiano Gerolamo Ghilini en 1589. Posteriormente se convirtió en el sustantivo «estadística», empleado para designar la disciplina que tenía por objeto la descripción cuantitativa de las características más importantes de un país.

Con el paso del tiempo y con el desarrollo de los aspectos matemáticos de la disciplina, y en particular con la introducción del cálculo de probabilidades, que nació como rama autónoma de las matemáticas en la primera mitad del siglo XVII, el término «estadística» adquirió el significado más general de *método científico para el análisis cuantitativo de los fenómenos colectivos.*

Hoy en día podemos distinguir dos significados en el término «estadística»: el de «ciencia que estudia los fenómenos colectivos con métodos matemáticos basados en el cálculo de probabilidades»; y el de «recopilación de datos con el objetivo de dar una visión de conjunto, aunque sólo sea orientativa, de determinados hechos o fenómenos» [Devoto y Oli, 1995]. Por lo tanto, podemos distinguir entre la estadística como «ciencia» y la estadística como «dato».

En este capítulo, cuando hablemos de estadística(s) nos referiremos a la segunda acepción del término. Las estadísticas demográficas son las tablas que reproducen la totalidad de la población, los nacimientos, las muertes, los matrimonios, las migraciones, etc. ordenados por territorio y año. Las «estadísticas judiciales» son las tablas relativas a delitos, procesos, detenidos, y en general a la actividad de las instituciones judiciales. Y así sucesivamente, para las estadísticas sanitarias, de empleo, de enseñanza, etc.

Estos datos se distinguen del resto que suele utilizar el investigador social, y, en particular, del dato que proviene de la encuesta por muestreo, en cuatro aspectos: la producción, la unidad de análisis, el contenido y el tamaño del registro.

7.2.1. La producción

En el título de este capítulo el atributo «oficiales» (que se podría sustituir por «públicas») indica que estas recopilaciones de datos las realiza, casi en su totalidad, la administración pública. El origen de los censos es muy antiguo: se han encontrado indicios de ellos incluso en civilizaciones primitivas. En general podemos decir que la necesidad de «contar» las poblaciones, las familias, los bienes, surgió con las primeras formas de organización del Estado, y se impuso sobre todo con el nacimiento de los grandes imperios de la antigüedad. Los restos arqueológicos sugieren que ya existían estas formas de registro estadístico en la antigua civilización de los Sumerios (IV-II milenio a. C.), los Asirios, los Babilonios, en el antiguo Egipto, el imperio chino, así como en la Grecia y Roma antiguas (el término «censo» proviene del latín).

Con respecto a las estadísticas que produce la administración pública, conviene distinguir entre los datos generados por la actividad administrativa normal, y los datos recopilados expresamente con una finalidad determinada.

Empecemos por los primeros. En toda sociedad con una estructura burocrática mínima tienen lugar una serie de *actos administrativos* que el individuo o las instituciones realizan con fines burocráticos. Si estos actos se recogen y se organizan adecuadamente, producen los datos estadísticos. Pensemos, por ejemplo, en las es-

tadísticas demográficas: cada vez que se producen nacimientos, muertes, matrimonios, divorcios, cambios de domicilio, los ciudadanos deben cumplimentar una serie de impresos, que desde el momento en que se registran ya son parte de la información necesaria para producir el dato estadístico.

Veamos otros ejemplos. Las estadísticas sobre delitos realizadas a partir del registro de las denuncias presentadas por los ciudadanos, las estadísticas sobre el comercio exterior a partir de los registros aduaneros, las estadísticas sobre la enseñanza a partir de los datos de alumnos inscritos en los centros escolares y los que obtienen el correspondiente título, las estadísticas sanitarias (sobre abortos, enfermedades infecciosas, causas de muerte, hospitalización), a partir de los registros realizados por los médicos y los distintos centros sanitarios, las estadísticas de seguridad social y asistenciales a partir del suministro de dinero para pensiones y otras prestaciones sociales, las estadísticas sobre el ahorro familiar a partir de la información sobre depósitos bancarios, y así sucesivamente.

El uso del dato administrativo para elaborar estadísticas que nos permitan conocer mejor la sociedad no es generalizable a todas las sociedades actuales. Es una tradición típica de los Estados europeos, que gracias a sus aparatos de gobierno muy centralizados, desde el siglo XIX empezaron a crear registros de datos aprovechando su vasta organización administrativa (ayuntamientos, municipios, comarcas, etc.). Fuera de Europa, esta tradición es desconocida, por ejemplo, en Estados Unidos, tanto por la ausencia de un Estado centralizado, como por su fuerte tradición individualista, que desconfía de cualquier forma de intromisión del Estado en la vida de los ciudadanos.

Siempre que el dato estadístico sea el resultado de un acto administrativo, podemos hablar de *recogida indirecta*. Otras veces, en cambio, la producción del dato estadístico se realiza por *recogida directa,* en el sentido de que la información se recoge expresamente con la finalidad de conocer un determinado fenómeno social. Éste es el caso, ya antes citado, del *censo,* que organiza el Estado precisamente con la finalidad de conocer las características de la población.

Los censos modernos nacen en el siglo XIX. Los primeros censos en inglés y francés, por ejemplo, datan del año 1801, e inauguran series decenales que se prolongan hasta nuestros días. En España el primer censo se encargó en 1857[5].

Durante mucho tiempo las estadísticas sociales se han realizado con los datos procedentes de actos administrativos y del censo. Recientemente, a los censos se han sumado otras formas de recogida directa de datos. La utilización con fines estadísticos de datos procedentes de registros administrativos ofrece la ventaja de eliminar casi por completo los costes del proceso de recopilación de material empírico, pero presenta el problema de que, al tratarse de datos recogidos para otros fines, podrían resultar inadecuados para el conocimiento completo del fenómeno que se quiere estudiar. Además, existen datos sociales que no se pueden incluir en ningún tipo de acto administrativo, por ejemplo, los nuevos modelos de estructuras

[5] Para una historia de los censos en España, véase el capítulo sobre los censos en el libro de Reher y Valero: *Fuentes de información demográfica en España* (1995).

familiares (parejas de hecho, familias con hijos de matrimonios anteriores, convivencia entre ancianos, etc.), la inmigración ilegal, los delitos no denunciados a las autoridades, etc. Por estos motivos, en los últimos años los institutos nacionales de estadística realizan una serie de *investigaciones por muestreo ad hoc,* sobre sectores específicos, con el fin de estudiar determinados aspectos de la sociedad. En España, por ejemplo, el INE realiza cada año encuestas para contabilizar la producción industrial de las empresas, el comercio minorista, a fin de conocer los niveles salariales en España, el fenómeno de la inmigración, y otros datos relevantes. [6]

Aunque las investigaciones por muestreo *ad hoc* son cada vez más numerosas en todos los institutos nacionales de estadística, se prevé que en el futuro aumente también el uso del material de origen administrativo con fines estadísticos, gracias a las amplias posibilidades que ofrece la informática: a medida que se va informatizando la gestión ordinaria de los diversos sectores de la administración, se crea la base para una conexión cada vez mayor entre los distintos archivos administrativos existentes.

7.2.2. La unidad de análisis

Una diferencia fundamental entre las estadísticas oficiales y la encuesta por muestreo es que para las primeras la *unidad de análisis* no es el individuo, sino una agrupación de individuos en un *territorio*, que puede ser un distrito electoral o censal, un municipio, una provincia, una región, una comarca, un país, etc. (aunque esto está cambiando, como veremos al final de este apartado). Cuando Durkheim compara la tasa de suicidios en las provincias de Baviera (tabla 7.1), utiliza datos que tienen como unidad de análisis la provincia. Cuando analiza la evolución en el curso del tiempo de la tasa de suicidios para observar las variaciones en los años de guerra (tabla 7.3), utiliza datos que tienen como unidad de análisis el país. Así mismo, cuando compara las tasas de suicidios entre casados y solteros por género y edad, sus datos tienen también una unidad de análisis de base territorial (la tabla 7.2, por ejemplo, se refiere a Francia).

Aunque en un principio la información se recoge individuo por individuo (que sería la *unidad de registro*), los datos están disponibles y se pueden analizar sólo de forma colectiva (lo que sería la *unidad de análisis*). Imaginemos una tabla en cuyas filas situamos las regiones de un determinado país, y en las columnas tres estadísticas en forma de porcentajes (o «tasas») para cada una de las regiones: la tasa de divorcios (sobre el total de matrimonios), el porcentaje de votos al partido socialista (sobre el total de votos válidos), y el número de entradas de cine vendidas (sobre la totalidad de la población).

En este caso la unidad de análisis es un territorio, en concreto, la región, y

[6] En la página web del Instituto Nacional de Estadística de España (INE), www.ine.es, se puede consultar el catálogo de todas las encuestas y sus características muestrales en cada año. En el apartado 7.3 ofrecemos información más detallada sobre el INE.

sólo podemos interpretar y analizar los datos dentro de este ámbito o marco de referencia. Por ejemplo, podemos decir que la región A tiene una tasa de divorcios mayor que la región B, que la media regional de entradas de cine vendidas al año es de 3,8 por habitante, que existe una correlación (a escala regional), entre el voto al partido socialista y la tasa de divorcios. Como puede apreciarse, todas estas afirmaciones tienen como marco de referencia la región, y no los individuos.

Estos datos colectivos se denominan datos *agregados,* por contraste con los datos *individuales.* El adjetivo agregado alude a una operación de composición o unión, y eso es exactamente lo que caracteriza a los datos agregados.

El dato agregado procede de una operación de *recuento* realizada sobre los individuos de un colectivo, que genera un total. A continuación se lleva a cabo una normalización basada en la población total del colectivo con el fin de anular las diferencias de dimensiones de los colectivos y poder compararlos entre sí.

En el caso de las estadísticas oficiales, el colectivo (que podría ser cualquier conjunto de individuos, como empresas, sindicatos, centros escolares, etc.) lo constituye un *territorio.* De esta manera contabilizamos el número de divorcios, el número de votos a un determinado partido político, el número de entradas de cine vendidas y, si situamos estos datos en un marco de referencia, que puede variar (en nuestros ejemplos este marco es el total de matrimonios, los votos válidos y la población residente, respectivamente), obtenemos una *relación estadística* que permite hacer comparaciones entre las observaciones agregadas.

Los datos se suelen recopilar de manera individual (cada uno de los cónyuges que se divorcia, cada uno de los electores que vota), pero sólo están disponibles de modo agregado. Cabe preguntarse por los motivos de ello.

En primer lugar, la información puede haberse generado de forma individual y haber sido registrada de forma de agregada, por territorios, por cuestiones de confidencialidad, como sucede con las elecciones, donde el voto es emitido por el individuo pero se registra sólo como total de votos a un cierto partido en un determinado territorio.

Otras veces la información se recoge inicialmente de forma individual, pero con fines meramente administrativos. Por ejemplo, las estadísticas demográficas, donde se cuenta el número de nacimientos, muertes, matrimonios contraídos e inmigrantes en un determinado día y en un municipio concreto, sin registrarse más información sobre los individuos, salvo algunos rasgos personales básicos que se suelen limitar al género, la edad, y el estado civil (como en el estudio del suicidio, cuyos datos se derivaban de un registro de las causas de muerte). Por tanto, el dato se conserva y se hace público sólo de forma agregada con subdivisiones a partir de unas pocas variables (por ejemplo, el número de inmigrantes por género y edad en un determinado año y para un municipio concreto).

En otros casos sucede que, aunque la información ha sido recogida y *registrada* de forma individual con una cierta riqueza de detalles, el dato se publica (y está disponible para la investigación) sólo de forma agregada, por cuestiones de organización. Tomemos como ejemplo el caso del censo. La información la registra un entrevistador que plantea preguntas a una serie de individuos (o familias). En principio esta información no difiere de la recogida mediante el cuestionario de una

encuesta por muestreo. Pero antes de que aparecieran los ordenadores no existía la posibilidad de analizar los impresos censales de forma individual, tratándolos del mismo modo que los cuestionarios de la encuesta por muestreo. De hecho, no era posible almacenar, guardar y analizar millones de fichas individuales, salvo en forma de tablas sintéticas por municipios (u otras unidades territoriales).

Por último, en algunos casos, la información procede de datos administrativos registrados directamente de forma agregada (por ejemplo, las ventas de entradas en las salas de cine).

Todo esto en cuanto a las estadísticas oficiales tradicionales. En la actualidad se están produciendo ciertos cambios. Gracias a los enormes avances de la informática, cada vez es más frecuente que los datos procedentes de fuentes estadísticas oficiales estén disponibles no sólo en forma agregada, sino también en forma de registros individuales. Aún hoy la mayoría de las fuentes estadísticas oficiales se encuentra en formato agregado por territorios. Además, aunque los datos se ponen a disposición del investigador en forma de archivos de base individual, esto no excluye que se puedan utilizar en forma agregada y por territorios. La ventaja de los datos agregados es que se pueden consultar de forma inmediata, mientras que los datos individuales requieren una serie de operaciones para las que se precisa ciertos conocimientos técnicos.

7.2.3. El contenido

La información que se deriva de un acto administrativo no nace de una pregunta ni de la observación, sino del registro de un hecho (un nacimiento, una adquisición, un cambio de domicilio, un ingreso hospitalario, la inscripción en un centro escolar, etc.). Nos encontramos, por tanto, en el terreno lo que denominamos «datos factuales», de los cuales se excluyen todo lo que sean opiniones, actitudes o motivaciones.

Esta connotación de la información contenida en las fuentes estadísticas se mantiene incluso cuando éstas no proceden de fuentes administrativas, sino de encuestas por muestreo realizadas *ad hoc*, con objetivos concretos. Normalmente el censo se limita a preguntas sobre hechos. Y, por lo general, las investigaciones específicas que llevan a cabo los institutos nacionales de estadística sobre muestras y problemáticas sociales concretas, tampoco estudian las motivaciones.

Esta autolimitación se debe sobre todo a las características originarias de las estadísticas oficiales, que, al proceder de registros administrativos, se refieren exclusivamente a acciones, comportamientos y hechos. También se debe a que, al tratarse de material gestionado por el Estado (o en cualquier caso por entidades públicas), su gestor no puede permitirse invadir la esfera de lo privado con la misma desenvoltura que puede hacerlo una entidad privada. Por último, la referida autolimitación se deriva también de la magnitud de los datos tratados: incluso en la investigación por muestreo, las encuestas estadísticas oficiales se realizan sobre muestras de grandes dimensiones, con un número considerable de encuestadores, lo que implica que se concede más importancia a la representatividad de la distribución de las variables (incluso a escala subregional), que al conocimiento profundo de las motivaciones.

7.2.4. El tamaño del registro

Otra característica del dato estadístico tradicional es que se registra sobre toda la población. Podemos distinguir entre *el registro exhaustivo* (o *total*) y el *registro muestral* (o *parcial*). El dato estadístico tradicional de origen administrativo es un dato exhaustivo, ya que deriva de un registro realizado sobre toda la población: todos los nacimientos, las muertes, los ingresos hospitalarios, los procedimientos judiciales, etc. de un municipio en un determinado año (además, la técnica del muestreo es bastante reciente, pues no empezó a consolidarse hasta los años treinta o cuarenta, mientras que los registros estadísticos se hacían ya desde principios del siglo XIX).

Los propios censos se han venido realizando sobre la totalidad de la población, y esto sigue siendo así en la actualidad. En nueve de los doce países que formaban entonces la Unión Europea, las operaciones censales de 1990-91 se realizaron sobre la totalidad de la población. En dos países, Dinamarca y Holanda, las operaciones tradicionales se sustituyeron por la conexión informática de *archivos* administrativos; sólo en el caso de Alemania[7] el registro censal se realizó sobre una muestra (que representaba al 1 % de la población).

Las ventajas de los registros por muestreo en comparación con los registros totales son bastante obvias. El registro por muestreo permite: *a*) *reducir los costes; b*) *reducir el tiempo* de recopilación y elaboración de los datos (que son muy numerosos en los registros totales); *c*) *reducir la carga organizativa,* dado que no es necesario reclutar, formar y dirigir a legiones de encuestadores, necesarios para el registro total; y *d*) *posibilidad de profundizar*: como el número de encuestadores es menor, es posible formarlos adecuadamente para que realicen registros más complejos, y como la organización es menos compleja, es posible concentrar los recursos en el control de la calidad del registro.

Por otra parte, también es obvio que sólo es posible analizar grados menores de disgregación territorial mediante un registro total. Aunque para el estudio de la población de todo un país o de grandes agregados territoriales como las regiones puede ser suficiente un registro por muestreo, para conocer las características de la población de cada uno de los municipios es necesario un registro total. Por este motivo, el censo, precisamente por su condición de «inventario de los recursos humanos y materiales» de un país, se realiza sobre toda la población; la finalidad del censo es proporcionar una fotografía *del* estado del país cada diez años, mientras que las denominadas *informaciones de flujo* se obtienen con las estadísticas corrientes.

Las investigaciones *ad hoc* más recientes, realizadas por varios institutos nacionales de estadística sobre aspectos concretos y hasta el momento poco explorados de la sociedad, se suelen realizar sobre muestras. Sin embargo, se trata de muestras de grandes dimensiones (que por lo general contienen más de 10000 casos). Recorde-

[7] Hay que recordar que en el año 1987 Alemania había realizado el censo que tenía que haber realizado en 1983.

mos que una encuesta por muestreo normal realizada por una entidad privada sobre muestras de la población de un país suele incluir entre 1000 y 2000 observaciones.

El origen de esta diferencia se encuentra en el objetivo específico de la fuente estadística oficial, que no es sólo estudiar un hecho social, comprobar empíricamente una teoría y encontrar las posibles causas de un determinado fenómeno. El objetivo principal de las estadísticas oficiales es de carácter descriptivo. Además, estas estadísticas son muy sensibles a los aspectos geográfico-territorial, puesto que buscan obtener estimaciones fiables de los fenómenos estudiados incluso a escala subnacional. En España esta escala subnacional suele ser las comunidades autónomas. Con un total de 17 y partiendo de unos 1000 casos por comunidad, incluso para las más pequeñas, la muestra nacional puede alcanzar fácilmente dimensiones de unas 20000 observaciones.

7.3. LAS ESTADÍSTICAS OFICIALES EN ESPAÑA Y EN OTROS PAÍSES

El término «estadísticas oficiales» puede hacer pensar en un sistema de datos homogéneo y unificado. Sin embargo, las operaciones de producción, revisión, mantenimiento y distribución de los datos recopilados por organismos oficiales se realizan de manera muy variable y poco precisa. Algunas de las características que varían entre los distintos grupos de estadísticas son: la naturaleza de los datos, el grado de centralización o dispersión en la recopilación de los datos, las agencias implicadas en la producción de estadísticas, el nivel de profundidad y la cobertura geográfica, o la forma de distribución.

Las estadísticas pueden ser recopiladas y distribuidas por unidades de gobierno locales a partir de registros administrativos. También pueden recopilarse a escala local y publicarse desde instancias administrativas o regionales superiores. Por último, hay estadísticas que proceden de encuestas y son el resultado de un trabajo coordinado a nivel nacional para el proceso de recopilación de datos; posteriormente, una agencia central reúne estas estadísticas y las distribuye de forma sistemática.

7.3.1. Agencias oficiales

Las agencias u organismos públicos pueden cumplir diversas funciones en la organización del sistema estadístico de un país. En España, el Instituto Nacional de Estadística (INE) es el organismo encargado de coordinar todos los servicios estadísticos de la administración. El INE es un organismo autónomo creado en 1945 y adscrito al Ministerio de Economía. La Ley de la Función Estadística Pública (del 9 de mayo de 1989) asigna al INE un papel destacado en la actividad estadística pública, encomendándole expresamente la realización de operaciones estadísticas de gran envergadura, como, por ejemplo, censos demográficos y económicos, cuentas nacionales, estadísticas demográficas y sociales, indicadores económicos y sociales, coordinación y mantenimiento de los directorios de empresas, elaboración del censo electoral, etc.

La citada ley atribuye al INE las siguientes funciones: llevar a cabo el plan estadístico nacional (PEN) con la colaboración de los departamentos ministeriales y del Banco de España; proponer normas comunes sobre conceptos, unidades estadísticas, clasificaciones y códigos; encargarse de las relaciones en materia estadística con los organismos internacionales especializados, en particular, la oficina de estadística de la Unión Europea (Eurostat); y coordinar el trabajo con las oficinas estadísticas de las comunidades autónomas. De manera periódica, el INE publica los resultados más recientes de sus actividades, y edita también numerosas publicaciones de sectores específicos y de ámbito más general.

La mayoría de los datos del INE están disponibles en soporte magnético y es posible acceder a gran parte de ellos por Internet, a través del sistema de almacenaje de datos estadísticos denominado *Inebase*. Este banco de datos contiene información estadística procedente del INE y de todo el sistema estadístico nacional. Además, el banco de datos *Tempus,* que es el sistema clásico del INE para la difusión de series de datos temporales y que se inició en 1987, también está integrado en el Inebase. Algunas de las series de Tempus son de acceso restringido porque el modo de desagregación de los datos es reservado, ya sea por motivos de confidencialidad, o porque la calidad de los datos no siempre es alta (lo que suele deberse a un error de muestreo elevado).

Respecto a las series temporales disponibles, cabe mencionar, entre otras, las siguientes: contabilidad nacional de España, índices de precios al consumo, estadísticas de causas de mortalidad, producción editorial, migraciones, movimiento natural de la población, encuesta de salarios o de población activa, etc. La necesidad de desarrollar un banco de datos cronológico en el INE estuvo ligada desde el principio al mantenimiento de un sistema de indicadores de coyuntura a partir de los cuales se hacen estudios sobre la evolución y la situación económica, y a la elaboración de la publicación del boletín trimestral de coyuntura.

Además del INE, dentro de la administración general del Estado existen otros ministerios u organismos oficiales que producen datos estadísticos, muchos de ellos disponibles a través de Internet, y publicaciones. En la página web del INE existe una amplia selección de enlaces a servidores estadísticos oficiales. Para citar algunos, el Banco de España (www.bde.es) produce estadísticas con información financiera y monetaria y, en menor medida, las estadísticas del sector real destinadas a alimentar su boletín estadístico, su publicación oficial y otras publicaciones. Además, a través de Internet se pueden consultar series cronológicas de indicadores económicos generales, estado de la economía española, contabilidad nacional, convergencia real España-UE, zona euro, economía internacional, etc.

También son interesantes las estadísticas sobre educación y cultura ofrecidas por el Ministerio de Educación y Ciencia (www.mec.es). Por un lado, los datos sobre la cultura dan una idea del desarrollo cultural de España en su conjunto y de las distintas regiones. Por otro lado, las estadísticas sobre educación de este ministerio aportan información sobre el gasto público en educación, estadísticas sobre la enseñanza universitaria y no universitaria, el curso escolar, la sociedad de la información o las becas. La información se suele presentar desglosada por comunidades autónomas. Este ministerio dispone también de numerosas publicaciones, por ejem-

plo, la serie anual titulada: *Estadística de la enseñanza en España.*

El Ministerio de Trabajo y Asuntos Sociales (www.mtas.es) ofrece datos diversos en materia laboral, como estadísticas de accidentes de trabajo, convenios colectivos, huelgas, regulaciones de empleo o permisos de trabajo a extranjeros. Su banco de datos contiene información sobre éstas y otras series de datos laborales, y una de sus funciones principales es alimentar el boletín de estadísticas laborales con información sobre el mercado de trabajo, las relaciones y las condiciones laborales, o el nivel de protección social. La página de inicio de este ministerio está conectada con el Instituto de la Mujer, el Instituto de la Juventud, el INEM o el Inserso, que también disponen de estadísticas en sus respectivas áreas.

A escala regional, desde la aprobación de la Ley de la Función Estadística Pública, la mayoría de las comunidades autónomas (CCAA) han asumido competencias en materia de estadística, creando sus respectivos institutos u oficinas de estadística. Éstos suelen estar conectados con los servicios estadísticos de las consejerías, sobre todo las de economía y hacienda. La mayoría contienen información detallada sobre presupuestos, contabilidad autonómica, datos periódicos sobre coyuntura económica, y sobre la realidad social y demográfica de la comunidad autónoma correspondiente. Los datos más relevantes a este respecto se presentan en los denominados *Índices anuarios estadísticos,* que se presentan ordenados por grandes temas y resultan muy útiles cuando se desea obtener información a escala local, ya que la mayoría publican datos por municipios. Toda la información se encuentra disponible en Internet[8].

A escala local, las *corporaciones locales y ayuntamientos* de las principales ciudades españolas también suelen disponer de departamentos de estadística. La mayoría de estos organismos editan publicaciones con datos agregados y disponen de estadísticas en soporte magnético, así como página web. Por ejemplo, en las páginas web de las grandes ciudades españolas existe una conexión con los departamentos, servicios u oficinas de estadística de los ayuntamientos[9].

A escala comunitaria, la oficina estadística de la Unión Europea, *Eurostat* (www.europa.eu.int/comm/eurostat) ofrece información comparada de los distintos países que la forman. Este organismo recoge datos producidos por las agencias nacionales de los países miembros y los armoniza para facilitar su uso a escala global. Además, Eurostat suministra información estadística al público en general a

---

[8] Algunos ejemplos de este tipo de institutos son: el Instituto de Estadística de Andalucía, el Instituto Canario de Estadística, el Institut d'Estadística de Catalunya, el Instituto de Estadística de la Comunidad de Madrid, el Instituto Galego de Estatística, el Instituto Vasco de Estadística (Euskal Estatistica Erakundea), etc. En la página web del INE (dentro del apartado de enlaces: www.ine.es/serv/estadist.htm#es) se puede encontrar un listado muy completo de todos los institutos de estadística de las CCAA de España, con sus correspondientes páginas web. En ese mismo apartado se encuentra una recopilación detallada de la información estadística en ministerios, organismos e instituciones oficiales, algunos de los cuales se citan en el texto, aunque no de forma exhaustiva.

[9] Véase, por ejemplo, la página web del Ayuntamiento de Barcelona (www.bcn.es/estadistica/catala/index.htm), o de Madrid (www.munimadrid.es/principal/ciudad/informacion.asp), o el servicio de estadística del Ayuntamiento de Sevilla (www.sevilla.org/frames.htm), la oficina de estadística del Ayuntamiento de Valencia (www.ayto-valencia.es/ayuntamiento2/ndayunt.nsf?opendatabase), etc.

través de publicaciones impresas y electrónicas, y ofrece cooperación técnica con países que no forman parte de la Unión Europea. Su banco de datos estructurales y coyunturales es muy extenso y, en algunos casos, contiene datos sobre países no miembros. La temática de las series e indicadores es amplia: economía, industria y energía, población y condiciones sociales, educación, medio ambiente, transporte, investigación y desarrollo, etc. Dispone también de metadatos y un gran número de publicaciones y anuarios estadísticos.

Finalmente, y por lo que se refiere al ámbito latinoamericano, todos los países disponen de agencias oficiales o institutos públicos similares al INE español, que centralizan la producción oficial de estadísticas. Además, existen organismos internacionales que recopilan, almacenan y difunden datos estadísticos de grupos de países latinoamericanos. Cabe destacar el Banco Interamericano de Desarrollo (www.iaddb.org), que se centra sobre todo en datos económicos, financieros y de desarrollo, o la Comisión Económica de las Naciones Unidas para América Latina y el Caribe (CEPAL, www.eclac.cl), que lleva a cabo una importante labor en la armonización y difusión de estadísticas sobre Latinoamérica. Esta institución dispone de bases de datos macroeconómicos o de comercio exterior recopilados a partir de fuentes oficiales de los países latinoamericanos. En este ámbito, publica la *Base regional de datos de coyuntura,* que recoge indicadores socioeconómicos de carácter coyuntural de ocho países de América Latina (Argentina, Bolivia, Brasil, Chile, Colombia, Ecuador, México, Paraguay, Perú, Uruguay y Venezuela). Las series de datos se inician en el año 1990 y tienen frecuencias mensual, trimestral, semestral y anual. Además de estadísticas y proyecciones económicas, este organismo recoge y almacena también datos sociodemográficos sobre aspectos como la evolución demográfica del continente, mujer y desarrollo, políticas sociales o medio ambiente.

7.3.2. Distribución a través de organismos oficiales

Los datos estadísticos oficiales se están haciendo omnipresentes. Su producción se ha generalizado considerablemente. Si en sus orígenes podían ser simples enumeraciones centralizadas, en la actualidad existen programas de cobertura sistemática realizados por una amplia variedad de entes gubernamentales. El incremento del número de estadísticas disponibles y la diversidad de fuentes conllevan una abundancia excesiva de estadísticas que, en ciertas ocasiones, pueden llegar a ser incoherentes. Aunque en muchos países se sigue manteniendo un cierto grado de dispersión en las actividades estadísticas, tiende a utilizarse un criterio único para la distribución de los resultados. En cuanto al material publicado, suele existir un conjunto de libros o publicaciones de referencia que contiene una selección de estadísticas ordenadas por temas.

En España cabe destacar la revista mensual *Fuentes estadísticas,* que empezó a publicarse en diciembre de 1994, y que se encuentra disponible en Internet con formato de revista electrónica (www.fuentesestadisticas.com). Su creación fue fruto de un convenio entre el INE, Eurostat y la Universidad Autónoma de Madrid (UAM), con el objetivo prioritario de dar a conocer las estadísticas oficiales más

recientes que se realizan en España y Europa. Cada número trata sobre un tema monográfico.

Como fuentes de referencia básicas existen los anuarios estadísticos que producen la mayoría de instituciones oficiales. Desde 1858, aunque con algunas interrupciones, el INE ha venido publicando el *Anuario estadístico de España,* que es una obra de información estadística general, disponible en CD-ROM desde 1998. Este anuario recopila datos procedentes de diversas fuentes oficiales, y tiene como objetivo ofrecer una visión cuantitativa de la realidad económica, demográfica y social. Otra publicación anual destacada del INE es *España en cifras.* En ella se resumen los datos más relevantes y actuales de España y sus CCAA, dando una visión general y sintética sobre la realidad social y económica.

El INE también realiza el *Boletín mensual de estadística,* una publicación de síntesis estadística disponible en formato electrónico que presenta amplia información sobre los aspectos demográfico, social y económico en España e incluye descripciones metodológicas como apoyo a los datos numéricos. Este boletín incluye sobre todo datos nacionales, aunque aporta también algunos desgloses regionales y comparaciones con otros países. Se centra en indicadores con una frecuencia de aparición mensual o trimestral, mostrando su evolución en períodos recientes. En el año 2000 empezó a publicarse el boletín de noticias estadísticas *Cifras INE,* también creado por el INE. El contenido de cada número es monográfico, aunque se dedican secciones concretas a recoger las publicaciones aparecidas o los indicadores coyunturales económicos o sociales más recientes. El INE publica también desde 1958 la *Revista estadística española,* orientada a difundir las investigaciones españolas en materia estadística.

Otras fuentes de datos de estadística en España son los anuarios que editan algunos diarios, como *El País,* que se publica desde 1982, y *El Mundo,* desde 1993. En estos anuarios se pueden encontrar estadísticas muy diversas de ámbito tanto nacional como internacional.

En Latinoamérica las agencias estadísticas oficiales de cada país también editan publicaciones y recopilaciones estadísticas. La CEPAL produce un número importante de publicaciones sobre datos comparados de toda la región, algunas de ellas disponibles en formato electrónico. Cabe mencionar el *Anuario estadístico de América Latina y el Caribe,* el *Estudio económico de América Latina y el Caribe,* la revista de la CEPAL o su boletín demográfico. El *Anuario estadístico de la CEPAL* contiene una selección de las principales series estadísticas disponibles sobre la evolución económica y social de los países de la zona, que se actualiza periódicamente.

Afortunadamente para los científicos sociales que usan datos oficiales, cada vez hay más estadísticas disponibles en formato electrónico. Además, los datos de las publicaciones se extraen de series estadísticas que suelen estar disponibles en hojas de cálculo o bases de datos. En muchos casos se han diseñando aplicaciones interactivas para ayudar a los usuarios a seleccionar entre las series disponibles [10].

[10] Un ejemplo de ello es el *Historical Census Data Browser* (http://fisher.lib.virginia.edu/census), que permite seleccionar características de los censos realizados desde 1790 en Estados Unidos y disponer de datos estadísticos en varias formas de agregación.

También existe la posibilidad de realizar algún tipo de análisis univariado o bivariado, como, por ejemplo, tablas de contingencia o cálculo de medias.

7.3.3. Metadatos

Los metadatos, que contienen información sobre los propios datos, son particularmente importantes en un ámbito en el que las estadísticas oficiales disponibles son cada vez más numerosas. En el pasado, cuando cada estadística iba acompañada por un texto explicativo, o dentro de una tabla con notas a pie de página y encabezamientos contextuales y asociaciones, casi toda la información que permitía al investigador entender la estadística se percibía de manera inmediata. Incluso cuando la estadística se presentaba sin metadatos, como era el caso de los archivos de texto, el proceso habitual por el cual los datos se importaban al programa de análisis estadístico era lo suficientemente laborioso como para precisar una lectura detallada del libro de códigos con los metadatos.

La obsesión de los investigadores y centros de investigación por conseguir que los datos estén fácilmente accesibles en Internet ha provocado que diversas organizaciones, tanto gubernamentales como privadas, hayan empezado a difundir estadísticas de todo tipo sin aplicar criterios bien definidos para la ordenación de la información. Ello ha provocado su vez falta de información para interpretar adecuadamente los datos disponibles. Por ejemplo, las series de cifras de desempleo femenino pueden variar en función de un denominador: para algunos años puede ser que se incluya a todas las mujeres mayores de 14 años, y para otros, sólo las mujeres entre 16 y 64 años que están buscando empleo o están trabajando.

Si no disponemos de información sobre el universo del que se parte para construir el indicador, el proceso de recopilación de los datos, la forma de procesar la información, o el tipo de estadística (un recuento, un porcentaje o una media) que representa, podríamos malinterpretar estadísticas aparentemente sencillas. Aunque la información estandarizada sobre las características de los datos aún está en proceso de desarrollo, en las estadísticas más difundidas suele incluirse ya de manera estándar, junto con los datos, la siguiente información: descripción, fuentes, títulos, cobertura, organismo emisor, fecha, sujeto y tipo de datos.

El *Dublin Core Metadata Initiative* (DCMI, www.dublincore.org) y las definiciones de tipos de datos de la *Data Documentation Initiative* (DDI, www.icpsr.umich.edu/DDI) son proyectos que intentan establecer los criterios mínimos para caracterizar los metadatos, así como las etiquetas más adecuadas para los mismos. Esta información debería acompañar a todos los datos electrónicos que están disponibles en Internet. Los investigadores deben ser conscientes de que cuando esta información no está disponible, surgen dudas sobre las estadísticas que se están utilizando.

7.3.4. Unidades geográficas

Los datos estadísticos se refieren a una determinada área geográfica. Ésta puede ser compleja, e incluso variar entre los distintos sistemas de estadísticas. En España el INE creó el *Inventario de operaciones estadísticas* de la Administración General del

Estado (IOE), que es un instrumento fundamental para la coordinación y planificación de los distintos organismos oficiales que emiten estadísticas. Entre las diversas características presentes en este inventario, el IOE establece los distintos rangos de desagregación de los datos estadísticos que se producen en España. Éstos rangos están ordenados de mayor a menor como sigue: nacional, autonómico, provincial, municipal o inferior u otros.

El INE ha reservado la categoría «otros» para aquellas desagregaciones de rango inferior al nacional que no disponen de tablas para los niveles geográficos mencionados, especificando de qué clasificación se trata. Siguiendo la división administrativa oficial, los municipios españoles se dividen en distritos municipales, pero para rangos inferiores, el INE, junto con los ayuntamientos, hace una subdivisión adicional de los distritos en secciones censales. Las secciones se utilizan para todos los trabajos encomendados por el INE que requieran una división inframunicipal, por ejemplo, las secciones electorales. Esto exige, según estipula la Ley Electoral, que cada sección incluya un máximo de 2000 electores y un mínimo de 500. En toda España hay unas 32000 secciones. Por tanto, la sección censal puede considerarse un área geográfica con limitaciones claramente definidas, cuyo tamaño de población viene determinado por las condiciones mencionadas. Los individuos seleccionados (tanto el individuo concreto como el número de individuos escogidos) puede variar con el tiempo y se va actualizando cada vez que se realiza un censo o padrón.

7.3.5. Normas y clasificaciones

En condiciones ideales, las estadísticas oficiales deberían ofrecer una medida coherente y de la máxima calidad posible para cada dato que registran. La existencia de metadatos apropiados y referencias geográficas sirve como base para asegurar que el universo que se está cuantificando y el territorio sobre el cual se están realizando las estadísticas, sean coherentes y explícitos (o al menos que se sepa cuándo no lo están siendo). La coherencia de las categorías que se utilizan para la tabulación de los datos y la documentación de los cambios en las categorías son otra de las características de las estadísticas oficiales.

Aunque, como ya hemos dicho anteriormente, las estadísticas oficiales tienden a reflejar y describir más los hechos que las opiniones, las categorías de interés y la forma en la que se definen estos hechos vienen determinadas socialmente. Por ejemplo, las estadísticas sobre el suicidio pueden parecer claras (¿está muerta la persona?, ¿provocó ella misma su propia muerte?, ¿fue un acto intencionado?), pero la atribución del suicidio y su posterior registro por parte de los funcionarios pueden estar afectadas de cierta ambigüedad derivada de la valoración de los jueces de instrucción, los médicos, los miembros de la familia o la policía.

Para categorías más abstractas, como la ocupación, el desempleo, la clase social, las influencias sociales en los recuentos de los hechos son evidentes. Las agencias oficiales establecen normas para la clasificación de estos hechos sociales, en condiciones ideales a escala nacional, pero estas normas pueden variar. En España, desde los años noventa, el INE ha elaborado diversas clasificaciones estadísticas en conso-

nancia con clasificaciones internacionales. La clasificación nacional de actividades económicas (CNAE-93), la clasificación nacional de ocupaciones (CON-94), la clasificación nacional de productos por actividades (CNPA) o la clasificación nacional de educación (CNED-2000), se han hecho con el objetivo de imponer una definición consistente de categorías o clases.

En el CD-ROM Gescla-97 (gestión de clasificaciones estadísticas), el INE presenta las principales clasificaciones estadísticas con sus correspondencias y metodologías. La existencia de estos esquemas limita el grado de discrecionalidad en la interpretación y, por lo tanto, la variabilidad que introducen los individuos que registran los datos. Aun así, sigue existiendo un cierto grado de interpretación, tanto individual como impuesta oficialmente.

Los efectos de los cambios en las clasificaciones se hacen especialmente evidentes cuando se comparan los datos de algunas estadísticas oficiales en largos periodos de tiempo. Las interpretaciones actuales de la ocupación y la industria han evolucionado mucho a lo largo de los periodos sobre los que existen datos. Es inevitable un cierto grado de incoherencia, pero en la actualidad las agencias estadísticas nacionales deberían documentar todos los cambios que se produjeran en las normas aplicadas.

En algunos casos, las agencias nacionales no pueden imponer normas, y esto se puede deber tanto a la falta de consenso interno y como a la falta de acuerdo entre quienes recopilan los datos y quienes los producen. Cuando los datos se recogen localmente y son agregados y distribuidos desde ámbitos territoriales superiores, existe un riesgo elevado de que se produzca esa falta de acuerdo. Cuanto mayor sea el grado de descentralización, más probable es que se produzcan variaciones en la recogida de datos, la clasificación de las categorías, o el nivel de selectividad en los agregados totales que se envían a las oficinas regionales o nacionales. Por ejemplo, para las primeras estadísticas que se realizaron sobre la asistencia a clase de los escolares en Estados Unidos, se utilizaron distintos criterios de medición: en algunos estados la asistencia se medía por el número de estudiantes presentes en la escuela en un día concreto, mientras que en otros se calculaba a partir del número de estudiantes que asistía a la escuela en cualquier periodo del año [11].

7.3.6. Otros canales de distribución

Un canal fundamental para la difusión de estadísticas lo constituyen los propios organismos que los producen (que pueden difundirlos de manera directa o indirecta, a través de medios oficiales o páginas web). Los usuarios de estas páginas y fuentes son variados: agencias de noticias, profesores, otros organismos gubernamentales, empresas o individuos que buscan evidencia empírica de fácil acceso para reforzar un argumento. Otro canal de difusión de las estadísticas oficiales son los archivos de datos de las universidades, que proporcionan recursos mucho más es-

[11] Para conocer con más detalle cómo el reparto de estadísticas oficiales entre organismos y ámbitos de gobierno ha determinado las estadísticas sobre enseñanza, véase Weiss y Griber, 1987.

tructurados para dar apoyo a la investigación en ciencias sociales.

Entre los archivos de datos internacionales más importantes cabe citar el *Interuniversity Consortium for Political and Social Research* (ICPSR), de la Universidad de Michigan (www.icpsr.umich.edu), en Estados Unidos. En Europa existen diversos archivos de datos destacados, por ejemplo, el *Economic and Social Research Council* (ESRC), de la Universidad de Essex (www.data-archive.ac.uk), en el Reino Unido.

El ICPSR se fundó en 1962, y es probablemente el mayor archivo de datos del mundo. Además de las funciones de depósito y mantenimiento de sus amplios fondos, el ICPSR ofrece cursos de formación sobre técnicas de análisis cuantitativo, material para la enseñanza, archivos de datos combinados e información que permite a los investigadores reproducir un artículo, libro o tesis doctoral publicados. También ofrece análisis de bases de datos específicos a través de Internet, muchos de los cuales han sido utilizados por organismos gubernamentales para producir estadísticas oficiales. El ICPSR es una entidad sin ánimo de lucro mantenida por sus miembros. En la actualidad tiene unos 400 miembros de universidades y centros de investigación de todo el mundo.

El archivo del ESRC, que fue fundado en 1967, contiene la mayor colección de datos en ciencias sociales y humanidades del Reino Unido. Distribuye datos a usuarios en el Reino Unido y también, mediante acuerdos, a otros archivos internacionales e investigadores extranjeros. Además, para algunos temas permite realizar análisis de los datos por Internet, a través de NESSTAR, un consorcio que incluye archivos del Reino Unido, Dinamarca, Noruega y otros cinco socios. NESSTAR utiliza datos y metadatos incluidos en la DDI *(Data Documentation Initiative)*.

Con la expansión de Internet y la consiguiente integración de los recursos electrónicos, cada vez es menos importante el lugar donde nos encontremos, aunque puede influir en la especialización de los fondos, las limitaciones de acceso, o el tipo de ayuda que se presta a los usuarios. Por este motivo, un gran número de archivos facilita también el acceso a las estadísticas oficiales. El *Resource Centre for Access to Data on Europe* (www-rcade.dur.ac.uk) proporciona datos estadísticos procedentes de Eurostat, la UNESCO o la OIT. Podríamos mencionar muchas más fuentes, pero creemos que es más razonable empezar a consultar los archivos académicos indicados aquí, y acceder a otros archivos a través de sus enlaces. Es muy probable que estos enlaces contengan las actualizaciones de los fondos más útiles para la investigación.

7.3.7. Datos agregados y microdatos

Como ya hemos visto antes, una de las características que definen a las estadísticas oficiales es que son datos agregados: tabulaciones de cálculos, ratios, medianas, medias, o cualquier otro resumen sobre un área geográfica concreta. La aparición de las encuestas para la recopilación de estadísticas oficiales, junto con la publicación de muestras de microdatos procedentes de censos nacionales, han hecho que el uso de datos a nivel individual sea una posibilidad atractiva. A partir de estos datos individuales se pueden extraer los datos agregados.

Las estadísticas agregadas, sobre todo las que proceden de censos, suelen estar

disponibles para rangos de agregación muy pequeños, pero el grado de detalle geográfico que se utilice dependerá de los intereses de la investigación. Algunas medidas, como los índices de segregación, aunque reflejan procesos sociales de tipo individual, sólo se pueden definir a partir de rangos superiores de agregación. Con frecuencia, el uso de estadísticas agregadas en una investigación consiste precisamente en relacionar características contextuales de un zona, por ejemplo, áreas con un alto grado de pobreza, niveles de delincuencia elevados, o segregación étnica, con datos individuales, que suelen ser confidenciales. A partir de esta comparación se puede investigar la influencia de las características del entorno en los comportamientos individuales.

Si no existen limitaciones importantes, los científicos sociales pueden aprovecharse de la flexibilidad que ofrecen los microdatos. En el capítulo 5 mencionamos la existencia de fuentes de microdatos disponibles en el INE, en España, o las muestras integradas de microdatos de uso público (IPUMPS), en Estados Unidos. En este sentido, para aquellos que estén interesados en investigar con datos individuales oficiales y al mismo tiempo establecer comparaciones internacionales, cabe mencionar la existencia de las muestras de microdatos de utilidad pública internacionales (*International Integrated Public Use Microdata Samples,* www.ipums.umn.edu/international/index.shtml). Las IPUMS proporcionan microdatos censales que han sido codificados y armonizados de modo que sean equivalentes en el mayor número posible de categorías. El objetivo de este proyecto es disponer de microdatos internacionales para 21 países de seis continentes, entre ellos España y diversos países latinoamericanos, con múltiples censos para cada país, lo que permitirá la realización de comparaciones entre países y en el curso del tiempo.

7.4. EL ESTUDIO DE LAS ESTRUCTURAS SOCIALES

Para concluir el capítulo, en este apartado queremos subrayar que las fuentes estadísticas oficiales ofrecen una oportunidad extraordinaria para la investigación social, ya que proporcionan una cantidad de datos inmensa, que engloba todos los sectores de la vida social y abarca toda la vida del individuo, en su trayectoria escolar, laboral, familiar, sanitaria y en sus relaciones con las instituciones. En definitiva, las estadísticas oficiales son una fuente fundamental para la investigación en ciencias sociales, sobre todo en cuatro aspectos.

En primer lugar, las estadísticas oficiales sirven *a) para analizar y describir la estructura de la sociedad:* la estructura demográfica de la población, los fenómenos migratorios, la distribución del voto y del empleo, la estructura de clase, la movilidad social, las dinámicas del mercado laboral, la organización empresarial, las actividades comerciales, y en general todos los fenómenos ligados a la economía, la estructura de la familia (hijos, convivencia, ancianos), los fenómenos de desviación (delitos y actividad judicial), la marginalidad social, la asistencia y la previsión social (ancianos, pensiones), etc.

En segundo lugar, las estadísticas oficiales permiten *b) estudiar las demarcaciones territoriales* de un país. Pensemos, por ejemplo, en un análisis sobre las diferen-

cias territoriales (entre norte y sur, entre grandes zonas geográficas, entre regiones, entre áreas urbanas y rurales, etc.) en las tasas de criminalidad, de empleo, de estructura de la economía; o en los estudios de geografía electoral, religiosa, etc. Sólo los datos que abarcan todo el territorio (o procedentes de muestras representativas también a escala infranacional) permiten realizar estos estudios, y este requisito es exclusivo de los datos procedentes de estadísticas oficiales.

En tercer lugar, las estadísticas oficiales son casi imprescindibles para *c) realizar estudios comparados entre países.* Aunque recientemente se han desarrollado encuestas por muestreo que, gracias a la coordinación de grupos de investigación nacionales, utilizan instrumentos análogos para la recopilación de datos (cuestionario, tipo de entrevista, estructura de la muestra), de tal modo que los resultados sean comparables entre los distintos países, se trata de episodios esporádicos y sobre sectores limitados [12]. Además, la creciente coordinación entre los institutos centrales de estadística de diversos países (por ejemplo, *Eurostat*, para la Unión Europea) favorece una integración cada vez mayor y mayores posibilidades para contrastar las fuentes estadísticas oficiales de los distintos países.

En cuarto lugar, un aspecto en el que resulta irrenunciable acudir a los datos estadísticos oficiales son los *d) estudios longitudinales.* No hace falta decir que la evolución en el curso de los años (a veces, también, décadas o siglos) de fenómenos como la estructura demográfica de la población, las tendencias de voto, la distribución del empleo en los diversos sectores productivos, los fenómenos migratorios, etc., sólo se puede estudiar recurriendo a las fuentes estadísticas oficiales. La fuente estadística oficial constituye un medio potente, no sólo para estudiar las dinámicas temporales que hemos mencionado *(análisis de series temporales),* sino también para estudiar las sociedades que han precedido a la actual: estadísticas electorales, judiciales, sobre el empleo, los ingresos, el comercio, los impuestos, etc. son instrumentos fundamentales para reconstruir la estructura y las relaciones sociales en épocas históricas anteriores.

Obviamente, el recurso de las estadísticas oficiales no está exento de limitaciones. Aquí mencionaremos las tres más importantes. En primer lugar, el hecho de que el investigador deba utilizar datos ya existentes y recopilados con finalidades distintas de las suyas (a veces, como hemos visto, con fines puramente administrativos), puede hacer que *la naturaleza de los datos no satisfaga las exigencias del investigador.* Por ejemplo, cuando utilizamos el número de ejemplares de una revista católica vendidos por cada 1000 habitantes, o la tasa de abortos, como indicadores de religiosidad, estamos utilizando indicadores muy parciales para el concepto que queremos operacionalizar, con una «parte de indicación» muy reducida. En este caso, sería mucho más apropiado para nuestro objetivo disponer de los datos sobre el número de personas que van a misa los domingos, pero esta información no está

[12] Por ejemplo, la *World Value Survey,* la *European Social Survey,* los eurobarómetros, la *International Social Survey*, el proyecto *Comparative Study of Electoral Systems*, etc. Todas ellas recogen datos utilizando el mismo cuestionario y el mismo tipo de muestreo en distintos países. Además, en la mayoría de estos proyectos los datos se recopilan en varios momentos distintos. Algunas de estas encuestas ya se han comentado en el capítulo 5.

disponible en ninguna fuente estadística.

Más allá de este ejemplo, el problema de los indicadores tiene gran relevancia en la investigación basada en las estadísticas oficiales, y el investigador debe prestar la máxima atención al grado en que dichos indicadores se corresponden con los conceptos estudiados (la cuestión de la validez).

En segundo lugar, las estadísticas oficiales están *limitadas a variables fácticas,* es decir, a datos objetivos y de conducta. De este modo, aunque la estadística oficial nos puede permitir realizar estudios profundos sobre la evolución del abstencionismo electoral en un determinado país, su distribución geográfica, las correlaciones territoriales con indicadores económicos y de estructura social, su distribución por género, edad, profesión, nivel de estudios, etc., no podrá decirnos nada sobre las razones por las que los ciudadanos no van a votar. El único instrumento que podría ayudarnos a estudiar esta dimensión del fenómeno sería una encuesta por muestreo *ad hoc*.

En tercer lugar, y en términos más generales, los datos estadísticos oficiales, cuando se conocen sólo a escala territorial, *c*) ***no son adecuados para el análisis de los comportamientos individuales*** [13]. Aunque encontremos una correlación a escala territorial entre el porcentaje de obreros y el porcentaje de votos a los partidos de izquierda, no podemos llegar a la conclusión de que los obreros votan en mayor medida a los partidos de izquierda. Éste es el conocido problema de la *falacia ecológica* que establece rigurosas limitaciones sobre el tipo de conclusiones que se pueden extraer de las estadísticas oficiales al nivel agregado.

7.5. SÍNTESIS DEL CAPÍTULO 7

1. El primer estudio importante basado en estadísticas oficiales fue realizado por Émile Durkheim a finales del siglo XIX (1897), y trataba sobre el suicidio. Basándose en las tasas de suicidio de los distintos países europeos, Durkheim logró confirmar empíricamente su compleja teoría de la integración social, teoría que sigue manteniendo su validez en la actualidad.
2. Las estadísticas oficiales se diferencian del resto de datos que se suelen utilizar en investigación social, sobre todo los recogidos mediante encuestas por muestreo, en cuatro aspectos: *a*) la producción: generalmente las estadísticas oficiales son realizadas por la administración pública y producidas mediante procedimientos administrativos ordinarios; *b*) las unidades de análisis, que en las estadísticas oficiales no son individuos, sino agregados territoriales de individuos, y que por este motivo se denominan *datos agregados,* por contraste con los *datos individuales*; *c*) el contenido, pues la información recopilada por la

[13] Es decir, cuando la unidad de análisis es un agregado territorial. La observación que se expone a continuación no es aplicable cuando el investigador dispone de datos individuales, como sucede, por ejemplo, cuando realiza una investigación secundaria sobre bases de datos ya existentes, como los barómetros del CIS.

administración pública consiste en registros de acontecimientos o hechos, por lo cual, como norma, las opiniones, las actitudes y las motivaciones están excluidas de estos datos; y *d*) la magnitud del registro, ya que las fuentes estadísticas tradicionales, como los censos, engloban a toda la población, y no a una muestra de la misma.

3. La naturaleza de los datos recogidos, el grado de centralización o difusión de los procedimientos de recogida, los organismos implicados en la obtención de éstos, los modos de distribución, el ámbito geográfico y las normas de clasificación, pueden variar mucho entre los distintos países. En este capítulo describimos algunas de las características principales de las estadísticas oficiales en España y en otros países (Latinoamérica y Europa).

7.6. LECTURAS COMPLEMENTARIAS

Sobre los «indicadores sociales» ya citamos en el capítulo 3 la obra de Carley (1981), que los sitúa en el contexto general de la «medición» en ciencias sociales. Un libro que ha retomado el tema más recientemente (aunque con un punto de vista más pragmático, de hecho se refiere no a los «indicadores sociales», sino a los «indicadores estadísticos»), es el de R. V. *Horn, Statistical Indicators for the Economic & Social Sciences* (Cambridge U. P. 1993, pp. 227).

Para comprender la utilidad de las estadísticas oficiales en el estudio de la transformación y el análisis comparativo entre países, podemos recomendar los diversos volúmenes publicados por el programa internacional *Comparative Charting of Social Change (CCSC),* para Alemania, Francia, Québec, Grecia, España, Rusia y Canadá. En cuanto al Reino Unido, remitimos a A. H. Halsey y J. Webb (eds.), *Twentieth-Century British Social Trends* (Macmillan 2000, pp. 760), el último volumen de una serie que periódicamente toma una fotografía de los cambios producidos en el Reino Unido desde principios del siglo XX.

Para una descripción similar de los cambios sociales acaecidos durante el último siglo en Estados Unidos, remitimos a la obra de T. Caplow, L. Hicks y B. J. Wattenberg, *The First Measured Century: An Illustratred Guide to Trends in America, 1990-2000* (AEI Press 2001, pp. 307).

Para las fuentes de información demográficas en España, remitimos al libro de D. Reher y A. Valero de la serie de *Cuadernos Metodológicos del CIS* (n.° 13): *Fuentes de Información Demográfica en España* (Madrid: CIS, 1995, pp. 109) o al número 24 de la misma colección, *Banco de datos,* de M. Cordero (Madrid: CIS, 1998, pp. 199), que ofrece información detallada sobre las bases de datos de encuestas representativas del CIS.

CAPÍTULO 8

EL MUESTREO

8.1. POBLACIÓN Y MUESTREO 272

8.2. ERROR DE MUESTREO 274

8.3. TAMAÑO DE LA MUESTRA 277

8.4. MUESTRAS PROBABILÍSTICAS 282

8.5. MUESTRAS NO PROBABILÍSTICAS 287

8.6. PROBLEMAS DE MUESTREO EN INVESTIGACIÓN SOCIAL 289

8.6.1. Error de cobertura 290
8.6.2. Error de no respuesta 291

8.7. PONDERACIÓN 293

8.8. REPRESENTATIVIDAD Y TAMAÑO DE LA MUESTRA 296

8.9. SÍNTESIS DEL CAPÍTULO 8 298

8.10. LECTURAS COMPLEMENTARIAS 299

En investigación social casi nunca es posible estudiar la totalidad de la realidad social en cuestión, y es necesario seleccionar una parte o muestra de esta realidad para su estudio. Por ello el procedimiento del muestreo suele ser la primera operación empírica que el investigador debe realizar. En este capítulo estudiaremos la técnica del muestreo de forma sistemática, con una breve historia de su aplicación a la investigación social, la presentación de sus fundamentos matemáticos, y la ilustración y el planteamiento de los diseños de muestreo más destacados. Al final del capítulo trataremos los problemas que plantea la investigación social con relación a los procedimientos clásicos de muestreo elaborados por la estadística.

8.1. POBLACIÓN Y MUESTREO

El muestreo, es decir, la observación de una parte de una población para extraer información sobre la totalidad de la misma, es un acto humano casi instintivo. Pensemos en el cocinero que prueba la comida que está preparando para decidir si le debe añadir sal, el profesor que examina al estudiante planteándole preguntas sobre una parte del programa de la asignatura, el periodista que entrevista a un taxista para saber qué opina «la gente» sobre una noticia de actualidad. Pese a ser habitual, el procedimiento del muestreo no ha logrado una sistematización científica hasta hace muy poco tiempo, gracias al desarrollo de las nociones de azar y casualidad.

Cuando en el lenguaje común se habla de elegir a alguien o algo «al azar», en un conjunto de individuos u objetos, se suele considerar una elección «a la buena de Dios» más que la aplicación de un procedimiento riguroso en el que el azar se entiende como una cuestión de probabilidades. Una elección casual no es una elección sin reglas, ya que el procedimiento de muestreo aleatorio debe realizarse siguiendo unos criterios muy precisos, y el azar, el verdadero azar probabilístico, tiene sus leyes. De hecho, a diferencia de lo que el sentido común parece sugerir, uno de los fenómenos mejor conocidos por la ciencia, hasta el punto de que lo podemos expresar en fórmulas matemáticas, es precisamente el *azar*. Por ejemplo, podemos calcular perfectamente la probabilidad de obtener una pareja de seises en diez lanzamientos de dados.

El *muestreo* es el procedimiento por el cual, de un conjunto de unidades que forman el objeto de estudio (la *población*), se elige un número reducido de unidades *(muestra)* aplicando unos criterios tales que permitan generalizar los resultados obtenidos del estudio de la muestra a toda la población. El muestreo ha sido estudiado exhaustivamente en sus formulaciones matemáticas, dando lugar a una rama específica de la estadística conocida como *teoría de las muestras*. Para realizar investigación social no es necesario dominar esta teoría, basta con conocer sus fundamentos y sus líneas generales.

Los motivos por los que se recurre al muestreo para realizar investigación social son obvios. La encuesta por muestreo: *a*) *permite reducir los costes* de la encuesta; *b*) *permite reducir el tiempo* necesario para la recogida y elaboración de los datos; *c*) *conlleva ventajas de organización,* en el sentido de que no es necesario reclutar, formar y dirigir a las legiones de encuestadores que son necesarias para la

encuesta total; y *d) permite un estudio más profundo y esmerado,* ya que los recursos se pueden concentrar en el control de la calidad de la encuesta. En algunas situaciones, la encuesta por muestreo es la única opción posible, por ejemplo cuando no se conoce totalmente la composición y el tamaño de la población de referencia.

Desde el punto de vista de las aplicaciones prácticas, las empresas de sondeos de opinión aportaron una contribución fundamental a la difusión del procedimiento de muestreo probabilístico. Para ilustrar las ventajas de esta técnica citaremos un episodio que, en cierto sentido, marcó su consolidación definitiva. En 1936, con el objetivo de pronosticar el resultado de las elecciones presidenciales de ese año, la popular revista estadounidense *Literary Digest,* envió por correo un facsímil de papeleta electoral a más de diez millones de nombres extraídos de guías telefónicas y registros de propietarios de vehículos. La iniciativa tuvo un éxito enorme, pues se recibieron dos millones de respuestas, lo que constituye una muestra de dimensiones colosales. Según este estudio, el demócrata Franklin D. Roosevelt obtendría sólo el 41 % de los votos, lo que hacía prever el triunfo de su contrincante, el republicano Alf Landon. Pero estas previsiones no se cumplieron, pues Roosevelt recibió cerca del 61 % de los votos del electorado. En cambio, las empresas de sondeos Gallup, Roper y Crossley, sí predijeron correctamente el resultado de las elecciones. Estas empresas utilizaban muestras de dimensiones mucho menores, pero extraídas de forma aleatoria de la población total.

¿Cuál fue el fallo de *Literary Digest*? Recordando la exposición sobre los distintos tipos de error del apartado 3.5, se trató de un *error de cobertura:* las listas de la población utilizadas no eran completas. A falta de un registro de los ciudadanos (que en Estados Unidos no existe), los organizadores del sondeo recurrieron a listas

Resumen 8.1. LOS ORÍGENES DE LA MUESTRA

Las primeras propuestas para sustituir la totalidad de la población por una parte de la misma para la recopilación de los datos las hizo el estadístico noruego Anders Kiaer en 1895. Pero estas propuestas no tuvieron una buena acogida. El método de Kiaer no recurría a la elección aleatoria, sino que utilizaba información sobre la población procedente de fuentes censales, según un procedimiento que luego sería definido como *muestreo subjetivo (por elección razonada).* Varios años después, en 1926, la intuición inicial de Kiaer fue formulada en términos más rigurosos por el estadístico A. L. Bowley, que introdujo el concepto de *muestreo aleatorio simple.* Sin embargo, habría que esperar todavía una década hasta la sistematización teórica definitiva de todo el campo, que llegó con los trabajos del estadístico polaco Jerzy Neyman, que estableció la distinción entre *muestreo subjetivo* y *muestreo probabilístico,* y fijó las bases teóricas del segundo, que deja en manos del azar la elección de las unidades que se incluirán en la muestra.

de propietarios de coches y titulares de líneas telefónicas. Sin embargo, en aquella época de depresión económica, estos ciudadanos no representaban a todo el país, sino sólo a las clases más acomodadas, que solían votar a los republicanos.

El segundo problema de la muestra con relación a la población total fue un *error de no-respuesta,* o error de autoselección. En efecto, los individuos que respondieron no eran iguales a los que no respondieron; probablemente los primeros tenían un nivel de estudios superior, eran lectores habituales de la revista, y tenían características similares en lo que concierne al voto. Las empresas de sondeos pronosticaron el voto con mayor precisión precisamente porque su muestra, aunque más pequeña, era más *representativa* de la población. Este episodio, que tuvo gran resonancia en los medios de comunicación y en la opinión pública, supuso la consagración oficial de las pequeñas muestras probabilísticas frente a las grandes muestras basadas en la lógica del censo.

8.2. ERROR DE MUESTREO

La *población* se define como un conjunto de N unidades, también llamadas unidades estadísticas o unidades de análisis, que constituyen el objeto de un estudio, donde N es el *tamaño de la población.* El término «población» podría sugerir un conjunto de seres humanos. Sin embargo, en estadística, este término tiene un significado mucho más general y se refiere a un conjunto cualquiera de objetos (hombres, casas, empresas, territorios, sucesos, etc.): pensemos en el conjunto de municipios de una determinada comarca, en los delitos cometidos en un año concreto, o en los productos fabricados diariamente por una máquina. [1]

Nuestro objetivo es estudiar las propiedades o variables de estas unidades *(X, Y, Z,* etc.). Estudiarlas significa conocer los valores característicos que adoptan sobre toda la población y que pueden servirnos para describir la distribución total de las variables o las relaciones entre estas variables, es decir, los llamados *parámetros.* Por ejemplo, si la unidad es el individuo y la variable X es la renta, y queremos conocer la renta media, el parámetro de la población que queremos conocer es una *media.* Si la variable Y es el voto y queremos conocer la proporción de personas que han votado al partido A, el parámetro de la población que queremos conocer es una *proporción.* Si queremos estudiar la relación entre las variables X e Y, el parámetro que queremos conocer es un *coeficiente de correlación.*

Para llegar a conocer la renta media (o algún parámetro) de la población, tendremos que conocer las rentas X_i de todas las unidades que componen la población. Pero no las conocemos, y a menudo resulta imposible conocerlas. Entonces es cuando

[1] El uso del término «población» deriva del que fue uno de los primeros campos de aplicación de la estadística, la demografía. Muchos estadísticos han empleado el término «población» como sinónimo del término «universo», mientras que otros utilizan el primer término para designar al conjunto *real* (y finito) de los casos, y el segundo para hacer referencia a conjuntos *hipotéticos* (e infinitos). Nosotros preferimos la segunda opción. Por ello, dado que en el texto nos referimos siempre al conjunto real de las unidades, utilizaremos siempre el término población.

recurrimos a la recopilación de los datos en un subconjunto de dicha población.

Una muestra es un conjunto de n unidades de muestreo, los llamados casos, seleccionadas entre las N unidades que componen la población, de modo que sean representativas de la misma para los fines del estudio, donde n es el *tamaño de la muestra*. La población es el objeto que se quiere conocer, la muestra es el instrumento para conocerla. El *muestreo* es el procedimiento que seguimos para elegir las n unidades de muestreo del total de las N unidades que componen la población [2].

Pese a las ventajas obvias enumeradas anteriormente, el hecho de trabajar con un conjunto reducido de n unidades en lugar de estudiar la totalidad de la población N, también tiene una serie de inconvenientes. En efecto, mientras que la investigación sobre el total de la población proporciona el valor exacto del parámetro que se quiere conocer, el estudio por muestreo ofrece sólo una estimación o valor aproximado de éste. Por ejemplo, si queremos conocer la renta mensual media de una población y realizamos una investigación sobre la población total, al final del estudio podremos formular una conclusión de este tipo: «La renta media de esta población es de 950 euros». En cambio, si realizamos una investigación por muestreo, nuestra afirmación será del tipo: «Existe un 95 % de probabilidad de que la renta media de la población sea de 925 ± 25 euros». Esta segunda afirmación contiene dos elementos de indeterminación, o mejor dicho, un elemento de probabilidad y uno de aproximación. En primer lugar, la afirmación es sólo probable: el investigador no está seguro al 100 % de que la renta media esté dentro del intervalo de 900-950 euros, peor sabe que existe un 5 % de probabilidad de que el valor real de la renta media en la población quede fuera de este intervalo. En segundo lugar, el investigador no puede concretar el valor de la renta media de la población, sólo puede indicar un intervalo en el que se sitúa este valor, entre un mínimo y un máximo. En este caso diríamos que su estimación tiene un cierto *grado de confianza* y consiste en la determinación de un *intervalo de confianza* en el que se sitúa el valor de la estadística de la población.

Por tanto, la estimación de la muestra estará afectada por un error, el llamado *error de muestreo*. Si V es el valor (desconocido) del parámetro en la población (una media, una proporción, etc.), v es el valor hallado en la muestra (es decir, su estimación), y e es el error de muestreo, podemos escribir la fórmula siguiente [3]:

$$\underset{\text{parámetro de la población}}{V} = \underset{\text{estimación de la muestra}}{v} \pm \underset{\text{error de muestreo}}{e}$$

[2] Conviene que el lector entienda bien la distinción entre *diseño de muestreo* (el procedimiento teórico), *muestreo* (la ejecución del diseño) y *muestra* (el resultado del muestreo). Por ejemplo, un diseño de muestreo puede ser aleatorio (si se establece que las unidades de muestreo se elijan mediante extracción por sorteo), y la muestra resultante podría no serlo, debido a que el proyecto no se puede realizar completamente. En las encuestas por muestreo, si no se logra realizar la entrevista a un porcentaje significativo de los nombres elegidos porque resulta imposible localizarlos o porque se niegan a responder, la muestra resultante no se podrá considerar una muestra aleatoria.

[3] Seguimos la convención de indicar con letras mayúsculas los valores en la población y con letras minúsculas los correspondientes valores en la muestra.

(desconocido)

La estimación *v,* se puede calcular fácilmente con los datos de la muestra. Pero *el verdadero problema consiste en calcular el error de muestreo.* Se trata de un problema de gran relevancia. De forma intuitiva, para determinar la magnitud del error, tendremos que saber también algo de la población: ¿pero qué hacemos si no sabemos nada de ella? Si la muestra se ha elegido según un proceso rigurosamente aleatorio, es decir, si se trata de una *muestra probabilística,* la teoría estadística nos permite calcular la magnitud de dicho error.

En el muestreo aleatorio simple, que es el caso más sencillo, el parámetro a estudiar es una media, y el error de muestreo se calcula mediante la fórmula:

$$e = z\sigma(\bar{X}) = z\frac{s}{\sqrt{n}}\sqrt{1-f}\ ,$$

donde:

$\sigma(\bar{X})$ = error típico de la media muestral;
z = grado de confianza de la estimación (que, para el 95 % es 1.96);
s = desviación típica muestral de la variable analizada;
n = tamaño de la muestra;
$1-f$ = factor de corrección para poblaciones finitas, donde f = fracción de muestreo = n/N

Obsérvese que el error es mayor:

— cuanto mayor sea el grado de confianza que el investigador quiere tener en su estimación; si se conforma con el 95 % [4], entonces z = 1,96; si quiere una probabilidad mayor, por ejemplo del 99 %, el valor de z aumenta a 2,58, y así sucesivamente.
— cuanto más alta sea la variabilidad de la variable estudiada; si queremos estimar la renta media de los habitantes de un barrio de una ciudad, manteniendo constantes el resto de condiciones, el error de muestreo será más alto en un barrio mixto que en un barrio socialmente homogéneo.
— cuanto menor sea el tamaño de la muestra (lo cual es obvio).

Por último, tenemos el término $\sqrt{1-f}$, que se denomina *factor de corrección para poblaciones finitas,* donde f es la fracción de muestreo, es decir, la relación entre el tamaño de la muestra y el tamaño de la población *(n/N).* Nótese que si la población es infinita, o en cualquier caso mucho mayor que n (cuando la muestra es inferior al 5 % de la población), el factor de corrección se acerca tanto a 1 que

[4] El 95 % de probabilidad de que el intervalo de confianza $\bar{X} \pm e$ contenga la media desconocida de la población.

se puede neutralizar. En este caso la dimensión N de la población ni siquiera interviene en la fórmula de cálculo del error. Por tanto, podemos decir que la dimensión de la población influye en el error de muestreo sólo en una segunda aproximación, cuando la muestra no es demasiado pequeña respecto a la población. *De hecho, es el tamaño de la muestra, más que la fracción de muestreo, lo que determina la entidad del error.* Pero volveremos más adelante sobre esta importante cuestión.

Si trabajamos con variables categóricas (nominales y ordinales con pocas categorías), la medida sintética más común es la proporción (por ejemplo: proporción de votos al partido *A*, proporción de partidarios de una medida adoptada por el gobierno, proporción de católicos, etc.). En este caso la fórmula del error de muestreo es algo diferente:

$$e = z\sqrt{\frac{pq}{n-1}}\sqrt{1-f}$$

Donde z, n y f tienen el mismo significado de la fórmula anterior, mientras que:

p = proporción en la muestra para la categoría a examen
$q = 1 - p$

Obsérvese que esta fórmula se corresponde con la fórmula anterior para la media, teniendo en cuenta que para variables categóricas la desviación estándar es $(\sigma = \sqrt{PQ})$ [5]. Con respecto a la población, dado que P y Q no se conocen, en la fórmula se han sustituido por los correspondientes valores p y q se han calculado con los datos recogidos en la muestra.

8.3. TAMAÑO DE LA MUESTRA

Probablemente, la pregunta «¿qué tamaño debe tener mi muestra?» es la primera que se plantea el investigador que haya decidido utilizar una muestra para su estudio. En efecto, la determinación previa, aunque sea aproximada, del tamaño de la muestra representa uno de los criterios básicos para determinar la magnitud de la investigación y hacer una previsión de sus costes.

Usando las fórmulas anteriores para el error de muestreo podemos deducir fácilmente el tamaño de la muestra: basta con sustituir e por el error que estamos dispuestos a aceptar, y resolver la ecuación tomando como incógnita a n. Supongamos por ahora que la población es tan grande en relación con la muestra que se puede obviar el factor de corrección para poblaciones finitas; en tal caso, calculan-

[5] Nos referimos a una variable dicotómica, que sólo puede tomar los valores 0 y 1 (aunque la variable sea politómica, siempre se puede reinterpretar como dicotómica, oponiendo en cada caso una modalidad a todas las demás juntas).

do n con la ecuación, para el caso de la estimación de una media o de una proporción, tenemos:

$$n = \left(\frac{zs}{e}\right)^2 \qquad n \cong n-1 = \frac{z^2 pq}{e^2}$$

Por tanto, el tamaño de la muestra es directamente proporcional al nivel de confianza para la estimación *(z)* y la variabilidad del fenómeno estudiado, e inversamente proporcional al error que el investigador está dispuesto a aceptar. En otras palabras, el tamaño de la muestra es directamente proporcional a la precisión deseada. Nótese que el investigador decide el valor de z y e (establece la fiabilidad y la precisión que quiere que tengan sus estimaciones), mientras que s, p (y q) no son conocidos, ya que el investigador quiere saber cuántos casos debe incluir en la muestra *antes* de empezar a recoger datos. Por tanto, no tendrá más remedio que recurrir a valoraciones aproximadas de s y p extraídas de estudios anteriores, de la opinión de expertos o de un estudio piloto cuyos casos pueden constituir la primera parte de la muestra.

Es importante hacer hincapié en la presencia en las fórmulas del tamaño de la población N. Es una opinión bastante difundida que el tamaño de la muestra debe ser proporcional al de la población. Si por ejemplo, hay que realizar dos muestras sobre las poblaciones de dos municipios, de 10000 y 100000 habitantes respectivamente, el investigador poco experimentado que se deja llevar por el sentido común, sugerirá probablemente que en el segundo caso se utilice una muestra 10 veces superior a la del primer caso. Pero este criterio es erróneo.

De lo dicho hasta ahora queda claro que el tamaño de la población ni siquiera interviene en la primera aproximación, en la fórmula que establece el tamaño n de la muestra. Interviene en la segunda aproximación, con la introducción del factor de corrección para poblaciones finitas, sólo cuando se requiere que el tamaño de la muestra no sea despreciable en relación con el tamaño de la población (es decir, cuando n sea superior al 5 % de N). Para determinar el tamaño de la muestra, si la fiabilidad y el error de la estimación se mantienen constantes, en primera instancia sólo interviene la dispersión de la variable estudiada, s, para variables cardinales, o $\sqrt{pq}$, para variables categóricas. Es perfectamente posible que, para los dos municipios antes mencionados, una muestra de 1000 casos sea suficiente para obtener estimaciones con el misma grado de precisión para ambos municipios, a pesar de que los tamaños de sus poblaciones son muy distintos.

Hasta ahora hemos establecido el tamaño de la muestra tomando como referencia la estimación de los parámetros de una única variable, pero, por lo general, el investigador se plantea estimar los parámetros de muchas variables. En ese caso, podrá aplicar el método indicado con cada una de las variables más importantes para la recogida de los datos, y después tomar como tamaño de la muestra la n más elevada de todas ellas. Si, como sucede a menudo, el investigador debe estimar sólo proporciones, la cuestión se simplifica. En efecto, al medirse la dispersión mediante $\sqrt{pq}$, es fácil constatar que este índice toma el valor más elevado cuando $p = q = 0{,}50$. Por lo tanto, se calcula el tamaño requerido para la muestra en el caso más desfavorable,

es decir, cuando $p = q = 0{,}50$; con ello nos aseguramos de que dicha muestra será suficiente para las variables que presenten proporciones distintas de 0,50.

La tabla 8.1 presenta, con un grado de confianza del 95 %, el tamaño de la muestra para tres valores distintos del error «absoluto» (5, 2 y 1 %), y para varios valores de la dimensión N de la población. Por ejemplo, con una población de 10000 unidades y un error máximo de 2 %, el tamaño requerido para la muestra es de 2000 casos; si nos contentamos con una precisión de 5 %, el tamaño requerido es de 400 casos. Obsérvese que mientras que un error del 5 % es bastante elevado, un error de 2 % es muy bajo [6].

De la tabla se deduce que 2500 casos son suficientes para obtener estimaciones con una precisión del 2 % (y sólo 400 casos para estimaciones con una precisión de ± 5 %) *cualquiera que sea el tamaño de la población* (podría tratarse incluso de la población de todo el país o de la población mundial). Este hecho puede resultar sorprendente, pero recordemos que todo ello es válido siempre que tengamos una *muestra aleatoria simple* rigurosa. Esto significa, por ejemplo, en el caso de la

Resumen 8.2. EL FACTOR DE CORRECCIÓN PARA POBLACIONES FINITAS EN EL CÁLCULO DEL TAMAÑO DE LA MUESTRA

Las fórmulas indicadas para calcular el tamaño n de la muestra no tienen en cuenta el factor de corrección para poblaciones finitas, igual a $\sqrt{1-f}$, que se debe introducir cuando el tamaño de la muestra sea superior al 5 % de la población ($f = n/N > 0{,}05$). Si tenemos en cuenta este factor, las fórmulas quedan como sigue:

$$n = \left(\frac{zs}{e}\right)^2 (1-f) \qquad n = \frac{z^2 pq}{e^2}(1-f)$$

Sustituyendo f por su valor n/N, y resolviendo la ecuación para n, obtenemos unas fórmulas bastante complejas. Para ejemplificar su utilización, estas fórmulas se suelen separar en dos pasos: primero se aplican las anteriores, que nos proporcionan una primera medida n_0 del tamaño de la muestra [donde $n_0 = n/(1-f)$]:

$$n_0 = \left(\frac{zs}{e}\right)^2 \qquad n_0 = \frac{z^2 pq}{e^2}$$

Si el valor de n_0 calculado es inferior al 5 % de N, se usa como tamaño de la muestra. Pero si es más elevado, es necesario introducir el factor de corrección, y se puede pasar de n_0 al valor correcto de n mediante la fórmula:

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

población mundial, que debemos contar con una lista de todos los habitantes del planeta, extraer por sorteo 2500 nombres, y recorrer el mundo hasta que los hayamos entrevistado a todos (suponiendo que los encontramos a todos y que ninguno de ellos se niega a responder), lo cual es imposible de realizar.

Esta consideración sobre la viabilidad del muestreo en las ciencias sociales, nos lleva a otro problema, también relacionado con la especificidad de la investigación social. La lógica estadística que hemos usado para la determinación del tamaño de la muestra es una lógica «monovariada», es decir, considera las variables por separado y no estudia las relaciones entre ellas. Para ilustrar la posibilidad de relacionar variables entre sí, consideremos el siguiente ejemplo.

Supongamos que, con el procedimiento de muestreo aleatorio simple, hemos extraído una muestra de 420 sujetos de la población de un barrio de 50000 habitantes, y que utilizando un cuestionario hemos encontrado un porcentaje de católicos practicantes del 25,7 %. Aplicando la fórmula anterior obtenemos lo siguiente (al tratarse de una muestra muy pequeña en relación con la población, hemos omitido el factor de corrección):

$$e = z\sqrt{\frac{pq}{n-1}} = 1{,}96\sqrt{\frac{0{,}257 \cdot 0{,}743}{419}} = 0{,}042$$

Esto significa un error del 4,2 %. Simplificando mucho podemos decir que, con una probabilidad del 95 %, el porcentaje de católicos practicantes es del 25,7 % ± 4,2 %.

Sin embargo, el investigador social no se contentará con este dato, y querrá saber también, por ejemplo, si la práctica religiosa varía con el nivel de formación.

Tabla 8.1. Tamaño *n* requerido para la muestra, en un muestreo aleatorio simple, según la precisión de la estimación (error *máximo* cometido) y el tamaño *N* de la población (con un nivel de confianza de la estimación del 95 %).

Precisión de la estimación					
5 %		2 %		1 %	
N	n	N	n	N	n
100	80	100	96	100	99
300	170	300	270	300	296
500	220	500	415	500	475
1.000	285	1.000	715	1.000	910
5.000	370	5.000	1.660	5.000	3.330
>8.000	400(n_0)	10.000	2.000	10.000	5.000
		>50.000	2.500(n_0)	20.000	6.350
				>200.000(n_0)	10.000

Para ello construirá una tabla en la que la práctica religiosa se representará por separado para los sujetos con estudios y los sujetos sin estudios.

Supongamos que hay 265 personas que tienen estudios universitarios de grado superior, y que entre éstas el 22,6 % son practicantes. ¿Cuál es el error de esta estimación? Aplicando las mismas fórmulas de antes, encontramos que las estimaciones son 22,6 % ± 5,0, es decir, el porcentaje de practicantes entre los sujetos que tienen un nivel de formación alto está comprendido entre el 17,6 y el 27,6 %. El error, por tanto, ha aumentado notablemente en comparación con la estimación anterior. ¿Cómo se explica esto? Simplemente por el hecho de que para calcular la tasa de práctica religiosa entre personas con distintos niveles de formación hemos dividido la muestra en dos submuestras de dimensiones reducidas, y, como sabemos, al disminuir el tamaño de la muestra aumenta el error de muestreo.

Si, por otra parte, el investigador, después de haber constatado la existencia de una relación entre nivel de estudios y práctica religiosa, decide dividir la muestra en tres grupos de edad para asegurarse de que la relación entre educación y práctica religiosa no es espuria por efecto de la edad, habrá fragmentado de nuevo la muestra, y ello provocará un nuevo aumento del error de muestreo.

Por tanto, en investigación social, el investigador no puede establecer la dimensión de la muestra considerando las variables por separado (análisis univariado), sino que debe tener en cuenta el tipo de análisis que pretende hacer, y las subdivisiones que aplicará a la muestra para analizarla (análisis multivariado).

Por ejemplo, en una investigación sobre el comportamiento electoral, si se quiere estudiar las características sociales de los electores de partidos pequeños que representan en torno al 5 % de los votos válidos, hay que tener presente que, sobre 1000 respuestas válidas, esta cuota corresponde a unos 50 casos; esta cantidad podría ser suficiente para hacer una estimación sobre la cantidad de votos que recibe un partido pequeño, pero es insuficiente para hacer un análisis de las características de este electorado. Por ejemplo, no podemos relacionar el voto a ese partido con el género, el nivel de estudios, etc., de sus seguidores, ya que ello significaría dividir esos pocos 50 casos. Por lo general, el tamaño de la muestra deberá ser tanto más grande cuanto más minoritario sea el fenómeno que se quiere estudiar.

Del mismo modo, si realizamos un estudio sobre movilidad social y queremos relacionar la profesión del sujeto con la de su padre, y considerar las variables del género y la generación, tendremos que fraccionar considerablemente la muestra. Una tabla de movilidad que relacione la profesión del sujeto con la del padre, considerando siete tipos de ocupación, fraccionaría la muestra en 49 celdas; si a continuación dividimos de nuevo los casos en dos grupos de género y dos grupos de edad, tendremos un total de 196 celdas. De ello se deriva que, para un análisis de este tipo, no bastaría con una muestra de 2000 casos, puesto que si tomamos una media de diez casos por celda (ya de por sí modesta), tendríamos muchas celdas vacías (con cero casos), que corresponderían a las profesiones menos comunes. No es casual que los estudios de movilidad social se suelan realizar sobre muestras de tamaño mucho mayor.

8.4. MUESTRAS PROBABILÍSTICAS

Decimos que una muestra es *probabilística* [6] cuando cada unidad que la compone se extrae de la población con una probabilidad conocida (y distinta de cero). La muestra del *Literary Digest* citada en el apartado 8.1 no era probabilística, ya que, por una parte, las personas que no tenían vehículo ni teléfono tenían una probabilidad igual a cero de ser incluidas en la muestra, y, por otra, las personas que disponían de teléfono y poseían uno o más vehículos tenían una mayor probabilidad de estar en la muestra. Del mismo modo, si construimos una muestra de estudiantes universitarios extrayendo por sorteo un cierto número de estudiantes entre los presentes un día cualquiera en la universidad, no obtendremos una muestra probabilística, ya que, *a)* los estudiantes que no suelen asistir a clase tienen una probabilidad nula de formar parte de la muestra; *b)* los matriculados en el primer curso, que como ya se sabe suelen asistir más a clase, tienen más probabilidad de ser incluidos en la muestra que los estudiantes de cursos superiores; y *c)* éstos últimos tienen más probabilidad de ser seleccionados que los estudiantes que tengan sólo algunas asignaturas pendientes [7]. En este apartado presentaremos las características básicas de los principales modelos de muestreo probabilístico.

Muestreo aleatorio simple. Dentro de las muestras probabilísticas, el caso más básico es el *muestreo aleatorio simple.* Formalmente, hablamos de muestreo aleatorio simple cuando todas las unidades de la población de referencia tienen la *misma probabilidad* de ser incluidas en la muestra. Para poder realizar este tipo de muestreo, el investigador deberá disponer, en primer lugar, de la lista completa de los miembros de la población. De este modo podrá asociar a cada una de las N unidades de dicha lista un número, y extraer después de forma aleatoria los n números correspondientes a los sujetos que formarán la muestra [8].

A pesar de su sencillez, el muestreo aleatorio simple raras veces se usa en la investigación social, por una parte, porque la selección se deja en manos del azar y no se utiliza la información conocida a priori sobre la población, y por otra, porque en las investigaciones a gran escala conllevaría un plan de investigación muy costoso y complejo desde el punto de vista organizativo; además, requiere una lista completa de todos los miembros de la población, que con frecuencia no está disponible.

Muestreo sistemático. Un procedimiento estadísticamente equivalente al anterior, en el sentido de que también produce una muestra aleatoria simple, es el *muestreo sistemático.* Éste difiere del muestreo aleatorio simple sólo en la técnica de

---

[6] Los adjetivos «probabilístico» y «aleatorio» tienen, en este contexto, el mismo significado, por lo que se puede hablar indistintamente de «muestra probabilística» o «muestra aleatoria».

[7] El hecho de que las probabilidades de ser escogidos sean desconocidas es más grave que el hecho de que dichas probabilidades sean distintas entre sí.

[8] La extracción puede efectuarse de una urna en la que se habrán introducido N bolas numeradas, o con ayuda de tablas de números aleatorios, o bien generando una secuencia de números aleatorios con un ordenador.

selección de los sujetos. Las unidades de la muestra no se escogen por sorteo (o tabla de números aleatorios), sino utilizando la lista de sujetos y seleccionando sistemáticamente uno cada cierto intervalo. Si se conoce el tamaño N de la población de referencia y se ha determinado el tamaño n de la muestra, se elige una unidad cada $k = N / n$ unidades de la población, comenzando con un número extraído al azar entre 1 y k (k es el llamado intervalo de muestreo). Por ejemplo, si de una población de $N = 8235$ unidades se quiere extraer una muestra de $n = 500$ unidades, se elige una unidad cada 16 ($k = 8235/500 = 16{,}7$), empezando con un número extraído al azar entre 1 y 16. Si éste fuera 12, entonces los sujetos elegidos serían el 12°, el 28°, el 44°, y así sucesivamente; de esta forma se seleccionarán 514 sujetos, 14 de los cuales serán descartados [9].

El muestreo sistemático es equivalente al aleatorio, salvo en las situaciones, poco frecuentes, en que la lista de los miembros de la población de referencia presenta periodicidades coincidentes con el intervalo de muestreo. Por ejemplo, al extraer a los empleados de una fábrica de una lista en la que están agrupados por departamentos y, dentro de cada departamento, por antigüedad laboral, si el intervalo de muestreo coincide con el número de personas de los departamentos, podría suceder que la muestra estuviera integrada por trabajadores con la misma antigüedad.

El interés del muestreo sistemático no radica tanto en la simplificación que introduce en la operación de extracción, como en el hecho de que permite obtener muestras aleatorias aunque no se disponga de la lista de los miembros de la población y se desconozca el tamaño la misma *(N)*. En la investigación social el muestreo sistemático se utiliza con frecuencia, precisamente porque en muchos casos la lista de los miembros de la población de referencia no está disponible. Por ejemplo, en el caso de las encuestas realizadas a pie de urna, en las que se le pide al elector que repita su voto en una ficha ofrecida por el entrevistador, se suele entrevistar a un elector por cada k electores que salen del colegio electoral. Del mismo modo, en los estudios de mercado se entrevista a un cliente por cada k clientes que salen del establecimiento tras haber efectuado su compra [10].

Muestreo estratificado. Como hemos apuntado al hablar del error de muestreo, la precisión de las estimaciones de la muestra depende, entre otras cosas, del tamaño de la muestra y la variabilidad de la distribución del fenómeno estudiado dentro de la población de referencia. Obviamente, la precisión de las estimaciones disminuye cuando la dimensión de la muestra es menor y la variabilidad de la distribución del fenómeno estudiado es mayor. Ello significa que si el fenómeno estudiado presenta

---

[9] ¿Cómo se eligen los sujetos a descartar? Para responder a esta pregunta hay que tener en cuenta la siguiente propiedad formal de las muestras: una muestra aleatoria simple de una muestra aleatoria simple es también una muestra aleatoria simple. Por tanto, los sujetos extraídos en exceso podrán ser descartados utilizando también un procedimiento de selección aleatoria.

[10] Este k es distinto del k citado en el párrafo anterior, en el sentido de que no se puede calcular como relación entre N y n.

una variabilidad muy elevada, para mantener un cierto nivel de precisión en las estimaciones de la muestra será necesario aumentar el tamaño de la misma. Como alternativa, si el fenómeno objeto de estudio presenta áreas de mayor homogeneidad, es posible aumentar la eficiencia (el grado de precisión en relación con el tamaño) de la muestra recurriendo al *muestreo estratificado.*

Este diseño de muestreo consta de tres fases: *a*) primero hay que subdividir la población de referencia en subpoblaciones o *estratos* lo más homogéneas posible en relación con el fenómeno estudiado, empleando como criterio de estratificación una variable correlacionada con ese fenómeno; *b*) en segundo lugar se extrae una muestra de cada estrato mediante un procedimiento aleatorio simple; y *c*) por último, se unen las muestras correspondientes a cada uno de los estratos para obtener la muestra total.

Por ejemplo, si el fenómeno objeto de estudio es la renta, los sujetos de la población de referencia pueden subdividirse en estratos en función de la variable «ocupación», que, como es sabido, está muy relacionada con la renta. Si los estratos son, por ejemplo, becarios, asalariados, trabajadores autónomos y profesionales liberales, extraeremos una muestra de cada uno de estos estratos por separado. Después, uniremos en una sola muestra estas cuatro muestras parciales. De este modo, dividimos una población heterogénea, que habría requerido una muestra muy amplia, en estratos relativamente homogéneos que pueden estudiarse utilizando muestras pequeñas; la suma de estas muestras más pequeñas sigue siendo inferior al tamaño de la muestra que habría sido necesario extraer de la población total de referencia.

Dentro de cada uno de los estratos, la extracción se realiza mediante muestreo aleatorio simple. La muestra se denomina *estratificada proporcional* (o *autoponderada*), cuando reproduce la misma composición de los estratos en la población. Por ejemplo, si los becarios, los asalariados, los trabajadores autónomos y los profesionales liberales son, respectivamente, el 40, el 40, el 15 y el 5 % de la población, y nosotros construimos una muestra total de 1000 casos formada por 400 becarios, 400 asalariados, 150 trabajadores autónomos y 50 profesionales liberales, obtendremos una muestra estratificada proporcional.

Si, en cambio, decidimos que unos estratos tengan una representación mayor y otros una representación menor, obtendremos una muestra estratificada no proporcional. Por ejemplo, podríamos decidir sobrerrepresentar los estratos menos numerosos para realizar un estudio más exhaustivo sobre ellos, u optar por una muestra en la que todos los estratos estén igualmente representados. En todos estos casos la muestra total no reproduce la composición de la población, y en los análisis de los datos se realizará una operación de *ponderación* (hablaremos de ella en el apartado 8.7), para restablecer la composición correcta de la población dentro de la muestra. Entre los distintos tipos de muestreo estratificado no proporcional el más eficiente teóricamente es la muestra *estratificada óptima,* en la que el tamaño de la muestra extraída de cada estrato es proporcional a la variabilidad del fenómeno estudiado dentro de ese estrato.

Muestreo a estadios múltiples. Esta técnica no conlleva una mejora en la eficiencia

con respecto al muestreo aleatorio simple, pero sí una simplificación del procedimiento de extracción y una reducción de los costes de la investigación. El muestreo a estadios múltiples es la única opción cuando no disponemos de la lista completa de los miembros de la población de referencia, o cuando, debido a la excesiva dispersión de la población, con el muestreo aleatorio simple o estratificado se obtendría una muestra de unidades distribuidas en un territorio demasiado amplio y, por tanto, de difícil acceso.

En el muestreo a estadios múltiples la población se subdivide en varios niveles ordenados jerárquicamente, que se extraen sucesivamente con un procedimiento «de embudo». Por ejemplo, si tenemos que construir una muestra de profesores de enseñanza primaria de todo un país, éstos pueden subdividirse en unidades primarias, representadas por circunscripciones didácticas (siempre que cada profesor pertenezca a una sola unidad primaria), y unidades secundarias, que serían los propios profesores. El muestreo se desarrolla en dos fases, es decir, en dos extracciones sucesivas. En la primera fase se extrae una muestra de unidades primarias (por ejemplo, 50 circunscripciones didácticas), para lo cual debemos disponer de la lista completa de unidades primarias. En la segunda fase se extrae aleatoriamente una muestra de unidades secundarias (profesores) de cada una de las unidades primarias seleccionadas en la primera extracción, para lo cual debemos disponer de la lista de profesores, pero *sólo* para las unidades primarias de la primera fase. Osbviamente, las fases de muestreo pueden ser más de dos.

Este procedimiento ofrece varias ventajas: *a*) no es necesario disponer de la lista de toda la población de referencia, sino sólo de las listas de las subpoblaciones de las unidades de orden superior extraídas; *b*) la investigación se concentra en las unidades extraídas, con una considerable reducción de los costes. Pero también hay inconvenientes. El muestreo a estadios múltiples comporta una pérdida de eficiencia debido a que los casos pertenecientes a una misma unidad de orden superior tienden a ser en parte similares. Por otro lado, la teoría estadística ligada a este modelo de muestreo es bastante compleja.

Muestreo por conglomerados. Este procedimiento es similar al muestreo a estadios múltiples, y se emplea cuando la población se subdivide de forma natural en grupos de unidades. Pensemos en las agrupaciones de individuos constituidas por familias, clases escolares, departamentos de trabajo, habitaciones de hospital, etc. Estos grupos se llaman *conglomerados,* de ahí el nombre del procedimiento. En este caso no se extraen las unidades básicas (los individuos), sino los conglomerados, y después todas las unidades básicas pertenecientes a los conglomerados extraídos se incluyen en la muestra. De esa forma nos beneficiamos de la proximidad que hay entre los individuos pertenecientes al mismo conglomerado. Por ejemplo, en una investigación sobre la afición futbolística se puede tomar como población de referencia el conjunto de los aficionados que viajan en un tren reservado para ir a ver el partido de su equipo fuera de casa, elegir con muestreo sistemático dos compartimentos por cada vagón del tren, y entrevistar a todos los individuos presentes en los compartimentos seleccionados. Es un procedimiento muy útil cuando, como en este ejemplo, no disponemos de la lista de las unidades básicas, y existe la

posibilidad de extraer los conglomerados con un procedimiento probabilístico.

También existe un procedimiento que combina el muestreo a estadios múltiples con el muestreo por conglomerados. Por ejemplo, en la encuesta sobre población activa que realiza el Instituto Nacional de Estadística en España, primero se selecciona una muestra de secciones censales entre las más de 30000 en que está dividida España; luego, dentro de cada sección censal se extrae una muestra de familias (elegidas entre las viviendas familiares), cuyos componentes serán entrevistados [11].

Muestreo por áreas. Una variante del muestreo a estadios múltiples es el muestreo por áreas. Esta técnica se aplica cuando no se dispone de la lista de los miembros de la población de referencia. Por ejemplo, una muestra de la población residente en una serie de municipios se suele construir mediante un muestreo a estadios múltiples, seleccionando primero los municipios y después, dentro de éstos, la muestra de los ciudadanos a partir de los censos municipales. Pero si el censo municipal no está disponible, el procedimiento no puede realizarse. En este caso, se recurre al muestreo por áreas.

Un ejemplo de ello sería la muestra que utiliza el instituto académico de investigación social más conocido de Estados Unidos, el *Survey Research Center* de la Universidad de Michigan, que se construye del siguiente modo. El país se subdivide en áreas primarias (comarcas, grupos de comarcas o áreas metropolitanas). Entre éstas se escogen 74. Cada una de estas áreas primarias se subdivide en zonas, por ejemplo, una gran ciudad, un grupo de cuatro o cinco ciudades de tamaño mediano, un área rural bastante amplia, etc. A continuación, cada zona se subdivide en partes, áreas que suelen estar definidas por límites naturales, como ciudades medianas-pequeñas, barrios de las ciudades más grandes, áreas rurales delimitadas por carreteras o ríos, etc. Cada una de estas partes se subdivide en segmentos, cada uno de los cuales contiene entre cuatro y dieciséis unidades familiares. Las entrevistas se realizarán en una selección de estas unidades familiares.

Todas las selecciones se realizan con un procedimiento probabilístico, que suele ser el muestreo aleatorio simple, excepto para las áreas primarias iniciales, donde se aplica un procedimiento de muestreo estratificado, para tener una representación proporcional de la población de todas las zonas geográficas del país. Si hay que hacer también una elección casual dentro de las unidades familiares, como cuando la entrevista se realiza a los ciudadanos de más de 18 años y en la unidad familiar elegida hay tres personas con estas características, se aplica siempre un procedimiento aleatorio, por ejemplo, se elige a la persona que cumple años antes, o se ordenan los sujetos por edades decrecientes y se elige uno con una tabla de números aleatorios.

La técnica del muestreo por áreas es útil no sólo cuando no se dispone de las listas de la población a nivel subterritorial, sino también cuando estas listas están incompletas. Por ejemplo, con el muestreo por áreas hay más posibilidades de que

[11] Este tipo de muestreo se denomina «muestreo de dos etapas con estratificación de las unidades de la primera etapa (secciones censales)».

estemos incluyendo también a los inmigrantes ilegales, residentes temporales, etc.

8.5. MUESTRAS NO PROBABILÍSTICAS

Cuando el diseño probabilístico no es posible, o cuando se sabe a priori que no podrá realizarse, se recurre desde el principio a las denominadas *muestras no probabilísticas*. En este apartado presentaremos los tipos más comunes.

Muestreo por cuotas. La muestra por cuotas tal vez sea el diseño de muestreo más difundido, sobre todo en los estudios de mercado y los sondeos de opinión. Para realizar este procedimiento, en primer lugar hay que dividir la población de referencia en un cierto número de estratos definidos por algunas variables cuya distribución es conocida (por ejemplo género, edad y estudios). Entonces se calcula el «peso» proporcional de cada estrato, es decir, la parte de la población total que representa (obviamente, la suma de estos pesos debe ser igual a uno). Por último, multiplicando cada uno de estos pesos por el tamaño n de la muestra se determina la cuota, es decir, el número de entrevistas que se deben realizar en cada estrato.

Hasta aquí el muestreo por cuotas coincide con el muestreo estratificado. La diferencia reside en el hecho de que, observando el requisito de la cantidad de unidades que debe tener cada cuota (por ejemplo, se sabe que debería entrevistarse a veinte varones-jóvenes-con formación), *el entrevistador es libre de elegir a los sujetos que va a entrevistar.* Por tanto, el muestreo por cuotas es un muestreo estratificado en el que la elección de las unidades se encomienda al entrevistador, y en el que la magnitud de la cuota limita la libertad de elección, haciendo que la muestra total reproduzca la distribución de la población con relación a las variables planteadas al establecer las cuotas.

Este procedimiento tiene algunas limitaciones. La libertad que concede al entrevistador hace que, una vez satisfechas las exigencias prefijadas, aplique criterios de selección utilitaristas, dando prioridad a los casos que le resulten más fáciles de localizar (círculos de conocidos, sujetos residentes en una misma región o del mismo grupo social, etc.), eludiendo cualquier tipo de dificultad y no insistiendo demasiado cuando un individuo se muestra reacio a realizar la entrevista.

El ámbito de los sondeos electorales nos ofrece el fracaso más sonado del procedimiento de muestreo por cuotas. En 1948 los tres institutos de sondeos más prestigiosos de Estados Unidos predijeron, utilizando el muestreo por cuotas, que el republicano Thomas Dewey superaría al demócrata Henry Truman en las elecciones presidenciales. Sin embargo, los resultados dieron a Truman el 50 % de los votos, y a Dewey el 45 %. En este caso el error estuvo en las elecciones subjetivas de los entrevistadores, que, después de cumplir las exigencias planteadas al establecer las cuotas (lugar de residencia, edad, raza, nivel socioeconómico), dieron preferencia a sujetos con tendencias republicanas, sin ser conscientes de ello. Esta experiencia llevó a las empresas de sondeos a adaptar su estrategia e restringir más la libertad de elección de los entrevistadores, asignándoles un itinerario fijo en su búsqueda de los sujetos a entrevistar. Según este procedimiento, para cubrir sus cuotas, los entrevis-

tadores deben visitar los hogares que les asignen, y sólo deben buscar otros en caso de rechazo o falta de correspondencia entre las personas identificadas y las definidas por las exigencias de las cuotas.

El problema es que, aunque en este caso ya no existe el riesgo de que el entrevistador escoja a los entrevistados de manera arbitraria, la selección sigue estando distorsionada a favor de los individuos más fácilmente localizables: cuando el entrevistador no encuentra a una persona en su casa, simplemente llama a la puerta siguiente, con lo cual en la muestra quedan infrarrepresentados los sujetos menos accesibles o reacios a ser entrevistados, mientras que las connotaciones de estos dos grupos sociales podrían estar relacionadas con las variables estudiadas.

El muestreo por cuotas ha generado una larga polémica. Los estadísticos lo consideran poco recomendable por su carácter poco científico. Además, mientras que en una muestra tradicional la más mínima distorsión respecto a la población puede aflorar si comparamos la distribución de alguna variable observada en la población y la correspondiente observada en la muestra, en el muestreo por cuotas la distorsión queda enmascarada por las propias cuotas. Sin embargo, el muestreo por cuotas siempre ha tenido mucho éxito entre los especialistas en estudios de mercado y sondeos de opinión, quienes afirman que produce resultados satisfactorios, y un notable ahorro en los costes. Ellos consideran que no vale la pena emplear muchos recursos en las muestras probabilísticas, ya que, de todos modos, en la investigación social sigue habiendo otras fuentes de error relevantes.

Muestreo de «bola de nieve». Este modelo de muestreo resulta particularmente útil para el estudio de poblaciones *clandestinas.* Nos referimos a grupos sociales cuyos miembros, por motivos éticos, legales, ideológicos o políticos, tienden a ocultar su identidad, como inmigrantes ilegales, miembros de sectas religiosas, activistas de grupos políticos, evasores fiscales, trabajadores no regularizados, personas sin techo, drogadictos, etc. Este procedimiento se utiliza también cuando la población objeto de estudio está constituida por «elementos raros», grupos poco numerosos y dispersos por el territorio que, sin embargo, están de alguna manera en contacto entre sí (miembros de minorías religiosas, grupos específicos, asociaciones, etc.).

El muestreo de «bola de nieve» consiste en identificar a los sujetos que se incluirán en la muestra a partir de los propios entrevistados. Se parte de una pequeña cantidad de individuos que cumplen los requisitos exigidos, y que se utilizan como informadores para localizar a otros individuos con características idénticas. Conforme avanza la investigación, el número de sujetos debería crecer de manera exponencial, de ahí la analogía con la bola de nieve.

Este tipo de muestreo presenta el inconveniente de que selecciona a las personas más activas socialmente y más visibles (incluso en ámbitos tan restringidos). Además, existe el riesgo de que la cadena de identificación lleve a vías demasiado específicas. En general, cuando se usa este muestreo se deben plantear exigencias basadas en la información disponible sobre el fenómeno, para evitar los riesgos indicados.

Muestreo subjetivo por decisión razonada. En este caso, las unidades de la muestra no se eligen usando procedimientos probabilísticos, sino en función de algunas de

sus características. Este tipo de muestreo se emplea, por ejemplo, cuando el tamaño de la muestra es muy limitado y se quieren evitar oscilaciones casuales que distancie excesivamente a la muestra de las características de la población.

Por ejemplo, en la elección de los barrios de una ciudad se puede decidir incluir en la muestra el mismo número de barrios céntricos y periféricos, o de barrios obreros y burgueses. En un estudio realizado en Italia se utilizó este tipo de muestreo para seleccionar 100 municipios de los 8000 que hay en todo el país. Para ello fue necesario clasificar estos 8000 municipios en veinte estratos según un criterio que combinaba el tamaño (cinco tipos) y la zona geográfica (cuatro tipos). Dentro de estos estratos, el número de municipios incluidos en la muestra fue proporcional al peso demográfico de la población correspondiente, y se aplicaron también criterios de elección racionales. Por ejemplo, se procedió de tal manera que las veinte regiones estuvieran representadas por al menos una capital de provincia, que los municipios menores distaran de la capital no menos de quince kilómetros, que en cada región hubiera una ciudad con universidad, etc. El proceso de elección se hace, por tanto, aplicando criterios racionales, sin recurrir nunca a la selección casual.

Una variante de este procedimiento es el *muestreo compensado* o *equilibrado*, que consiste en seleccionar las unidades de tal forma que la media de la muestra, para determinadas variables, se acerque a la media de la población. Por ejemplo, la muestra de municipios antes mencionada puede construirse con el requisito de que, en cada estrato de la muestra, la media de los votos al centroizquierda y al centroderecha en las elecciones no presente una diferencia superior al 3 % con respecto a la media del voto del correspondiente estrato en la población.

Obviamente, el muestreo subjetivo no ofrece las garantías de la muestra probabilística ni las ventajas de la sencillez de ejecución del muestreo por cuotas. Sin embargo, como hemos dicho, resulta adecuado para muestras muy pequeñas, sobre todo cuando las unidades de análisis no son individuos, sino instituciones, municipios, etc., o en situaciones particulares en las que la importancia de ciertas unidades exige su inclusión en la muestra para garantizar que la información recopilada esté completa.

8.6. PROBLEMAS DE MUESTREO EN INVESTIGACIÓN SOCIAL

En ciencias sociales, la aplicación más común de la técnica del muestreo es la encuesta por muestreo, en la que de una población de individuos se extrae una muestra para entrevistarlos. En este caso resulta ineludible recurrir a la técnica de la muestra. A pesar de esta necesidad fundamental, uno de los campos en los que la teoría de la muestra resulta más difícil de aplicar es precisamente la encuesta por muestreo, es decir, cuando la población de referencia (en sentido estadístico) está compuesta por individuos y el instrumento de investigación es un cuestionario o una entrevista. Esto se debe a que, como ya indicamos en el apartado 3.8, el denominado «error de selección» no se limita al error de muestreo, sino que presenta otros dos componen-

tes: el error de cobertura y el error de no respuesta.

8.6.1. Error de cobertura

En primer lugar se plantea el problema de la *lista de la población.* Hemos dicho que para poder obtener una muestra probabilística debemos poder asignar a todas las unidades de la población una probabilidad conocida de que sean incluidas en la muestra. Pero, ¿qué sucede cuando no conocemos las unidades?, ¿cómo podemos, por ejemplo, extraer una muestra aleatoria de los inmigrantes en una región si no contamos con una lista de todos los inmigrantes presentes en la misma? Tampoco podemos recurrir a un muestreo sistemático, dado que los inmigrantes no están situados físicamente en un único lugar, en el que podamos incluir a cada uno en un determinado intervalo [12].

En líneas generales podemos decir que no existen problemas graves cuando el objeto del estudio es la población total (incluso de todo un país). Existen censos municipales (de todos los residentes) y listas electorales (de los ciudadanos de más de 18 años) que nos proporcionan la relación completa de la población de la cual podemos extraer la muestra con procedimientos aleatorios.

El problema se plantea cuando queremos estudiar segmentos específicos de la población. En efecto, salvo en los casos en que los sujetos estudiados formen parte de asociaciones formalmente constituidas o pertenezcan a una institución, o por alguna razón se hayan puesto en contacto con la administración pública, en la mayoría de los estudios sobre segmentos particulares de la población no disponemos de una lista de éstos (pensemos en estudios sobre obreros, desempleados, inmigrantes, católicos practicantes, etc.). Tampoco existe, en ninguno de estos casos, un lugar en el que podamos localizar y pasar revista a todos los sujetos, para poder realizar un muestreo sistemático.

Esto vale para el caso en que la unidad de análisis es un individuo. Cuando la unidad es un colectivo, la situación es algo más fácil, dado que las agrupaciones de individuos, (escuelas, universidades, empresas, hospitales o establecimientos comerciales), suelen estar institucionalizadas y registradas. Pero todo se complica cuando vamos más allá de las formas oficiales (pensemos, por ejemplo, en cursos de informática, asociaciones deportivas, compañías teatrales, etc.).

Siempre que no dispongamos de una lista de la población (o un equivalente que

[12] En rigor no es necesario disponer de una lista real de la población para asignar a todos los sujetos una probabilidad conocida de que sean incluidos en la muestra. En el muestreo a estadios múltiples basta con disponer de una lista de las unidades primarias y las unidades básicas pertenecientes a las unidades primarias elegidas; en el muestreo por áreas basta con tener la lista de las áreas en los distintos niveles. El muestreo sistemático puede realizarse también sin una lista previa, siempre que exista un lugar donde se localice toda la población y el investigador pueda pasar revista, en el curso de la selección, a todos los sujetos (por ejemplo, entrevistando a un individuo por cada diez que salen de un colegio electoral o entran en un museo). En cualquier caso, se trata siempre de situaciones en las que el investigador está en condiciones de identificar a todos los componentes de la población, y éstos pueden tener una probabilidad conocida (generalmente igual) de ser incluidos en la muestra.

permita al investigador acceder teóricamente a toda la población) *tendremos que renunciar a las técnicas de muestreo probabilístico.* Y es que en estos casos no podemos asignar a *todas* las unidades de la población una determinada probabilidad de extracción.

Naturalmente no basta con que existan listas de la población. Hay que tener en cuenta los problemas de la actualización, las duplicaciones y las omisiones. El problema de *las listas incompletas* es el más grave. Recordemos de nuevo el caso del *Literary Digest*: el uso de las listas de titulares de líneas telefónicas y propietarios de vehículos como sustituto de la lista de electores excluyó a una parte importante de la población que no poseía ni una cosa ni la otra. La clave no es la omisión en sí misma, ya que si los sujetos presentes en la lista fueran una muestra aleatoria de la población total, el problema no existiría. La cuestión es que los sujetos que están en la lista son, en cuanto a tendencias, distintos de los que han quedado fuera [13].

En estos casos el investigador social tiene tres posibilidades:

— *Redefinir a la población.* Podría especificar, por ejemplo, que la investigación estudia a los comerciantes, sino a los miembros de asociaciones de comerciantes, motivando la elección y extrayendo las consecuencias de ello también en la fase de interpretación de los resultados.
— *Obviar a los excluidos.* Si las omisiones no representan una parte muy importante del total (por ejemplo, son sólo un 10-15 %), y existen razones motivadas para considerar que los excluidos de la lista no son muy distintos de los incluidos, con respecto a los temas estudiados, el error podría considerarse aceptable en el contexto de las ciencias sociales, donde, como sabemos, existen varias fuentes de errores [14].
— *Completar la muestra.* Por ejemplo, en un país donde se sabe que el 20 % de la población no tiene teléfono, se podría completar la muestra incluyendo en ella un 20 % de individuos que no tengan teléfono y que se seleccionarían con una técnica de muestreo distinta.

8.6.2. Error de no respuesta

Cuando el procedimiento de muestreo se aplica en ciencias sociales, el hecho de disponer de la lista de la población y extraer de ella una muestra aplicando rigurosamente el proceso de extracción por sorteo, no es suficiente para garantizar que la muestra sea efectivamente probabilística. Si la población está compuesta por seres

[13] El razonamiento se extiende, por analogía, a los muestreos probabilísticos que no emplean una lista en sentido estricto, por ejemplo, el muestreo sistemático, por conglomerados, o por área. Si entrevistamos, con un muestreo sistemático, a los aficionados que van en tren a ver a su equipo cuando juega fuera de casa, obtenemos una muestra distorsionada de dichos aficionados, porque excluimos a aquellos que se desplazan en su propio vehículo, que son en efecto distintos (por ejemplo, en edad y poder adquisitivo), de los que van en tren.

[14] Por ejemplo, en un país donde el 85 % de las familias tenga teléfono y donde sólo el 5 % de los números no estén publicados en las guías telefónicas, es una costumbre generalizada aceptar la lista de abonados como un inventario fiel de la población de las familias del país.

humanos, una vez extraída la muestra sobre el papel, se plantea el problema de realizar la encuesta, y es fácil que una muestra que empezó siendo probabilística, al concluir la fase de recogida de datos haya dejado de serlo. Todo ello genera el *error de no respuesta,* que puede deberse a dos causas muy distintas: *a*) la falta de contacto con los sujetos extraídos; o, *b*) la negativa de una parte de los sujetos a realizar la entrevista.

El problema de la *falta de contacto con los sujetos extraídos* es, sobre todo, un problema de costes, pero no por ello menos grave o fácil de resolver. Este problema tiene dos vertientes. Una es la *dificultad para acceder* a los sujetos, que deben encontrarse dentro de unos límites impuestos por el presupuesto disponible para la investigación. Por ejemplo, en una investigación sobre la población de un país, aunque dispongamos de una lista de la misma, es impensable realizar un sorteo aleatorio simple, y luego ir a entrevistar a los sujetos dispersos por todo el territorio. La otra vertiente del problema es la *imposibilidad para localizar* a los sujetos de la muestra. También en este caso hay limitaciones de carácter económico que imponen que tras dos o tres intentos fallidos de encontrar al sujeto de la muestra en su casa [15], éste sea sustituido por otro nombre de la lista de reserva.

Hay que añadir que aunque la dificultad para acceder a los sujetos de la muestra y la imposibilidad de localizarlos representan obstáculos graves para la correcta aplicación del procedimiento probabilístico, pueden salvarse, por ejemplo, con diseños de muestreo alternativos (como el muestreo a estadios múltiples) o invirtiendo más recursos en la fase de localización de los sujetos (por ejemplo, estableciendo que se visite la casa de los sujetos de la muestra tantas veces como sea necesario, hasta localizarlos).

Mucho más grave es el problema de la *negativa del sujeto a hacer la entrevista.* Existen motivos fundados para considerar que quienes se niegan a responder son distintos de quienes responden. A menudo se trata de personas ancianas, con poca formación, mujeres, personas que viven solas, o personas muy ocupadas. También puede haber una motivación para responder distinta en función de las características del entrevistado: por ejemplo, en una investigación sobre profesores es más probable que los más desmotivados y menos comprometidos con su profesión no quieran hacer la entrevista. Otras veces puede haber un rechazo de carácter ideológico o una preocupación ligada al contexto específico de la entrevista: por ejemplo, en una encuesta a los trabajadores de una fábrica encargada por los jefes podrían negarse a responder los más concienciados políticamente. Otro ejemplo serían los rechazos a responder a preguntas específicas: por ejemplo, a la pregunta de si está afiliado a algún partido es más fácil que el afiliado se muestre más reticente que el no afiliado, el afiliado a un partido de la oposición más que el afiliado al partido gobernante, el obrero que el profesional, etc.

¿Cuál es la tasa de ausencia de respuesta habitual en las encuestas por muestreo? En Estados Unidos la tasa para una muestra de la población mayor de edad de

[15] En este ejemplo, nos referimos a entrevistas cara a cara en casa del entrevistado, pero el problema se plantea también en las otras modalidades de entrevista.

todo el país gira en torno a una media del 25 % para las entrevistas cara a cara [Schuman y Presser, 1981, 16-17]. En este punto hay que detenerse en las entrevistas telefónicas. A partir de los estudios realizados, éstas presentan tasas de no respuesta superiores a las correspondientes entrevistas cara a cara. La diferencia en Estados Unidos parece situarse entre el 5 y el 10 % [Collins *et al.* 1988, 213]. Una comparación puntual realizada en el Reino Unido entre investigaciones idénticas pero efectuadas en parte con entrevistas cara a cara y en parte con entrevistas telefónicas revela desfases aun más altos [*ibidem,* 217] [16].

El lector ya habrá percibido la importancia cuantitativa de los sujetos que han sido incluidos en la muestra pero por un motivo u otro no llegan a ser entrevistados. En la práctica, el problema se resuelve sustituyendo a quienes no se ha conseguido entrevistar por nombres de la lista de reserva, que también han sido extraídos de la población por sorteo. Pero este hecho no elimina la posible distorsión de la muestra. Si los no entrevistados son, por ejemplo, ancianos, personas sin estudios, habitantes de las periferias, etc., y los sustituimos con otros nombres extraídos de la población al azar, estaremos sustituyendo un segmento de la población con unas características particulares (marginales) por un grupo «medio» (ya que son sujetos extraídos de forma aleatoria), y, por consiguiente, obtendremos una muestra en la que los ciudadanos marginales están infrarrepresentados.

El investigador dispone de pocos los recursos para hacer frente al error de la no respuesta. En primer lugar, aunque la sugerencia resulte obvia, podrá intentar reducir el número de contactos fallidos con intentos repetidos de ponerse en contacto con las personas que no pudo localizar para realizar la entrevista; esta forma de proceder debería tenerse en cuenta a la hora de calcular los costes de la investigación. En segundo lugar, una vez concluida la fase de recogida de datos, podría recurrir a técnicas de ponderación. Este procedimiento consiste básicamente en atribuir a los sujetos que no se haya podido entrevistar las respuestas medias dadas por el grupo social al que pertenecen. En el apartado que sigue lo explicamos con más detalle.

8.7. PONDERACIÓN

La «ponderación» es el procedimiento con el que modificamos de manera artificial la composición de la muestra para hacerla más próxima a la distribución real de la población. La ponderación se realiza durante la elaboración de los datos y mediante operaciones matemáticas, atribuyendo un «peso» a cada uno de los casos de la muestra, que variará en función de sus características. Los procedimientos de ponderación pueden ser muy complejos, pero aquí presentaremos sólo las formas y situaciones básicas, que pueden reducirse a tres, según se basen en: *a*) probabilidades de inclusión de las unidades en la muestra; *b*) conocimientos disponibles sobre la población; y, *c*) conocimientos disponibles sobre las no respuestas. Como vere-

[16] Para ampliar información sobre la tasa de falta de respuestas en las entrevistas telefónicas y en los cuestionarios postales, remitimos a Dillman [1978].

mos, será necesario introducir una distinción clara entre el primer caso y los otros dos.

• Cuando la ponderación se realiza a partir de las *probabilidades de inclusión de los sujetos en la muestra*, su descripción es fácil. Con frecuencia, las muestras empleadas en investigación social no se corresponden con la situación ideal en la que todas las unidades tienen la misma probabilidad de ser seleccionadas, y que puede ilustrarse con la imagen de las bolas extraídas de la urna. Ya hemos mencionado el muestreo estratificado con un tamaño de la muestra para cada estrato intencionadamente no proporcional al peso del estrato en la población (para sobrerrepresentar a unos estratos e infrarrepresentar a otros). Supongamos que las mujeres empresarias son el 2.5 % de la población y en la muestra incluimos el 5 % con el fin de disponer de un número suficiente de casos para hacer estimaciones sobre esta categoría específica. En la fase de elaboración de los datos, tendremos que asignar a las mujeres empresarias su peso real en la población si queremos obtener estimaciones de parámetros sobre la población total. Por lo tanto, a cada caso incluido en la muestra de mujeres empresarias deberá aplicarse un coeficiente que reduzca a la mitad su peso dentro de la muestra. La operación de ponderación la exige, en este caso, la adaptación del propio diseño de la muestra, y los términos de esta operación se conocen desde el principio.

También existen situaciones en las que el mismo procedimiento de extracción produce diferentes probabilidades de inclusión en la muestra. Pensemos, por ejemplo, en el objetivo de obtener una muestra de familias teniendo en cuenta que en los municipios no existe un censo de familias, aunque sí existe un censo de individuos. En este caso, podría extraerse una muestra de individuos del censo e incluir en la muestra las familias a las que pertenecen estos individuos. Sin embargo, con este procedimiento, las familias más numerosas tendrían más probabilidades de ser elegidas. Por ello, en la fase de elaboración será necesario ponderar los datos, atribuyendo a cada familia un peso inversamente proporcional a su probabilidad de inclusión en la muestra, es decir, al número de miembros que la componen [17].

En todos estos casos la probabilidad de inclusión no es la misma para todos los sujetos, pero se conoce, y a partir de este conocimiento se calculan los pesos para la ponderación. Por tanto, nos mantenemos dentro de la lógica del muestreo probabilístico. La situación es muy distinta cuando la muestra no es probabilística, es decir, cuando no se conoce la probabilidad de inclusión, o cuando está afectada por graves errores de cobertura y no respuesta, lo cuales, de hecho, la convierten en no probabilística, aunque el diseño inicial se planteara en términos probabilísticos. En estos casos, las operaciones de ponderación consisten en simples ajustes para limitar los errores y distorsiones, que se realizan a partir de datos sobre la población, o bien de información sobre los sujetos que no contestan.

---

[17] Otro caso de desigualdad en la probabilidad de inclusión generada por el procedimiento de muestreo es, por ejemplo, el muestreo a estadios múltiples en el que las unidades primarias no tienen la misma dimensión.

• El caso más común de ponderación realizada *en función de los conocimientos disponibles sobre la población* es la denominada postestratificación. Se trata, como el término sugiere, de una estratificación de la muestra realizada después de la fase de recogida de datos. Por ejemplo, si se conoce la distribución de algunas variables en la población gracias a fuentes de información externas como el censo electoral, esta distribución conocida se compara con la resultante de la muestra y se corrigen los datos de la muestra de forma que se correspondan con los datos de la población en relación con las variables relevantes. La operación se realiza multiplicando cada caso de la muestra por un coeficiente de ponderación *(peso)* igual a la relación *proporción verdadera / proporción observada* de la categoría (o estrato, si se consideran varias variables juntas) de pertenencia.

Por ejemplo, si los varones resultan ser el 49 % (proporción verdadera, conocida por el censo), en lugar del 58 % (proporción observada en la muestra, según los datos recogidos), multiplicamos cada sujeto varón por el peso: 49/58 = 0,84 [18], y cada mujer por el peso: 51/42 = 1,21. Esta operación viene a ser como si redujéramos el número de varones, imponiendo que cada varón cuente no como una unidad, sino como 0,84 unidades, y aumentáramos el número de mujeres imponiendo que cada una de ellas cuente como 1,21 unidades, todo ello sin variar el tamaño de la muestra. De este modo, el número total de sujetos sigue siendo el mismo, mientras que la proporción de varones y mujeres cambia.

La operación de ponderación a partir de la información disponibles sobre la población puede realizarse en función de cualquier variable cuya distribución en la población se conozca. Por ejemplo, en un sondeo preelectoral, para atenuar el riesgo de distorsiones de la muestra sobre las orientaciones políticas de los sujetos incluidos en ella, se puede preguntar a los entrevistados qué votaron en las últimas elecciones, y ponderar así la muestra a partir del conocimiento de la variable «resultado electoral en las últimas elecciones» en la población.

• La tercera forma de ponderación se realiza utilizando los *conocimientos disponibles sobre los sujetos que no responden.* Este procedimiento se aplica para contrarrestar el error de no respuesta, pero no sirve para contrarrestar el error de cobertura. Por ejemplo, en las encuestas a pie de urna, el encuestador clasifica a las personas que rechazan responder según el género y edad. Las encuestas realizadas se ponderan teniendo en cuenta los rechazos. Si del total de personas interpeladas, las mujeres ancianas fueron el 9 %, pero de ellas sólo respondió el 5 %, a las que han aceptado responder se les atribuirá un peso igual a 9/5 = 1,8. En este caso el coeficiente de ponderación es el inverso de la tasa de respuesta de cada categoría [19].

---

[18] Este cálculo procede de una sencilla proporción: si los individuos con peso 1 suman 58, ¿cuánto deberían pesar para dar como suma 49? (X : 49 = 1 : 58). Como se deriva de este ejemplo, si optamos por la ponderación, tendremos que ponderar todos los estratos en los que se subdivide la población.

[19] Otro caso en el que es recomendable la ponderación de las no respuestas son los sondeos telefónicos, que suelen presentar un número elevado de no respuestas y éstas suelen concentrarse en segmentos específicos de la población (ancianos con un nivel de formación bajo). En estos casos, si el sujeto rechaza ser entrevistado, es recomendable identificar el género por la voz e intentar que revele su edad (e incluso su nivel de estudios); de este modo será posible ponderar las respuestas.

¿Cuál es el sentido de esta operación? Si no se ponderara y se aceptaran sin modificaciones los resultados obtenidos a partir de la muestra (generalizándolos a toda la población), sería como atribuir a los que no han respondido el comportamiento medio de los que sí lo han hecho. Esto es un error, pues sabemos con certeza que quienes no responden son predominantemente ancianos, mujeres, personas con un bajo nivel de formación, etc. Por tanto, es preferible atribuir a quienes no responden el comportamiento medio de las personas pertenecientes a su mismo grupo social, más que el comportamiento medio de toda la población.

Las operaciones de ponderación en las formas *b* y *c*, que se aplican para compensar errores de cobertura o de no respuesta, suelen considerarse poco adecuadas e incluso se piensa que son una manipulación incorrecta de los datos. Sin embargo, hay que recordar que si la ponderación (en las formas *b* y *c*) conlleva un componente de arbitrariedad, la falta de ponderación conlleva un error indudable, que se comete al generalizar a toda la población los resultados derivados de muestras claramente distorsionadas en relación con los fenómenos estudiados. En estos casos la ponderación permite reducir la distorsión presente en los datos, aunque no soluciona el problema de forma definitiva y completa.

El verdadero problema estriba en que la ponderación viene a disfrazar, mediante un artificio contable, las distorsiones de la muestra. Por ello, es necesario explicar de manera explícita los procedimientos de ponderación aplicados; de lo contrario es preferible exponer claramente las distorsiones presentes en la muestra, para que, tanto el analista como el lector, estén informados de la inadecuación de los datos y la tengan en cuenta a la hora de interpretarlos.

8.8. REPRESENTATIVIDAD Y TAMAÑO DE LA MUESTRA

Hay dos factores que intervienen en la «buena calidad» de una muestra, es decir, en su validez científica: su representatividad y su tamaño. La *representatividad* es lo contrario del error estándar de las estimaciones, es decir, lo que en el apartado 8.2 hemos definido como «error de muestreo». La representatividad es lo contrario del «error de selección» (figura 3.5), que incluye el error de muestreo anterior (también llamado «error muestral») y los errores de falta de cobertura y de no respuesta («distorsiones» o «errores no muestrales») [Kish, 1965, 25].

La representatividad de una muestra depende en gran medida de su tamaño: si la muestra es pequeña, el error estándar de las estimaciones será alto y la muestra tendrá una representatividad baja. Por ejemplo, un total de diez casos, aunque se hayan seleccionado de manera perfectamente aleatoria, sólo puede dar lugar a muestras tan imprecisas que resultarán prácticamente inútiles.

Sin embargo, el tamaño no es suficiente para tener una buena muestra, ya que hay que tener en cuenta también su representatividad: una muestra muy grande, pero que tiene errores no muestrales graves, por ejemplo, errores de cobertura o de no respuesta, es, efectivamente, una muestra de mala calidad. En ciencias sociales es prácticamente imposible evitar totalmente los errores de cobertura y no respuesta. Sin embargo, el investigador puede ***minimizarlos,*** intentando que el procedi-

miento de muestreo sea lo más parecido posible al modelo de muestreo probabilístico.

A veces los objetivos de reducción del error muestral y reducción de las distorsiones (errores de cobertura y de no respuesta) pueden entrar en conflicto. Si se invierten recursos para reducir los errores no muestrales, habrá que hacer concesiones en cuanto al tamaño de la muestra. Es fundamental alcanzar un equilibrio entre ambos objetivos. De hecho, tendría poco sentido invertir muchas energías para garantizar la ausencia total de errores de no respuesta o de falta de cobertura de una muestra que tuviera solo 50 casos, pero también sería inútil invertir muchos recursos para recopilar, codificar y analizar 20000 cuestionarios sobre una muestra que estuviera claramente distorsionada.

En líneas generales se puede afirmar que, de los dos requisitos, error muestral reducido y ausencia de distorsión, la ausencia de distorsión es más relevante. Por ejemplo, si se quiere realizar una investigación sobre los profesores de una provincia y se cuenta con pocos recursos, es preferible entrevistar a 300 profesores elegidos aleatoriamente de la lista completa ofrecida por el organismo de enseñanza correspondiente, que a 1000 seleccionados con ayuda de conocidos, amigos, cuestionarios entregados en las salas de profesores, etc. En un estudio sobre el cuerpo médico es preferible entrevistar a 300 individuos elegidos aleatoriamente e invertir el máximo esfuerzo en localizarlos hasta que todos hayan realizado el cuestionario, que conseguir 1000 entrevistas enviando un cuestionario postal a 10000 individuos (de los que 9000 no responden). La elección inadecuada de los sujetos, debida al medio utilizado para su identificación y a la autoselección (responden sólo los más motivados), conlleva una distorsión de la muestra que está fuera del control del investigador y que puede influir de forma determinante en los resultados.

Llegados a este punto, abrimos una reflexión sobre los fines de la investigación. El objetivo del estudio es un factor determinante para la estructura de la muestra. En un extremo tenemos los estudios cuyo objetivo es estimar con precisión extrema las medias (o las proporciones) de unas cuantas variables. Un ejemplo de ello sería un pronóstico de los resultados electorales: un error mínimo podría cambiar radicalmente el significado político de las elecciones (por ejemplo, atribuyendo la victoria a un candidato y no a otro). En el otro extremo tenemos los estudios que se centran en el análisis de las relaciones entre las variables. Siguiendo con el ejemplo de las previsiones de resultados electorales, el investigador podría estar interesado en conocer la influencia del género, la edad, la clase social y la religiosidad en el voto. En este caso es irrelevante la precisión de la estimación total (el porcentaje de electores que han votado al partido A), y es muy importante disponer de un gran número de casos para poder fraccionar la muestra y ver, por ejemplo, qué votan los jóvenes-varones-obreros-no religiosos.

En este contexto, puede resultar útil la distinción entre *estudios descriptivos* y *estudios explicativos*. Si el objetivo es de tipo descriptivo, es decir, hacer una descripción lo más exacta posible de la distribución de cada una de las variables, será necesario limitar al máximo las distorsiones en la muestra. Si el objetivo es de tipo explicativo, este requisito no es tan restrictivo. En un estudio sobre la relación entre la edad y el voto, aunque en la muestra los jóvenes estén sobrerrepresentados con

respecto a los ancianos, nada hace pensar que la relación entre las variables se verá alterada por este hecho: si los ancianos votan preferentemente a los partidos conservadores, esta tendencia se mantendrá aunque hayamos incluido en la muestra a menos ancianos de lo debido [20]. No obstante, hay que decir que aunque las investigaciones exclusivamente descriptivas son bastante frecuentes, es mucho más difícil imaginar una investigación que no tenga absolutamente ningún interés en conocer las frecuencias marginales de las variables.

8.9. SÍNTESIS DEL CAPÍTULO 8

1. El muestreo ofrece múltiples ventajas para la investigación social, en cuanto a los *costes* de la investigación, el *tiempo* requerido para la recopilación de datos y la elaboración de los mismos, la *organización,* puesto que reduce el personal necesario para la investigación, la posibilidad de *profundizar*, ya que la menor complejidad organizativa permite concentrar los recursos en asegurar la máxima calidad de los datos.
2. Los estudios que se realizan sobre la población total dan como resultado el valor exacto del parámetro estudiado, en cambio, los estudios realizados sobre muestras sólo nos dan una *estimación* del mismo, es decir, un valor aproximado. Esta estimación implica un cierto *grado de confianza*, y la determinación de un *intervalo de confianza* dentro del cual podría encontrarse el parámetro de la población, con una cierta probabilidad. Por tanto, las estimaciones basadas en muestras están afectadas de un error que se denomina *error de muestreo*. Si la muestra se ha construido observando un procedimiento probabilístico, la magnitud de este error se puede calcular usando la teoría estadística.
3. El *tamaño de la muestra* es directamente proporcional al grado de confianza de la estimación y a la variabilidad del fenómeno estudiado, e inversamente proporcional al error que el investigador está dispuesto a aceptar, o, dicho de otro modo, directamente proporcional a la precisión requerida. El error de muestreo apenas depende de la dimensión de la población, *N*: una muestra de 1000 casos produce estimaciones con errores muy similares, tanto si se han extraído de una población de 20000 sujetos como de una de 100000. En cambio, el tamaño *n* de la muestra tiene una gran influencia sobre el error de muestreo: una muestra de 1000 casos conlleva errores de muestreo muy superiores a los que conlleva una muestra de 2000 casos.
4. Una muestra se denomina *probabilística* cuando todas las unidades que la forman tienen una probabilidad de selección distinta de cero, esta probabilidad es conocida para todas las unidades, y el proceso de selección es aleatorio. Los tipos de muestreos probabilísticos más importantes son: el muestreo aleatorio simple, el muestreo sistemático, el muestreo estratificado, el muestreo a estadios múltiples, el muestreo por conglomerado y el muestreo por áreas.

---

[20] En otras palabras, las frecuencias marginales varían más que las correlaciones.

5. Los principales tipos de muestreo *no probabilístico* son: el muestreo por cuotas, el muestreo de «bola de nieve» y el muestreo subjetivo, o por elección razonada.
6. En la encuesta por muestreo la aplicación de la teoría de la muestra se hace difícil por la existencia de los denominados «errores de selección», que incluyen, además del error de muestreo, otros dos tipos de error: el error de cobertura y el de no respuesta. El *error de cobertura* procede de que en investigación social con frecuencia no se dispone de una lista completa de la población. El *error de no respuesta* puede deberse a dos causas distintas: la falta de contacto con los sujetos seleccionados o el rechazo de los propios sujetos a ser entrevistados.
7. El procedimiento de *ponderación* consiste en una modificación artificial de la composición de la muestra mediante operaciones matemáticas realizadas en la fase de elaboración de los datos, para conseguir que la muestra se parezca más a las características de la población. Los procedimientos de ponderación más comunes se basan en la probabilidad de inclusión de las unidades en la muestra, en los conocimientos disponibles sobre la población y en los conocimientos disponibles sobre las no respuestas.

8.10. LECTURAS COMPLEMENTARIAS

La obra de G. T. Henry, *Practical Sampling* (Sage, 1990, 139 pp.) ofrece un texto introductorio y no demasiado técnico sobre el muestreo; otra síntesis breve y accesible sobre la teoría del muestreo (más centrada en la teoría estadística que la anterior) es el libro de G. Kalton, *Introduction to Survey Sampling* (Sage 1983, 96 pp.).

Una obra bastante completa sobre el muestreo en investigación social, con muchos ejemplos y que ha constituido el texto básico de los especialistas en sondeos durante generaciones, es el clásico de L. Kish, *Survey Sampling* (Wiley, 1965, 643 pp.); otro texto estadístico básico es el de W. G. Cochran, *Sampling Techniques* (Wiley, 1977, 448 pp.). Existe traducción al español de estos dos últimos libros: L. Kish, *Muestreo de Encuestas* (Trillas, 1979, 2ª impresión) y W. G. Cochran, *Técnicas de Muestreo* (Compañía Editorial Continental, 1987, 7ª impresión)

Por último, remitimos al lector a los dos libros de J. Rodríguez Osuna de la serie de *Cuadernos Metodológicos del CIS* (número 1): *Métodos de muestreo.* (Madrid, CIS, 1991, pp. 115), y el número 6: *Métodos de muestreo: casos prácticos.* (Madrid, CIS, 1993, pp. 157). Sobre todo el segundo volumen incluye una descripción pedagógica sobre el proceso de construcción de tres muestras distintas.

LA RECOPILACIÓN DE LOS DATOS: TÉCNICAS CUALITATIVAS

Los tres capítulos siguientes están dedicados a la investigación cualitativa, de la que ya hablamos ampliamente en el capítulo 2. Conviene enfatizar que esta parte de la investigación social utiliza unas técnicas completamente distintas de las cuantitativas que hemos visto en capítulos anteriores, así como una disposición mental y una filosofía radicalmente diferentes. Recordemos que entre estos dos planteamientos hay un salto de paradigma: pasamos del paradigma positivista al paradigma interpretativo.

Desde el punto de vista conceptual y terminológico las técnicas de análisis cualitativo no difieren mucho entre sí: por ejemplo, los términos de investigación etnográfica, investigación de campo, estudios de comunidades, observación participante o investigación naturalista son todos más o menos sinónimos, así como las entrevistas en profundidad, las entrevistas libres, las entrevistas no estructuradas, las entrevistas clínicas, los historiales de vida, el método biográfico, etc. son todas técnicas de recopilación de datos bastante similares.

Las técnicas de análisis cualitativo tampoco presentan grandes diferencias en cuanto a sus aplicaciones, y se suelen emplear de manera simultánea en una misma investigación.

Por último, es difícil dividir el procedimiento de la investigación cualitativa en fases separadas y bien diferenciadas entre sí, ni siquiera los dos momentos clave de la recopilación de los datos y el análisis de los mismos, tan claramente diferenciados en la investigación cuantitativa, tienen un orden invariable en investigación cualitativa, y con frecuencia se alternan y solapan entre sí.

Este hecho ha permitido a Bryman y Burgess [1994, 2] afirmar que en investigación cualitativa hay que hablar, más que de «técnicas» o de «fases», de un «proceso» de investigación: «la investigación cualitativa no puede reducirse a técnicas específicas, ni a una sucesión de estadios, sino que consiste más bien en un proceso dinámico que une problemas, teorías y métodos»; por consiguiente, «el proceso de investigación no es una sucesión bien definida de procedimientos con un diseño claro, sino una interacción imprecisa entre el mundo conceptual y el empírico, donde la deducción y la inducción son simultáneas».

Esta situación se ha acentuado con los últimos avances en investigación cualitativa, que han ampliado el abanico de técnicas empleadas, con la exploración de nuevos enfoques y nuevas perspectivas de investigación, añadiendo a las técnicas clásicas de la observación participante de origen etnográfico y las entrevistas no

estructuradas, toda una «colección variada de material empírico, estudios de casos, experiencias personales, introspección, historiales de vida, textos y productos culturales, textos históricos, material visual, de observación, que describen la rutina, los significados y los momentos clave en la vida de los individuos» [Denzin y Lincoln, 2000: 3].

Es indudable que el proceso de investigación cualitativa carece de la linealidad que suele caracterizar a la investigación cuantitativa. Sin embargo, esto no significa que no podamos delimitar técnicas o grupos de técnicas dentro del mismo. Además, el hecho de que éstas se empleen con frecuencia de manera simultánea no impide estudiar cada una de ellas por separado. Esto es lo que haremos en los tres capítulos que componen esta tercera parte del libro.

Las técnicas de recopilación de datos de la investigación cualitativa se pueden agrupar en tres grandes categorías, basadas respectivamente en la *observación directa,* las *entrevistas en profundidad* y el *uso de documentos,* que pueden remitirse a tres acciones básicas que el hombre utiliza para analizar la realidad social que lo rodea: *observar, preguntar* y *leer.*

Mediante la observación el investigador estudia un determinado fenómeno social, participando primero en él, para vivirlo desde dentro y poder ofrecer una descripción directa del mismo. Con la entrevista, el investigador registra los comportamientos y las motivaciones de los actos mediante la descripción que dan de ellos los propios sujetos, a los que se interroga sobre sus experiencias, sentimientos y opiniones. Por último, el uso de los documentos consiste en analizar una determinada realidad social, a partir del material, por lo general escrito, que la propia sociedad ha producido y produce, a través de los individuos (relatos autobiográficos, cartas, etc.) y de las instituciones (boletines, prensa, actas, fichas, etc.).

Aunque las técnicas de investigación cualitativa se puedan clasificar en estos tres grupos, es mucho más difícil aislar las fases en las que se desarrolla. Por ello, en esta parte del libro no hemos incluido un capítulo dedicado a la presentación del proceso de investigación y sus fases.

Tampoco presentamos una parte dedicada expresamente al análisis de los datos. De hecho, en investigación cualitativa no existe un cuerpo de técnicas formales comparable con el análisis estadístico de los datos propio de la investigación cuantitativa, que tampoco hemos tratado en este libro. Las elaboraciones del material recopilado se hacen según criterios personales, no formalizados o difícilmente formalizables, en los que prevalecen la sensibilidad subjetiva del investigador, su imaginación sociológica, su capacidad personal para relacionar acontecimientos y hacer generalizaciones a partir de la realidad observada, y que no se pueden sintetizar en forma de esquemas aplicables a otros investigadores y otras situaciones de investigación.

Ni siquiera la reciente incursión de la informática en este ámbito, con programas orientados expresamente al análisis de los datos cualitativos, parece haber modificado esta situación ni haber introducido procesos-guías de aplicación general, más allá de sugerir formas estandarizadas de organizar el material empírico y ayudas para identificar las conexiones entre sus distintas partes.

CAPÍTULO 9

LA OBSERVACIÓN PARTICIPANTE

9.1. OBSERVACIÓN Y OBSERVACIÓN PARTICIPANTE 304

9.2. CAMPOS DE APLICACIÓN Y DESARROLLO DE LA OBSERVACIÓN PARTICIPANTE 307

9.3. OBSERVACIÓN DECLARADA Y ENCUBIERTA: EL ACCESO Y LOS INFORMADORES 313

9.4. QUÉ OBSERVAR 318

9.5. REGISTRO DE LA OBSERVACIÓN 323

9.6. ANÁLISIS DEL MATERIAL EMPÍRICO 325

9.7. LA OBSERVACIÓN EN «MICROSOCIOLOGÍA» 331

9.7.1. Del estudio de «los otros» al estudio de «nosotros mismos» 331

9.7.2. La etnometodología 333

9.8. LIMITACIONES Y RECURSOS 337

9.9. SÍNTESIS DEL CAPÍTULO 9 339

9.10. LECTURAS COMPLEMENTARIAS 341

En este capítulo presentamos la técnica de investigación cualitativa más clásica, que traduce el instrumento de recopilación de datos sociales de la *observación* en una inmersión integral en el segmento de la sociedad estudiado. Expondremos una breve historia de este planteamiento de origen etnográfico, las diversas fases de la investigación que utiliza la observación participante, y los problemas prácticos y teóricos que plantea esta técnica de exploración de la realidad social. Así mismo, presentamos los distintos campos de aplicación de esta forma de conocer la sociedad, que si originariamente fue un método para estudiar las sociedades «distintas», ha ido avanzando hasta convertirse también en el principal instrumento para estudiar los actos cotidianos de la vida social.

9.1. OBSERVACIÓN Y OBSERVACIÓN PARTICIPANTE

La distinción entre «observación» y «observación participante» es clara: la primera indica la técnica para la recopilación de datos sobre comportamiento no verbal, mientras que la segunda hace referencia a algo más que una mera observación, e incluye la intervención directa del investigador en el objeto estudiado. En un capítulo sobre «observación» podríamos incluir técnicas como la observación en laboratorio de la interacción entre individuos. Por ejemplo, los experimentos de Bales [1951], que consistían en reunir a varios sujetos en una habitación, hacerlos discutir sobre una determinada cuestión, y observar desde fuera su comportamiento durante la discusión, codificándolo en una serie de acciones básicas. También podríamos incluir técnicas de observación en un entorno natural. Por ejemplo, el estudio de las dinámicas de una clase escolar, en el que la interacción entre profesor y alumnos se graba en vídeo y a continuación se analiza en todos sus aspectos. En estos casos falta un elemento constitutivo de la técnica de observación participante, la intervención del investigador en la situación social estudiada y su interacción con los actores sociales. En otras palabras, mientras que estos ejemplos podrían clasificarse dentro del enfoque positivista, la observación participante queda dentro del *paradigma interpretativo*.

Entonces, ¿por qué «observación» y por qué «participante»? *Observación,* puesto que implica mirar y escuchar. Pero al mismo tiempo esta técnica conlleva un contacto personal e intenso entre el sujeto que estudia y el sujeto estudiado, una larga interacción que puede durar incluso años, con una *participación* del investigador en la situación objeto de estudio, que constituye su elemento distintivo. El investigador *observa* y *participa* en la vida de los sujetos estudiados. Pero es el segundo aspecto el que hace que esta técnica sea diferente de todas las demás técnicas de investigación social.

Ni en las entrevistas estructuradas, ni en el análisis de las fuentes estadísticas, ni en el experimento, ni en el análisis de documentos, ni siquiera en las entrevistas en profundidad, el investigador participa en el fenómeno social estudiado. En la observación participante el investigador «baja al campo», se adentra en el contexto social que quiere estudiar, vive *como* y *con* las personas objeto del estudio, comparte con ellas la cotidianidad, les pregunta, descubre sus preocupaciones y sus

esperanzas, sus concepciones del mundo y sus motivaciones al actuar, con el fin de desarrollar esa «visión desde dentro» tan importante para la *comprensión*.

Son dos, por tanto, los principios de fondo de este enfoque: *a*) que el conocimiento social pleno sólo se puede alcanzar mediante la comprensión del punto de vista de los actores sociales, mediante un proceso de identificación con sus vidas; y *b*) que esta identificación sólo se puede realizar mediante la integración en su cotidianidad, con una interacción continua y directa con los sujetos estudiados.

Por consiguiente, podemos definir la observación participante como *una técnica en la que el investigador se adentra en un grupo social determinado: a) de forma directa; b) durante un periodo de tiempo relativamente largo; c) en su medio natural; d) estableciendo una interacción personal con sus miembros; y, e) para describir sus acciones y comprender, mediante un proceso de identificación, sus motivaciones.*

Repasemos los distintos elementos de esta definición: *a*) la observación la debe realizar el investigador en primera persona, no puede delegar en otras personas la recopilación de los datos (como ocurre, por ejemplo, en las entrevistas de la encuesta por muestreo); *b*) el periodo de participación en el grupo estudiado debe ser relativamente largo (desde varios meses hasta varios años); *c*) esta participación tiene lugar en el hábitat natural del grupo y no en un ambiente artificial reconstruido para facilitar el trabajo de investigación; *d*) el investigador no se limita a observar desde fuera, sino que interactúa con las personas a quienes estudia; y *e*) la finalidad es describir y «comprender» en el sentido weberiano, es decir, lograr «ver el mundo con los ojos de los sujetos estudiados».

Desde esta perspectiva, el investigador no debe temer contaminar los datos mediante un proceso de interpretación subjetiva y personal, ya que la subjetividad de la interacción y la interpretación, es precisamente una de las características de la técnica; por tanto, la *implicación* y la *identificación* no deben evitarse, sino que deben buscarse, mientras que la *objetividad* y la *distancia*, que eran las premisas de los planteamientos neopositivistas, pierden su valor.

En este proceso de participación, el investigador debe mantener el equilibrio entre dos casos extremos, el del «marciano» y el del «converso» Davis [1973]. Ambos parten de la premisa de que la sociedad humana es sobre todo un constructo social, pero llegan a conclusiones muy distintas. El «marciano» trata de participar lo menos posible en la situación social estudiada, contempla con desconfianza lo que los miembros de la sociedad relatan sobre sí mismos y su mundo, considera estos relatos desorientadores porque están al servicio de una visión y una cultura determinadas, y opina que la tarea del investigador social es librarse de todas las premisas cognitivas y culturales (incluso las propias) para captar la esencia real de los procesos sociales.

El «converso», en cambio, no teme que la cultura de la sociedad estudiada sea una «trampa cognitiva», sino que piensa que sólo la inmersión total en ella puede proporcionar los instrumentos necesarios para su comprensión. Con la *inmersión total*, que se entiende como «una identificación total con los esquemas de vida de los demás» [*ibídem*, 338], el «converso» desencadena una experiencia existencial parecida a la de un individuo que se convierte a otra religión, que ve en la conver-

sión una transformación radical de su identidad.

Probablemente la solución al dilema *participación/distancia* está en el equilibrio entre estas dos posturas: si una excesiva distancia impide la comprensión, también la identificación total puede ser un obstáculo, puesto que, como veremos, la peculiaridad del investigador social radica en saber plantear preguntas sobre la situación estudiada que surgen de su cultura y su experiencia.

La técnica de la observación participante nace en el contexto de la investigación antropológica a principios del siglo XX, con el fin de estudiar a las sociedades primitivas. Con la desaparición de estas últimas, la antropología desplazó sus intereses hacia el estudio de las sociedades modernas, centrando su atención en áreas sociales o culturales específicas (antropología urbana, antropología de la medicina, etc.), y al mismo tiempo, su modelo de investigación fue adoptado por otras disciplinas, como la sociología. En particular, en el terreno sociológico, hay que recordar las notables aportaciones de la Escuela de Chicago, que, bajo la dirección de Robert Ezra Park y desde el Departamento de Sociología de la Universidad de Chicago, realizó en los años veinte y treinta una serie de estudios sobre la sociedad urbana americana, dedicando especial atención a los fenómenos de desviación y marginalidad social, que aún hoy representan un punto de referencia clave para el desarrollo de la investigación sociológica. La intención explícita de Park era aplicar al estudio de las culturas urbanas el mismo criterio que antropólogos como Francis Boas y Robert Lowie habían aplicado al estudio de los indios de América.

Resumen 9.1. LA OBSERVACIÓN PARTICIPANTE EN LA INVESTIGACIÓN ETNOGRÁFICA.

Fue el gran antropólogo inglés de origen polaco Bronislaw Malinowski quien formuló en la introducción de su libro *Argonautas del Pacífico Occidental*, que se considera la génesis de la antropología moderna, los principios de la observación participante. Malinowski aplicó estos principios en las investigaciones que realizó entre 1914 y 1920 sobre las sociedades primitivas de las islas de Melanesia. Sus planteamientos supusieron la crisis definitiva del modelo tradicional de la antropología del siglo XIX, que consideraba a los nativos «salvajes primitivos» a los que hay que educar según las normas de la civilización occidental, y cuyas investigaciones consistían en informes de las oficinas gubernamentales de las colonias o las misiones. Malinowski propuso el objetivo «de entender el punto de vista del indígena, su relación con la vida, [para] conocer *su* visión de *su* mundo» [Malinowski, 1922]. Él realizó sus investigaciones en las islas Trobriand, unos 200 km al noreste de Nueva Guinea, fuera de las rutas comerciales de la época. Su método consistía en pasar largos periodos viviendo dentro de las sociedades primitivas, evitando todo contacto con el mundo occidental durante el periodo de observación.

Estas experiencias permitieron enriquecer y desarrollar el modelo inicial de Malinowski, aunque sin alterar las características de fondo: convivencia del estudioso con los sujetos estudiados en el lugar de residencia de éstos, observación en el entorno natural de la interacción social y uso de personajes-clave como informadores.

En el contexto de la investigación de tipo antropológico-etnográfico, también llamado *estudio de campo*, la observación participante no es el único instrumento usado por el investigador. El observador participante debe observar, escuchar y preguntar, y al preguntar, los instrumentos que emplea son los de la entrevista. Al mismo tiempo, debe documentarse sobre los hechos que suceden y los sucedidos en el pasado, examinar el material documental existente sobre la comunidad que estudia y el que ésta produce (relatos históricos, autobiografías, cartas, actas, informes, artículos de periódico, etc.), y para hacerlo utiliza los instrumentos del análisis documental. En los capítulos siguientes veremos las técnicas de la entrevista y el análisis documental, que se usan también de manera autónoma en investigación social. No obstante, conviene indicar que, en todo caso, la observación participante constituye el instrumento principal de este tipo de investigación.

En las páginas que siguen presentaremos los rasgos fundamentales de la observación participante. Hay que mencionar que esta técnica no sigue un recorrido lineal, constituido por fases que se suceden de forma preestablecida. La observación participante es una técnica difícil de codificar y explicar, ya que se trata de una experiencia más que un conjunto de procedimientos coordinados, cuyo progreso depende básicamente de la compleja interacción que se establece entre el problema estudiado, el sujeto que estudia y los sujetos estudiados.

9.2. CAMPOS DE APLICACIÓN Y DESARROLLO DE LA OBSERVACIÓN PARTICIPANTE

La observación participante puede aplicarse al estudio de todas las actividades y agrupamientos de los seres humanos, sobre todo cuando se quiere descubrir «desde dentro» la visión de su mundo. Pero en algunos casos, esta técnica resulta especialmente útil, en función de los sujetos estudiados. Jorgensen [1989, 12] enumera cuatro:

- Cuando se sabe poco de un determinado fenómeno (un nuevo movimiento político; un acontecimiento social imprevisto, como una rebelión, etc.).
- Cuando existen grandes diferencias entre lo percibido, el punto de vista interno y el punto de vista externo (grupos étnicos, organizaciones sindicales, grupos profesionales, como médicos, abogados, etc.).
- Cuando el fenómeno no admite la presencia de miradas de extraños (rituales religiosos, vida familiar, relación entre médico y paciente, etc.).
- Cuando el fenómeno se oculta de manera deliberada a las miradas de extraños (conductas delictivas o desviadas, asociaciones secretas, sectas religiosas, etc.).

Hablamos de investigación mediante observación participante cuando el investigador pretende estudiar una realidad de la que él mismo ha formado o forma parte, dando lugar a lo que también se ha denominado *sociología autobiográfica.* De la extensa literatura existente sobre el tema podemos citar la investigación sobre los jugadores de azar de Scott [1968], asiduo durante mucho tiempo a timbas clandestinas, la del músico Becker [1963] sobre los músicos de jazz, la de Hall [1976], alumna de escuelas de baile durante 16 años, sobre la socialización de las niñas en la profesión de bailarina, la de Mary Romero [1992], que desde los 15 años trabajó como empleada del hogar junto a su madre y sus hermanas en una familia de El Paso, sobre las mujeres mexicanas que cada semana atraviesan ilegalmente la frontera entre Estados Unidos y México para trabajar en casas de familias estadounidenses, las investigaciones de Irwin [1970; 1980] sobre las cárceles americanas, tras su experiencia carcelaria de cinco años que empezó cuando tenía doce, la investigación de Jacqueline Wiseman [1979], propietaria de una tienda de ropa de segunda mano, sobre el mundo que rodea a este peculiar tipo de negocio.

También podemos incluir en esta categoría de sociología autobiográfica los estudios realizados a partir de una experiencia vivida transformada en objeto de investigación. Whyte [1984, 31], por ejemplo, cita el caso del estudiante José Moreno, enviado por la Universidad de Cornell a la República Dominicana para estudiar la vida en los pueblos, y que en cuanto llegó a la isla se encontró con una rebelión. Moreno simpatizó con los sublevados, se unió a ellos y después de su experiencia publicó un libro basado en el material recopilado durante la misma [Moreno, 1970].

Cabe decir que la experiencia autobiográfica que se convierte en investigación no es el caso ideal de observación participante, y que está expuesta a no pocas críticas. La investigación de planteamiento etnográfico pretende ser un *encuentro de culturas.* Precisamente esta capacidad de la cultura que estudia para mirar a la cultura estudiada con los ojos ingenuos del extranjero que sabe poco o nada y se adentra en un mundo nuevo sin conocer sus costumbres, lenguaje, normas y valores, permite al que observa descubrir aspectos y significados que de otro modo pueden quedar ocultos a quien ha formado parte de ese mundo durante toda su vida.

Hemos dicho que la observación participante, siguiendo la estela de la etnografía, es una técnica de investigación que se utiliza sobre todo para estudiar las culturas. En la investigación sociológica se ha aplicado fundamentalmente con dos objetivos: estudiar todos los aspectos de la vida de microcosmos sociales autónomos ubicados en territorios bien delimitados y con un universo cultural propio y cerrado, por ejemplo, una comunidad campesina, una pequeña ciudad de provincias, un pueblo de mineros, etc.; o estudiar subculturas que se han desarrollado dentro de segmentos sociales de las sociedades complejas, que pueden representar aspectos de la cultura dominante (la cultura de los jóvenes, los ricos, los abogados, los empleados de un gran complejo industrial, los militares, un partido político, los aficionados al fútbol, etc.), o pueden estar en conflicto con ella (una secta religiosa, un partido revolucionario, jugadores de azar, minorías étnicas, etc.), o enfrentadas a ella (grupos terroristas, presos, movimientos políti-

cos radicales, activistas, etc.). Los estudios del primer tipo se denominan *estudios de comunidades;* los del segundo tipo, *estudios de subculturas.*

Los *estudios de comunidades* son los que más se acercan al modelo etnográfico. Se trata de investigaciones realizadas sobre comunidades pequeñas (o relativamente pequeñas), ubicadas en un territorio concreto, que requieren el traslado del investigador a la comunidad estudiada, donde vivirá durante un periodo de tiempo determinado. Un estudio clásico en este campo es la investigación del matrimonio Lynd, que en 1924 se trasladó a una pequeña ciudad de provincias de Estados Unidos (Muncie, en el estado de Indiana) donde residió durante tres años y medio, con el fin de estudiar «las interacciones de la vida de una pequeña ciudad estadounidense [...] según el planteamiento de investigación de la antropología cultural. Después de todo, pese de las infinitas variaciones en los detalles, la gente no hace cosas tan distintas. Ya sea en un poblado Arunta de Australia central o en nuestra vida burocrática aparentemente tan complicada [...], el comportamiento humano parece consistir en variaciones de unas cuantas actividades básicas [...]. Este estudio parte, por tanto, de la tesis de que todas las cosas que hacen los individuos en esta ciudad estadounidense pueden enmarcarse en alguna de las siguientes actividades principales: ganarse el sustento, formar una familia, educar a los hijos, ocupar el tiempo libre [...], participar en prácticas religiosas, y participar en actividades públicas» [Lynd y Lynd, 1929, 3-4].

Como puede verse, y como subrayan los autores, se estudian todos los aspectos de la vida de la sociedad, empleando un criterio totalmente análogo al del antropólogo que estudia una tribu primitiva.

Un estudio de comunidades muy conocido es el que el investigador norteamericano Edward Banfield realizó en Italia en los años cincuenta: Bandfield residió durante nueve meses con su mujer y sus dos hijos en un pueblo de la provincia de Potenza, con tres mil habitantes, compuesto en su mayoría por campesinos pobres y jornaleros, para realizar «un estudio de las condiciones culturales, psicológicas y morales que forman la base de una organización política» [Banfield, 1958; 8 y 9]. Aunque utilizó el mismo método que el matrimonio Lynd, el autor se centró en un aspecto específico de la vida social, el de la participación política y social, poniendo de manifiesto una cultura muy individualista, centrada en la familia, que condicionaba toda la vida social y económica del pequeño pueblo, hasta el punto de que la «extrema pobreza y el atraso [podían] explicarse en gran medida [...] por la incapacidad de los habitantes de actuar juntos por el bien común» [*ibídem*].

Pasando a los *estudios de subculturas,* podemos decir que al principio la atención de los sociólogos se dirigió sobre todo a las culturas distintas de la cultura dominante. Los estudios de la Escuela de Chicago de los años veinte y treinta tuvieron como objeto fundamental el fenómeno de la marginalidad social producida por la inmigración y la urbanización: los vagabundos, las bandas, las prostitutas, los guetos negros, los suburbios, las comunidades de inmigrantes, etc.

Una de las investigaciones más conocidas, que se cita a menudo como prototipo de este tipo de estudios, es *Street Corner Society,* realizada por William Foote Whyte en 1936-37 en un barrio marginal de las afueras de Boston. Whyte, que entonces tenía 23 años y estudiaba en la Universidad de Harvard, decidió dedicar su tesis

doctoral al estudio de la criminalidad organizada a pequeña escala, por lo que localizó en su ciudad el barrio que le pareció más degradado y se fue a vivir allí como un habitante cualquiera. Para él fue crucial el encuentro con Doc, el líder de una de las bandas juveniles de la zona, que lo introdujo en todos los ambientes del barrio y en particular en el de las bandas. Su investigación está descrita en detalle en el apéndice de la segunda edición de su obra (publicada en 1955), que se convirtió pronto en una especie de guía metodológica para la observación participante.

Otro ejemplo es la investigación sobre bandas de Jankowski que citamos en el capítulo 2. Como ya mencionamos, esta investigación se basa en la observación participante de un número considerable de casos (37 bandas de tres ciudades distintas observadas durante un periodo de diez años), lo que es bastante inusual en esta técnica, que suele centrarse en el estudio de una sola situación social.

Otro ejemplo clásico es *The Hobo* (El vagabundo), de Nels Anderson, que inauguró en 1923 la *Sociological Series* del departamento de Sociología de la Universidad de Chicago, dedicada a los problemas urbanos de esa ciudad. Esta investigación estudia el proletariado marginal que había crecido en aquellos años en torno a la construcción de la gran red ferroviaria de Estados Unidos: los *hobos* eran trabajadores inmigrantes, sin casa fija, que trabajaban como peones sobre todo en el sector de la construcción y durante periodos de tiempo limitados (por lo general, el necesario para realizar un trabajo determinado), que se desplazaban al lugar donde hubiera trabajo, individuos pertenecientes a un mundo de connotaciones culturales bien definidas, donde la movilidad, el viaje y la sensación de provisionalidad eran rasgos existenciales. La investigación de Anderson es en parte producto de la propia biografía del autor (su padre fue un *hobo,* y él también conoció ese mundo durante los primeros años de su vida laboral), y en parte fruto de una intervención de observación participante, que se produjo cuando el autor eligió como tema de su tesis doctoral la vida de los *hobos,* para lo cual se trasladó al barrio de Chicago en el que se concentraba el mayor número de personas sin vivienda fija.

A partir de los primeros estudios sobre grupos marginales, se va recurriendo cada vez más a la observación para estudiar también la cultura de sectores sociales que integran la sociedad «normal». Son muy comunes, por ejemplo, los estudios sobre la vida y la condición social de ciertos grupos de trabajadores, para los cuales el investigador suele ser contratado para un determinado puesto de trabajo, a fin de describir su estilo y condiciones de vida y su visión del mundo.

Una investigación clásica de este tipo fue realizada en los años cuarenta por Donald Roy [1952], un joven investigador norteamericano que estudió el trabajo de la clase obrera, y en concreto intentó dilucidar la cuestión de la autolimitación en la productividad: los motivos que llevaban a los obreros de una fábrica determinada a limitar de manera consciente su productividad en las máquinas, ignorando incluso incentivos económicos que los premiaban por cada pieza adicional que produjeran. Para comprender este problema, y más en general para conocer la cultura obrera, Roy, entonces un joven estudiante de doctorado de la Universidad de Chicago, consiguió que lo contrataran durante diez meses (entre 1944 y 1945) en un taller que construía piezas de vagones ferroviarios. Lo hizo sin desvelar sus intenciones a nadie, ni al empleador cuando lo contrató, ni a sus compañeros de trabajo.

Roy rechazaba tanto el punto de vista radical de inspiración marxista, según el cual la autolimitación obrera (la tendencia a producir por debajo de sus posibilidades) era una expresión de la conciencia de clase y el efecto del inevitable conflicto entre capital y trabajo, como la interpretación funcionalista que atribuía las limitaciones de productividad a una falta de comunicación entre los obreros y la dirección, a una atención inadecuada a los aspectos humanos del trabajo. Roy constató que este comportamiento de autolimitación era un acto racional de los obreros, que se guiaban por consideraciones de lúcida conveniencia: o porque una mayor productividad habría inducido inmediatamente a la dirección a elevar el nivel de producción mínimo, o porque el incentivo económico era demasiado bajo como para merecer un esfuerzo mayor.

Para citar otro estudio de este tipo, podemos mencionar que, exactamente treinta años después, otro investigador estadounidense, también de la Universidad de Chicago, Michael Burawoy [1979], logró que lo contrataran en la misma fábrica para realizar una investigación similar, centrada sobre todo en el estudio de las transformaciones que había sufrido el trabajo obrero en los treinta años transcurridos desde la primera investigación. Cabe destacar una curiosidad: aunque el autor del segundo estudio utilizó también el instrumento de la observación participante, el hizo una observación abierta, revelando sus objetivos a la dirección de la empresa y a sus compañeros de trabajo. Este cambio fue posible gracias a la menor la tensión entre la dirección y la mano de obra en la fábrica que se había conseguido en los treinta años transcurridos, y al gracias también a un mejor conocimiento de los instrumentos de investigación social por parte de la dirección de la empresa.

En el proceso antes mencionado de desplazamiento del foco de la observación participante de los grupos marginales a los «normales», podemos citar las investigaciones orientadas a estudiar los valores, la red de relaciones sociales y las dinámicas interpersonales que se desarrollan dentro de instituciones y organizaciones sociales. Un ejemplo de ello es la corriente de estudios de la denominada *etnografía organizativa,* cuyo objeto de estudio es la cultura de las organizaciones, (los conocimientos implícitos comunes a los miembros de la institución, los modelos de referencia empleados para interpretar la realidad, las reglas no escritas que orientan los actos individuales), y la forma en que esta cultura está presente en las acciones e interacciones sociales (la organización formal e informal, los procesos de toma de decisiones, las relaciones interpersonales, los símbolos y ritos, etc.).

En el ámbito de las organizaciones productivas podemos mencionar los estudios realizados por Gouldner entre 1948 y 1951 y publicados en 1954 en las obras *Modelos de burocracia empresarial* y *La huelga del gato salvaje*. Las dos investigaciones son el resultado de una observación participante realizada durante tres años, entre 1948 y 1951, en un yacimiento minero estadounidense con unos 200 trabajadores, una parte de los cuales trabajaba en la extracción y la otra en la transformación del mineral. La recopilación de datos fue realizada por un grupo de investigadores constituido mayoritariamente por estudiantes universitarios, cuyo campus se encontraba a poca distancia del yacimiento. Los estudiantes tenían acceso libre a la fábrica y a las minas, y su presencia se hizo pronto familiar para obreros y mineros. El material empírico de la investigación es el producto de la observa-

ción y el análisis de los documentos empresariales, así como de las entrevistas no dirigidas realizadas por los estudiantes, previa adecuada formación y constante supervisión de su trabajo. Como puede verse, este modelo organizativo difiere en parte del modelo de observación partieipante clásico, tanto por la presencia de un equipo real de investigadores, como por la relevancia de la parte de las entrevistas.

En el primer libro, el análisis de Gouldner se centra en la organización empresarial, a partir del tipo ideal weberiano de burocracia, dentro del cual intenta identificar subtipos; por otra parte, el autor estudió también la cuestión de la sucesión en la burocracia, lo cual fue posible gracias al hecho de que la investigación se realizó cuando se produjo el relevo en la dirección del yacimiento, para pasar de un director mayor paternalista y permisivo, a un joven dirigente inspirado en un modelo más eficaz y ejecutivo. En su segundo libro, Gouldner parte de una afortunada coincidencia: durante su estancia en el yacimiento se produjo un hecho imprevisto, una huelga espontánea, que le permitió observar en directo un episodio social bastante inusual y convertir la ocasión en la base de observación para el estudio y la elaboración de una teoría general sobre las tensiones de grupo.

Otro sector de la sociedad actual que ha sido estudiado mediante observación participante son las instituciones sanitarias. Por ejemplo, Atkinson [1995] quiso estudiar la cultura de los médicos y las formas de definición de la enfermedad en la situación de especialización y subdivisión extrema de la medicina moderna. Para ello, se incorporó a un departamento de hematología-oncología de un hospital de Boston, y durante diez semanas siguió de cerca el trabajo de tres médicos, participando en todas sus reuniones y grabando sus conversaciones: desde la reunión matutina diaria del equipo para tratar las condiciones de los pacientes de la unidad, hasta las discusiones sobre los casos, los contactos con la generación más joven de los médicos que se estaban especializando y las reuniones con el grupo de médicos más veteranos y experimentados, observando tanto las discusiones informales como las reuniones formales. Todo ello con el objetivo de estudiar las formas de interacción entre los médicos, cómo describen éstos sus casos clínicos a sus compañeros, cómo intentan convencerse entre sí sobre sus respectivos diagnósticos, el proceso de formación de las nuevas generaciones, en resumen, cómo se produce y se reproduce la cultura médica en un sector muy especializado y avanzado tecnológicamente.

La observación participante se ha utilizado también para el estudio de las instituciones políticas. La investigación de Dahl, Polsby y Wolfinger sobre democracia y poder en una ciudad estadounidense (en concreto, New Haven, en Connecticut), es un clásico de estos estudios. El objetivo de la investigación fue estudiar la cuestión de «quién gobierna» la ciudad, en todos sus aspectos: los lugares efectivos del poder, la dinámica de toma de decisiones, el origen social y los itinerarios de selección de los líderes más influyentes, el papel de los grupos de presión, los mecanismos de negociación entre intereses opuestos, la relación entre representantes y representados, etc. [Dahl, 1961; Polsby, 1963; Wolfinger, 1973]. El estudio es complejo y para el mismo se utilizaron diversas técnicas de recopilación y análisis de datos, entre ellas la observación participante. Wolfinger obtuvo la autorización del alcalde para pasar un tiempo en su despacho, observando el

desarrollo de sus jornadas laborales, las personas con las que mantenía contacto, los filtros utilizados para seleccionar estos contactos, las dinámicas de las interacciones, el papel de los colaboradores y consejeros, etc., y con ello realizó una reconstrucción detallada de una serie de procesos de toma de decisión.

Los estudios sobre la «cultura de los niños» son una de las aplicaciones más originales de la observación participante al estudio de contextos culturales de sectores específicos de la sociedad. Podemos recordar las investigaciones realizadas en Estados Unidos e Italia por el investigador estadounidense William Corsaro [Corsaro y Rizzo, 1990], para las cuales el observador se integra en una comunidad infantil (Corsaro ha trabajado sobre todo en escuelas infantiles, con niños de entre tres y cinco años). El trabajo de campo se compone de observación (usando cámaras de vídeo para grabar comportamientos y diálogos), e interacción con los niños. Esta interacción se realiza mediante la integración del investigador en la clase, para observar a los niños, interactuar con ellos y participar en los juegos, tratando de adquirir el estatus de miembro del grupo para captar la visión y las interpretaciones de los niños, igual que el investigador que observa y participa en la vida de una comunidad de adultos. También en este caso, como en la observación participante clásica, se atribuye gran importancia a las interpretaciones y explicaciones dadas por los sujetos y observadas durante la conversación aparentemente casual con el observador, que mientras juega con los niños les pregunta por el significado de los comportamientos, y el análisis de las conversaciones entre los propios niños.

Estos últimos ejemplos sobre el estudio de subculturas que forman parte de la cultura de una sociedad compleja no nos deben hacer olvidar la utilidad de la observación participante para el estudio de «culturas específicas», universos culturales autónomos y a menudo muy divergentes de la cultura dominante: la observación participante sigue siendo la técnica ideal para el estudio de la desviación, la marginalidad social, las minorías étnicas, las sectas religiosas, las organizaciones cerradas y los grupos «alternativos» en todos los sentidos.

9.3. OBSERVACIÓN DECLARADA Y ENCUBIERTA: EL ACCESO Y LOS INFORMADORES

Al hablar de observación participante, es importante distinguir según si el papel del observador se hace explícito o no. Éste puede dar a conocer u ocultar sus objetivos reales: puede declarar previamente que es un investigador que quiere formar parte de un determinado grupo social, no para compartir sus objetivos, sino para estudiarlo, o puede integrarse en la situación social en cuestión, fingiendo formar parte de la misma como un miembro cualquiera.

La justificación principal para defender la observación encubierta es el hecho de que, cuando se sabe observado, el ser humano, se comporta presumiblemente de forma distinta a la habitual. Esto es lo que Labov [1972] ha llamado la paradoja del observador: nosotros queremos observar cómo se comportan las personas cuando no están siendo observadas. Y la observación encubierta permite captar de forma

más genuina el modo de actuar natural. Ya hemos subrayado varias veces que la cuestión de la «reactividad» del ser humano es uno de las diferencias fundamentales entre la investigación en las ciencias sociales y la investigación en las ciencias naturales.

Sin embargo, este procedimiento presenta varios inconvenientes. El primero es de carácter moral. Presentar una identidad diferente de la propia, asumir un papel que podría compararse con el de un «espía», es un hecho reprobable en sí mismo, que sólo resulta aceptable si se apoya en motivaciones éticas convincentes. Y es discutible que el objetivo de la investigación social tenga un valor moral tan alto como para justificar medios como el engaño y la instrumentalización de la buena fe de los demás. Además, saber que está engañando a otros puede generar un estado de incomodidad y falta de naturalidad en el investigador difícil de gestionar: incluso para personas con pocos escrúpulos puede resultar difícil profundizar en la relación interpersonal cuando están fingiendo. Además siempre existe el riesgo de ser descubiertos, lo que puede tener consecuencias totalmente imprevisibles (desde la interrupción brusca de la relación, hasta hechos incluso más graves cuando se trata de grupos desviados o que realizan actividades ilegales).

En ciertos casos la ocultación del papel del investigador puede ser un obstáculo para el objetivo de la observación, que consiste precisamente en comprender al grupo que se estudia: las entrevistas explícitas o las preguntas demasiado insistentes, cuyo fin es entender el punto de vista del actor social, pueden resultar una empresa imposible si el observador no manifiesta abiertamente su función y sus objetivos. Además, el observador participante reconocido como tal puede aprovechar su «incompetencia» declarada para formular las preguntas más ingenuas, para pedir explicaciones sobre los hechos más banales, lo que le permitirá recoger testimonios sobre el punto de vista de los miembros del grupo y su definición de las situaciones sociales.

A todo ello podemos añadir una observación citada por muchos investigadores con larga experiencia: el argumento que se suele usar para criticar la observación declarada (cuando uno sabe que lo están observando se comporta de forma distinta) es válido para las primeras fases de la observación, pero a medida que el observador es aceptado en el ambiente estudiado y su presencia se convierte en habitual, las desconfianzas se atenúan y el comportamiento de los «observados» vuelve a ser normal.

A veces no se plantea el problema de si el papel del observador se debe hacer explícito o no. Cuando el ambiente observado es público, abierto, no es necesario declarar que se va a hacer una observación, y en ciertas ocasiones podría ser sería imposible y ridículo. Pensemos, por ejemplo, en estudios sobre el comportamiento de la multitud en un estadio, sobre el cortejo entre jóvenes en un local nocturno, sobre la dinámica de las intervenciones en una asamblea política o sindical. Tampoco se plantea el problema, o éste es menos evidente, cuando el investigador estudia situaciones en las que ya está integrado de manera natural, por ejemplo, un profesor que estudia las dinámicas de la clase o un trabajador social que estudia el problema de la dependencia de las drogas en un centro de rehabilitación de toxicómanos.

Otras veces la ocultación del observador no es un engaño abierto, sino una

omisión, es decir, simplemente no se explicita su función. Pensemos, por ejemplo, en el investigador que consigue que lo contraten en una fábrica para estudiar las condiciones de vida de los obreros, o bien en el periodista que, aprovechando su fisonomía de rasgos árabes, se unió a un grupo de inmigrantes ilegales procedentes del norte de África para estudiar la inmigración y el trabajo sumergido en Italia. O incluso en los estudios de comunidades en los que el investigador se traslada a una ciudad o un pueblo para describir su vida. En estos casos suele suceder que el investigador combina una actividad normal de trabajo con su estudio, y no es necesario que explicite sus intenciones reales, a menos que quiera ir más allá de las relaciones normales (de trabajo, de vecindad, etc.), por ejemplo, realizar entrevistas en profundidad a observadores privilegiados a quienes sí revelará sus objetivos. Para estos casos en que la función del observador sólo se revela a ciertos miembros de la comunidad observada, se ha propuesto también el término de rol «semiencubierto» ([Whyte 1984, 30]).

En general, cuando el grupo estudiado es un grupo privado, no relacionado con el investigador y accesible sólo para individuos que cumplen determinados requisitos, es muy recomendable que el observador explicite las motivaciones de su participación. Como escribe Whyte [*ibidem*, 31], «en un estudio de comunidades se suele descartar la ocultación de la propia función; los miembros de la comunidad no estarán dispuestos a conceder entrevistas ni a expresar sus puntos de vista sin una motivación razonable». Para el resto de situaciones no es posible dar una indicación general, y la elección entre explicitación y encubrimiento deberá ser considerada en cada caso concreto.

Una vez elegido el caso a estudiar y establecida la modalidad de observación (encubierta o declarada), la primera cuestión que debe abordar el investigador es el acceso al entorno estudiado. Esta cuestión representa probablemente una de las fases más difíciles de la observación participante. Patton [1990, 250] dice que los informes de los investigadores a este respecto le recuerdan el relato de Kafka titulado *El castillo*, cuando el protagonista, al que se nombra sólo con la inicial K., sin más señas de identidad, merodea en torno al castillo: quiere formar parte de ese mundo, pero cualquier esfuerzo por entablar contacto con las autoridades invisibles que controlan su acceso le provoca frustración y ansiedad. Supone que existen unas reglas para acceder al mismo, pero él no las encuentra. Duda de sí mismo, entra en crisis y se siente culpable por su incapacidad.

El protagonista de la obra de Kafka no consigue entrar en el castillo. El observador participante suele conseguir acceder al campo estudiado, pero no le resulta sencillo. La forma más común de resolver el problema es la intervención de un *mediador cultural*. Esta táctica consiste en recurrir a la credibilidad y el prestigio de uno de los miembros del grupo para legitimar al observador y hacer que el grupo le acepte. El mediador cultural es una persona que goza de la confianza de la población estudiada y que al mismo tiempo, por sus características culturales, es capaz de comprender las motivaciones y las exigencias del etnógrafo.

En su exposición sobre diez experiencias de acceso diferentes extraídas de otras tantas investigaciones, Whyte [1984, 37 ss.] distingue los grupos informales de las organizaciones formales. El ejemplo clásico de los grupos informales lo constituye

la investigación ya mencionada del mismo Whyte, *Street Corner Society* (1943). En el apéndice a su investigación Whyte relata los esfuerzos iniciales que tuvo que hacer para entablar contacto con los habitantes de Cornerville (el nombre ficticio del barrio que estudió).

Empezó trabajando para una agencia que se dedicaba al mantenimiento de casas, a fin de tener una excusa para llamar a las puertas, entrar en las viviendas y hacer preguntas sobre las condiciones de vida de las personas que las habitaban. Sin embargo, el peso psicológico de esta intrusión y la innaturalidad de la situación lo obligaron a interrumpir pronto este experimento. A continuación pensó en acercarse a alguna chica en uno de los bares frecuentados por los jóvenes del barrio, con la esperanza de poder incitarla a contarle su vida invitándola a tomar algo. Pero el resultado fue desastroso. Se sucedieron otras experiencias frustrantes, hasta que un asistente social del barrio le sugirió que se pusiera en contacto con Doc, un joven inteligente que había trabajado durante un tiempo para los servicios sociales en barrios desfavorecidos. Según el propio Whyte, la noche en que encontró a Doc, supuso en cierto sentido el inicio de su investigación. Doc entendió las intenciones de Whyte y se ofreció a llevarlo por varias zonas del barrio presentándolo como un amigo suyo. Doc estaba tan bien considerado en el barrio que estas presentaciones bastaron para abrirle las puertas del mismo a Whyte y asegurarle su aceptación, en particular entre los jóvenes que eran el objeto más importante de su observación.

Estas experiencias sirven como ejemplo para ilustrar los procedimientos de acceso a un grupo informal. Cuando el grupo social estudiado es una institución, la situación es distinta. A veces existen reglas formales para acceder al grupo; otras veces es necesario solicitar la autorización de las personas encargadas del control del acceso. Un caso clásico es el del investigador que consigue que lo contraten como trabajador de una empresa para estudiar el mundo laboral. Whyte [1984] cita el caso de Robert Cole, un investigador estadounidense que en los años setenta se propuso estudiar la cultura laboral japonesa, tratando de aplicar el modelo weberiano de la racionalidad burocrática. Para conseguir acceder al entorno estudiado Cole utilizó tanto la vía de acceso formal (en este caso, respondiendo a ofertas de trabajo publicadas en los periódicos), como la vía de obtención de autorización (para lo cual expuso sus objetivos de investigación a la dirección de personal de algunas empresas y solicitó que lo admitieran como empleado durante un tiempo).

Ambas vías resultaron difíciles. La primera porque en el mundo laboral japonés una solicitud de trabajo de un occidental era algo totalmente inusual, salvo en el caso de trabajos muy especializados. La segunda, porque en aquella en Japón se desconocían completamente las investigaciones de este tipo, por lo que su solicitud no se entendía o, peor aun, se malinterpretaba. Finalmente consiguió encontrar trabajo en una pequeña fábrica. También en este caso la intervención de mediadores informales resultó fundamental para el trabajo de observación posterior. De hecho, fueron dos compañeros de trabajo, uno encargado de las tareas más modestas y que por tanto estaba en contacto directo con los trabajadores poco cualificados, y el otro, un líder sindical, quienes, después de informarlos de sus objetivos de investi-

gación, lo introdujeron en los departamentos y le presentaron a las personas de más interés para su estudio [Cole, 1979].

Cuando Cole emprendió esa investigación tenía 28 años, más o menos la edad media de los compañeros que encontró en la fábrica. Según él mismo, en este caso la edad le facilitó las cosas. Sin embargo, representaría un obstáculo en la segunda investigación de este tipo que realizó unos años después en una fábrica de coches, cuando era algo mayor, mientras que gran parte de sus compañeros de trabajo eran veinteañeros. Este detalle nos lleva a reflexionar sobre el efecto que pueden tener ciertas características del investigador (sus rasgos externos más visibles, como la edad, el género y la raza), en su relación con los sujetos observados. En algunas situaciones, la observación participante requiere que las características del primero no difieran demasiado de las características de los segundos. Por ejemplo, resulta difícil imaginar la investigación de Whyte sobre las pandillas juveniles de Cornerville si él hubiera sido una mujer o un investigador de edad avanzada.

Obviamente, una vez logrado el acceso al entorno que se desea estudiar, el trabajo del investigador no ha hecho más que empezar. Aún tiene que ganarse la confianza de los sujetos observados mediante el trabajo cotidiano y paciente. En este contexto entran en juego los rasgos psicológicos y de carácter del investigador, su sensibilidad, su capacidad para gestionar la relación no sólo con los demás sino también consigo mismo (frustraciones, implicación emocional, etc.) [Lofland y Lofland, 1995], cap. 4]. Hay que decir que en algunos casos puede plantearse el problema contrario, es decir, que se produzca una identificación excesiva del observador con el grupo estudiado, que podría hacerle perder su capacidad para realizar una lectura crítica de los hechos observados. Whyte, en su apéndice a *Street Corner Society,* explica que la excesiva integración en el grupo de jóvenes llegó a suponer el riesgo de que el «observador no participante» de los primeros tiempos se transformara en un «participante no observador».

Con la conquista de la confianza del mediador cultural y el acceso al grupo estudiado no se acaba la necesidad del observador de establecer relaciones privilegiadas con algunos de los sujetos estudiados. Se suele denominar *informadores* a los individuos pertenecientes a la comunidad que el observador utiliza para conseguir información e interpretaciones dentro de la cultura estudiada, y con los que establece una relación personal intensa y a veces de verdadera amistad. En ocasiones se trata de personas que ocupan una posición estratégica para el conocimiento del entorno (por ejemplo, el gestor de un local público para el objetivo del conocimiento de un barrio); otras veces son personas normales, a las que podríamos llamar sencillamente «interlocutores privilegiados».

A veces se hace una distinción entre «informadores institucionales» e «informadores no institucionales». Los primeros son personas que ejercen una función formal dentro de la organización (por ejemplo, en una investigación sobre prisiones podrían ser el director, los trabajadores sociales, etc.) y que, como tales, podrían ofrecer una lectura del contexto social que estuviera influida por su lealtad hacia la institución. Más valiosa es la labor de los informadores no institucionales, que forman parte del grupo y la cultura en cuestión (en el ejemplo anterior serían los propios presos), y como tales, están en condiciones de proporcionar elementos cru-

ciales para el investigador, con su interpretación de los hechos y sus motivaciones para actuar. Hay que decir que las personas que están más dispuestas a colaborar con el investigador no siempre son también las que están mejor informadas; por otra parte, si la elección recae en una persona impopular o poco apreciada dentro de la comunidad, esto podría tener repercusiones negativas para la integración del observador en la misma. Por este motivo, es recomendable que el investigador no consolide su relación con el informador hasta haber pasado algún tiempo dentro del grupo y haber sometido a dicho informador a un severo escrutinio, antes de iniciar la colaboración con el mismo.

9.4. QUÉ OBSERVAR

La acción del observador participante debe ser selectiva, ya que no es posible observarlo todo. La observación participante no puede ser una fotografía completa de toda la realidad: de hecho, algunos objetos sociales se sitúan en el centro de atención, otros se quedan en un segundo plano, y otros quedan excluidos del objetivo del investigador.

En primera instancia, la realidad a observar viene dictada por la teoría. Conviene hacer algunas distinciones sobre el papel de la teoría en la investigación cualitativa, en general, y en la observación participante, en particular. En el capítulo 3 propusimos un esquema estándar (figura 3.1) que asignaba a la teoría un importante papel en la determinación de los objetivos de la investigación. Pero, como ya hemos indicado, mientras que la investigación cuantitativa de planteamiento neopositivista se inspira en general en un criterio deductivo, atribuyendo a la teoría la función de guía, la investigación inspirada en el paradigma interpretativo otorga una autonomía mucho mayor a la recopilación de evidencia empírica. No obstante, recordemos que dentro de ambos paradigmas existen distintas variantes. En concreto, en el caso de la investigación cualitativa, ya hemos mencionado la postura según la cual la teoría debe descubrirse en el curso de la investigación, hasta el punto de que aconseja al investigador ignorar la literatura existente sobre el problema estudiado, para evitar ser condicionado por ideas preconcebidas.

El caso es que en una situación como ésta, el investigador puede verse abrumado por la riqueza y la repetición desordenada de los estímulos que le llegan de la realidad, y, ante la ausencia de puntos de referencia, verse incapaz de ordenar y escoger entre las diversas interpretaciones.

Por ello, nosotros adoptaremos una postura más moderada, siguiendo en líneas generales la disertación de Blumer sobre los «conceptos sensibilizadores» que presentamos en el capítulo 3. Recordemos que, según este enfoque, la investigación del observador participante no parte de un vacío o *tabula rasa,* sino que se guía, sobre todo al principio, por una «sensibilidad» especial hacia determinados conceptos. Por ejemplo, el investigador de la desviación juvenil prestará especial atención a los problemas ligados a la socialización infantil, la relación con los padres, las primeras experiencias con instituciones sociales como la escuela, la relación con el grupo de amigos, etc.; en cambio, el investigador de los conflictos empresariales será más sen-

sible a las cuestiones relacionadas con la ideología del ambiente laboral, la estructura jerárquica, las expectativas sobre el trabajo y las frustraciones, las relaciones entre compañeros, las formas de comunicación entre distintos niveles jerárquicos, etc.

Conviene tener siempre presente que la observación participante «difiere de los otros métodos de investigación en el hecho de que el investigador selecciona los objetos de observación, decide qué preguntar y va construyendo sus intereses *en el curso de la propia investigación*. Todo ello contrasta con los diversos métodos de investigación "guiados por la teoría" u orientados a la "comprobación de hipótesis", cuyos objetos de observación y tipos de análisis [...] se definen claramente antes de empezar a recopilar los datos» [Lofland y Lofland, 1995, 5].

Una vez establecida esta premisa sobre la función de la teoría a la hora de decidir qué observar, repasaremos ahora los posibles objetos de observación, proponiendo una clasificación que distingue los siguientes aspectos: *a*) el contexto físico; *b*) el contexto social; *c*) las interacciones formales; *d*) las interacciones informales; y, *e*) las interpretaciones de los actores sociales. Esta clasificación de las «cosas a observar» es sólo una de las muchas posibles. En todo caso, en la observación participante siempre podemos distinguir una fase preliminar descriptiva, que aquí recogemos bajo los títulos de «contexto físico» y «contexto social», cuyo fin es fijar las características visibles externas del entorno estudiado, y una fase de interacción real con los protagonistas.

• *El contexto físico*. Es importante que el investigador observe con atención la estructura de los espacios donde se desarrolla la acción social estudiada (incluyendo una descripción detallada de los mismos en su informe de investigación), no sólo para comunicar mejor sus experiencias de observación, mediante una especie de visualización verbal, sino también porque las características físicas suelen ser expresión de características sociales (así como condicionantes de la acción social). Patton escribe: «Durante nuestras visitas a las guarderías siempre hemos hallado una estrecha relación entre lo agradable del entorno (pósters dibujados por los niños en las paredes, material de trabajo bien ordenado, zonas de trabajo del profesor) y otros aspectos de la escuela (participación de los padres, ética de los profesores, programación didáctica, etc.)» [1990, 220].

Obviamente, la exposición del investigador no deberá tener el estilo ni el tono de una novela: deberá contener descripciones, más que valoraciones o impresiones. Más que hablar de un barrio «deprimido» o «desheredado», deberá describir el estado de las carreteras, las tiendas, la estructura de la vivienda media, el transporte, los espacios públicos (la iglesia, la escuela, la oficina de correos, etc.). En una investigación sobre una empresa, dará información sobre la distribución de las oficinas y los departamentos, la ubicación de las zonas comunes, la división del espacio dentro de las oficinas en función de los diversos cargos jerárquicos, etc.

• *El contexto social*. De la misma forma en que se ha descrito el ambiente físico, se debe describir el ambiente humano. En el ejemplo anterior del barrio, se describirá a las personas que caminan por la calle en las distintas horas del día y de la noche, su forma de vestir, los objetivos de sus desplazamientos (hacer la compra,

ir al trabajo, pasear, etc.); se darán datos sobre la economía de las familias, sus ingresos y sus hábitos de consumo, etc.

En una investigación sobre una empresa, se describirá el organigrama del personal y las oficinas, las características de las personas que trabajan allí, sus funciones y tareas. También se describirán, en caso de que una comunidad se subdivida en grupos o subgrupos, las dimensiones de éstos, las personas que los componen, la frecuencia de las interacciones, los canales y la dirección de las comunicaciones (por ejemplo, en un comedor de empresa: la composición de las mesas, si los grupos se constituyen según las características personales, las mujeres juntas, los jóvenes separados de los mayores, etc., o si los grupos reflejan la distribución por departamentos, si los individuos de jerarquías diferentes se mezclan entre sí o se mantienen una clara separación entre ellos, si existe una estructura estable o los grupos van cambiando, etc.). También es importante que, al menos en esta fase de presentación inicial del entorno, el investigador haga sobre todo una descripción (lo más analítica y detallada posible), evitando introducir elementos interpretativos.

Para comprender a la comunidad es importante conocer su historia, sobre todo cuando el estudio se centra en el cambio social. Whyte aporta interesantes ejemplos de ello, como el caso de una investigación realizada en un pueblo peruano para estudiar los cambios acaecidos en los últimos cinco años: «en ese periodo se registraron, en efecto, cambios importantes, pero hemos tenido que reconocer que para comprender la dinámica de los procesos en curso hay que retroceder décadas e incluso siglos en la historia de esa comunidad» [Whyte, 1984, 161].

• *Las interacciones formales.* Las interacciones formales son aquellas que se producen entre los individuos dentro de instituciones y organizaciones, donde las funciones están preestablecidas y los tipos de vínculos están prefijados. En este contexto, el investigador estudiará, por ejemplo, cómo se produce la comunicación entre los distintos niveles jerárquicos dentro de una empresa, o la dinámica de una asamblea pública, analizando sus momentos más destacados: para ello comenzará primero con una descripción del entorno físico y humano, indicando las dimensiones de la sala, el número y las características visibles de las personas presentes, (género, edad, raza, clase social, etc.), disposición de los oradores y el público, etc. Posteriormente pasará a estudiar las acciones y las interacciones: en el caso de una asamblea pública, quién abre la sesión, qué dice, el orden de los oradores, las reacciones del público, etc. O, por ejemplo, en un estudio sobre la dinámica de un proceso en una sala judicial observará los aspectos ceremoniales del acontecimiento (la sala, las togas de los jueces, la distribución de los asientos, las fórmulas rituales de apertura, el juramento, el interrogatorio, etc.), las distintas formas de plantear las preguntas del juez, los tipos de respuesta del imputado y los testigos, la forma en que interpretan su propio papel el ministerio fiscal y el abogado defensor, las modalidades de participación del público, las reacciones a las sentencias, etc.

En un estudio sobre los procesos de toma de decisión de una organización formal se podría hacer una tipología de las decisiones (desde las más importantes hasta las rutinarias), en función del tipo de decisión, quién la toma (un individuo o

un órgano colectivo), el proceso (quién es consultado, de qué forma, de manera formal o no, si se hace mediante reuniones, cómo y cuando se convocan, etc.), qué canales se utilizan para comunicar la decisión una vez adoptada, cuáles son las consecuencias si una decisión resulta errónea, etc.

• *Las interacciones informales.* En la mayoría de los casos las interacciones informales son el elemento central de la observación participante. Precisamente por su informalidad resulta tan difícil estudiarlos y, además, su observación conlleva miles de casos diferentes sobre los que es imposible formular no ya reglas, sino incluso indicaciones de carácter general. En cualquier caso, intentaremos delimitar algunos aspectos.

Podemos decir que el observador puede empezar por las interacciones físicas. Spradley dice que toda situación social se define por tres elementos: el lugar, los actores y las actividades. Al observar éstas actividades, el investigador «al principio quizás sólo vea un confuso bullir de comportamientos, centenares de actos aparentemente iguales. Tras repetidas observaciones podrá ordenar los actos individuales por acciones, como cazar, correr, pedir algo para beber, elegir un asiento en el autobús, hacer la compra en el supermercado» [Spradley, 1980, 41].

Muy a menudo el punto de partida de la observación participante es el comportamiento cotidiano, que precisamente por ser tan corriente es el más difícil de analizar, ya que está formado por una serie de actos mecánicos de los que el actor social es poco consciente. Por ejemplo, una persona que espera en la cola del autobús es poco consciente de lo que pasa a su alrededor (las otras personas en la cola, su ropa, quién mira el reloj con nerviosismo, quién lee el periódico, quién lleva la bolsa de la compra, los coches que pasan, el silbido de un tren en la lejanía, un mendigo que pide limosna, un chico en bicicleta, etc.), y también es poco consciente de su propio comportamiento, por su frecuencia, repetición y carácter mecánico. El observador participante, en cambio, debe ser consciente de todos estos fragmentos de vida cotidiana, debe saber observarlos y aislarlos, para poder analizarlos. Y debe saber captar la interacción de las personas que observa: quién empuja y quién cede el paso al subirse al autobús, la dinámica de elección del asiento, dos desconocidos que se chocan sin darse cuenta, los escolares que arman jaleo, las reacciones de los viajeros, las dos señoras que charlan, el tono de voz, los gestos, etc.

Es importante que el investigador se vaya centrando gradualmente en las interacciones que le interesan. Al principio, su campo de observación será muy amplio, pero a medida que avanza la investigación su actuación será cada vez más selectiva. Imaginemos un investigador que pretende estudiar las relaciones cliente-dependiente en las tiendas. Tras una observación sobre los distintos tipos de tiendas, el investigador se da cuenta de que debe acotar la variedad, por lo que decide limitarse a un determinado tipo de tiendas: las librerías. A continuación advierte que la mayoría de las relaciones entre dependientes y clientes se desarrolla a la salida del establecimiento, en el momento de pagar y controlar al cliente, por lo que se sitúa cerca de la salida y observa. Después puede decidir restringir el campo de observación, limitándose al estudio de los conflictos entre el cliente y el dependiente: por

ejemplo, los conflictos derivados del pago, o los conflictos que surgen cuando se descubre que un cliente intenta salir de la tienda con un libro sin haberlo pagado. Observará entonces la dinámica de los hechos, las características (género, edad, profesión, etc.) de las personas implicadas, las reacciones de los diversos actores sociales (cliente, dependiente, público, etc.), y el desenlace (denuncia, compensación económica, excusas, etc.). Al final resultará que su idea inicial de estudiar las relaciones dependiente-cliente se ha transformado en un estudio sobre los intentos de hurto en las librerías. El investigador no siempre tiene tanta libertad para definir el objeto del estudio en el curso de la investigación. A veces ciertas condiciones externas ponen límites a su arbitrio. Pero en todo caso, a medida que avanza la observación, deberá ir centrándose en objetos y procesos sociales cada vez más específicos.

• *Las interpretaciones de los actores sociales.* Sabemos que en el paradigma interpretativo el individuo estudiado no es sólo un objeto pasivo de la investigación, sino que se convierte en sujeto activo protagonista, y su interpretación de la realidad se hace parte constitutiva (y no sólo accesoria) del conocimiento científico. En este ámbito la interacción verbal con el observador participante se convierte en canal privilegiado de comunicación. Como escriben Hammersley y Atkinson, «cada acto humano tiene una dimensión expresiva propia. La organización del entorno, el atuendo, los gestos, las formas de conducta en general, son portadores de mensajes. Indican el género, el estatus social, la ocupación y hasta la personalidad de los actores sociales. Sin embargo, es la fuerza expresiva del lenguaje lo que constituye la fuente de conocimiento más importante. [...] En la vida cotidiana los individuos emiten continuamente descripciones lingüísticas de su propia visión del mundo: por ejemplo, cuando hablan sobre sus aspiraciones, cuando cuentan "lo que sucedió" [...], cuando piden perdón o justifican su conducta» [Hammersley y Atkinson, 1983, 107].

Todo ello sucede entre los actores sociales, pero puede producirse también entre el investigador y los sujetos estudiados. El investigador «participa», observa y pregunta, y la *interrogación,* en forma de coloquio informal y entrevista formal, se une a la observación como instrumento de investigación. La pregunta incorporada en la conversación informal, para solicitar información y comprender las motivaciones y el punto de vista del actor social, es parte constitutiva de la observación participante clásica, que implica siempre un solapamiento de miradas y preguntas. Pero también la entrevista programada y formalizada, durante la cual el investigador plantea de manera explícita una serie de preguntas, es parte del instrumental usual del observador participante. Obviamente, la entrevista explícita define una situación distinta y en cierto sentido «artificial» en comparación con la conversación informal, revelando de manera incluso demasiado evidente el papel del investigador y su disparidad respecto al mundo en el que se ha integrado, y podría restar naturalidad a la propia observación. Sin embargo, como escriben Hammersley y Atkinson, «esta disparidad de la entrevista debe verse como un recurso [...]. Dado que el objetivo del etnógrafo no es sólo ofrecer una descripción, [...] pueden ser ventajoso el hecho de someter a las personas a estímulos verbales diferen-

tes de los que experimentan en su vida cotidiana» [*ibíidem*, 119].

Como sabemos, la entrevista también puede utilizarse de forma autónoma como instrumento de recopilación de datos para la investigación social. Siguiendo el esquema tradicional de los manuales de metodología, dedicaremos un capítulo entero a la entrevista. No obstante, el lector debe ser consciente de la importancia de la misma dentro de la técnica de la observación participante.

9.5. REGISTRO DE LA OBSERVACIÓN

El proceso de registro de los datos observados es fundamental en toda investigación científica, tanto en ciencias naturales como en ciencias sociales. En el caso de la observación participante, ésta adopta la forma de apuntes cotidianos, una especie de cuaderno de bitácora donde se refiere con riqueza de detalles y reflexiones personales todo lo que el investigador ha observado a lo largo del día. La redacción de estas notas es parte de la observación participante. Los Lofland afirman que «tras la observación, la tarea fundamental del investigador es la de redactar notas: si no lo hace, puede ahorrarse la observación»; y añaden que «el parte diario es, en un sentido pleno y real, "el dato", y ello significa que los datos no son los recuerdos del investigador. [...] Éstos consisten más bien en lo que ha sido registrado a diario» [Lofland y Lofland, 1995, 67, 89]. Hammersley y Atkinson [1983, 150] sugieren que, en caso de que la observación sea demasiado ardua y deje poco tiempo para la redacción de notas, se alternen periodos de observación y periodos de escritura, para no correr el riesgo de descuidar esta última.

En efecto, debe evitarse confiar únicamente en la memoria para almacenar los datos recogidos, ya que es casi seguro que provocará errores. No sólo por las limitaciones obvias a la cantidad de información que el ser humano es capaz de memorizar, sino sobre todo por el hecho de que nuestra memoria es enormemente (e inconscientemente) selectiva, lo que implica un riesgo fundado de distorsión del material memorizado. Puede suceder, por ejemplo, que en la memoria del investigador el material quede registrado con una determinada interpretación, y excluya todo lo que contradiga esa interpretación. Esta distorsión es mucho menos probable si el material se transcribe inmediatamente después de ser recopilado.

Las notas cotidianas del investigador nacen de la interacción entre el observador y la realidad observada, por lo que constan básicamente de dos componentes: la *descripción* de los hechos, acontecimientos, lugares y personas, y la *interpretación* de los mismos del investigador, con sus impresiones, reflexiones y reacciones. Es importante destacar esta doble vertiente de la representación objetiva y la comprensión subjetiva. Obviamente, no es fácil separar los dos componentes, y hasta la descripción más «objetiva» puede resentirse de la visión cultural específica de quien describe [cfr. Geertz, 1973]. Las notas etnográficas constituyen la primera formalización del encuentro entre las dos culturas, la que estudia y la estudiada: el investigador debe ser consciente de esta circunstancia, y ello debe orientar su forma de redactar las notas y organizar el material observado.

A continuación estudiaremos en detalle el registro de las observaciones, des-

componiéndolo en sus tres aspectos fundamentales: «cuándo» realizarlo, «qué» registrar y «cómo» hacerlo.

• *Cuándo*. La respuesta inmediata es lo antes posible y lo más cerca posible del hecho. El tiempo que pasa puede hacer perder la vivacidad del detalle, los hechos nuevos se solapan con los pasados y difuminan sus detalles. Lo ideal sería tomar apuntes mientras se producen los hechos, pero esto raras veces es posible. No es posible cuando la observación es encubierta, pero también es difícil cuando es declarada, por la evidente perturbación que provocaría en el entorno. Se recomienda que el observador tenga siempre a mano un cuaderno, sobre el que pueda ir realizando breves apuntes (por ejemplo, ausentándose un momento con alguna excusa), que podrá desarrollar después cuando haya ocasión de ello, o que lleve una grabadora portátil, por ejemplo, con un pequeño micrófono escondido en la solapa de la chaqueta, para grabar sus impresiones durante la observación. En cualquier caso, todos los manuales de etnografía aconsejan, y hasta prescriben, que el observador lleve un diario cotidiano de su actividad donde podrá ordenar el material recopilado a lo largo del día, redactar sus observaciones y recapitular la situación.

• *Qué*. Hemos dicho que las notas deben reflejar: *a*) la descripción de los hechos y *b*) la interpretación del investigador. A estos dos componentes podemos añadir un tercero: *c*) las interpretaciones de los sujetos estudiados. La descripción de los acontecimientos debe lo más detallada posible. Pero no se puede describir todo, hay que seleccionar los acontecimientos que se describirán. Sin embargo, sobre todo al principio, es importante que el investigador haga descripciones extensas, que cubran al máximo la realidad observada; es probable que después se acote más el foco del análisis, pero este proceso de restricción se irá produciendo conforme avanza la investigación. Además, puede suceder que en el curso de la investigación, el centro de atención de la misma se desplace, por lo que en las primeras fases conviene tener un campo de observación lo más amplio posible.

La interpretación del investigador suele constar de dos partes: las reflexiones teóricas y las reacciones emocionales. La primera está directamente vinculada al objetivo principal de la observación y representa un primer esfuerzo reflexivo orientado a acumular material y puntos de partida (que deben fijarse inmediatamente, pues de lo contrario se perderían), que luego se incluirán en el informe final. La segunda parte, los sentimientos del investigador, también es importante. La observación participante requiere una implicación del investigador que va más allá del puro compromiso intelectual: es más, sabemos que la participación emocional es uno de los requisitos para la comprensión. Por tanto, la manifestación explícita y el registro de los propios sentimientos representa para el observador no sólo una forma de autoanálisis útil para el control de sí mismo durante el trabajo de campo, sino también una documentación provechosa para una reconstrucción *a posteriori* de la dinámica que se ha creado entre el observador y el sujeto observado mediante una relectura crítica de los datos recopilados, con el fin adicional de identificar posibles distorsiones generadas por las emociones del investigador.

Por último, las interpretaciones de los sujetos estudiados, tal y como surgen de

frases escuchadas, de conversaciones informales con el observador y de entrevistas más formalizadas constituyen el tercer aspecto de esta documentación, que no es en absoluto secundario; como escriben Hammersley y Atkinson, los relatos realizados por las personas estudiadas se deben tratar de forma idéntica a los producidos por el investigador» [1983, 126].

• *Cómo.* Hay que diferenciar bien los tres componentes de las notas que hemos presentado en el párrafo anterior: descripciones, interpretaciones del observador e interpretaciones de los sujetos estudiados. El parte detallado de un acontecimiento debe estar claramente separado del comentario del investigador. Dentro de este comentario también debe quedar clara la división entre reflexiones teóricas y reacciones emocionales. Por último cada una de las interpretaciones de los sujetos implicados se debe atribuir claramente a quien las ha expresado, evitando reunirlas todas en una mezcla confusa. Éste es el primer principio del registro de las notas, el principio de la *diferenciación.*

El segundo principio es el de la *fidelidad* del registro. Éste es relevante sobre todo para las notas que contienen las interpretaciones de los sujetos observados: sus expresiones verbales deben reproducirse con total exactitud, incluso las palabras en jerga, indicándolas con comillas. En efecto, la vivacidad pictórica del relato de un protagonista es muy distinta del resumen que puede ofrecer el investigador. El empleo de la grabadora en las entrevistas formales puede ser de gran utilidad en este contexto. Pero también hay que decir que ésta debe utilizarse con moderación, dejando la decisión al criterio del investigador. La grabación de la entrevista subraya ese carácter de artificialidad que ya hemos mencionado en la entrevista formal; por tanto, es oportuno realizarla sólo cuando existe una sólida relación de confianza entre el entrevistador y el entrevistado, y, en todo caso, cuando la función del investigador ya ha sido plenamente aceptada por la comunidad estudiada.

En relación con el «cómo», queremos hacer un apunte sobre las herramientas disponibles actualmente para el investigador que utiliza la técnica de la observación. Por una parte podemos mencionar las ventajas de los medios audiovisuales: a los tradicionales, como las grabaciones de sonido y las fotografías, se ha añadido recientemente la cámara de vídeo, cuyo uso ha abierto nuevos horizontes para la observación. En relación directa con la redacción de las notas escritas, recordemos que la difusión de los ordenadores y la aparición de programas para la gestión de observaciones etnográficas, ponen a disposición del investigador nuevos instrumentos para la organización del material recopilado, la lectura sinóptica de los mismos episodios, las conexiones entre datos heterogéneos con aspectos comunes, etc.

9.6. ANÁLISIS DEL MATERIAL EMPÍRICO

Hemos llegado a la que probablemente sea la parte más difícil de la investigación mediante observación participante: el análisis del material empírico recopilado y la redacción del informe final. Con el análisis de la documentación empírica se realiza la delicada conversión del lenguaje de los sujetos observados en categorías con-

ceptuales de la teoría sociológica.

Aunque el objetivo de la observación participante es comprender el punto de vista de los sujetos observados, también pretende hacer aflorar aspectos de los que los propios nativos no son conscientes. Como escribe Van Manen, el investigador debe observar la realidad social desde dos puntos de vista. Uno es la perspectiva desde el interior: ésta «obliga al investigador a entrar en el mundo de las personas estudiadas y salir de él con una lectura realizada con los instrumentos interpretativos que ha encontrado allí (presentes en las acciones, palabras, formas simbólicas y emociones de los sujetos estudiados)» [Van Manen, Dabbs y Faulkner, 1982, 17]. El otro punto de vista es aquel que, desde el exterior, trata de desvelar los aspectos del mundo social estudiado que son desconocidos para los propios protagonistas de la acción social.

La fase de análisis e interpretación del material empírico es muy laboriosa, cualquiera que sea la técnica empleada para la recopilación de datos; no obstante, como ya hemos dicho, en investigación cualitativa, la ausencia de procedimientos normalizados hace que esta fase dependa en gran medida de las capacidades personales del investigador. En el caso de la observación participante, es aun más complicado debido a la enorme cantidad de material recopilado, que suele ser heterogéneo y fragmentado, y puede resultar muy difícil de ordenar para el investigador.

Normalmente los manuales de metodología ofrecen reglas prácticas sobre la forma de organizar el material empírico y enseñan la forma de hacer clasificaciones, diagramas, esquemas, etc. con el fin de sintetizarlo y extraer de él generalizaciones. Esto se ha producido sobre todo a partir de los años ochenta, cuando la renovada popularidad de la investigación cualitativa chocaba con la ausencia de procedimientos normalizados, que desorientaba al investigador inexperto.

Sin embargo, no se puede decir que estos esfuerzos hayan dado resultados especialmente útiles: con frecuencia se trata de «recetarios» que despiertan sonrisas por la banalidad de las recomendaciones que proponen, o causan indignación porque lo único que proponen es una imitación de los procedimientos estadísticos de la investigación cuantitativa. En la relación empática entre el investigador y la realidad observada propia de la observación participante, los componentes individuales de los sujetos implicados son determinantes, y la interacción que se establece entre ellos es el principal generador de «comprensión» sociológica. Por tanto, creemos que sólo es posible ofrecer unas directrices generales sobre el modo de analizar los datos y extraer consecuencias teóricas, y por ello aquí nos limitaremos a exponer estas escuetas sugerencias.

Podemos empezar diciendo que el análisis de los datos de la observación participante es un proceso *continuo,* que tiene lugar en parte durante la propia observación. Este solapamiento de la observación y el análisis es incluso indispensable, puesto que gestionar muchos datos sin detenerse a reflexionar puede llevar a acumular grandes cantidades de material sin orden ni sentido y difíciles de manejar; en cambio, si se analizan a medida que se recogen, produciendo unas primeras elaboraciones teóricas provisionales, todo estará más claro, y podremos orientar mejor la observación y centrarnos en los temas principales. Por tanto, el análisis de los datos es también *retroactivo y cíclico,* en el sentido de que conforme avan-

za la reflexión teórica, vuelve sobre elementos observados ya analizados, para reinterpretarlos a la luz de las nuevas conclusiones o para reutilizar el material empírico en nuevas interpretaciones con un grado de abstracción superior, intercalando así distintos niveles de análisis.

Un problema común al que debe enfrentarse inevitablemente el investigador que utiliza la técnica de la observación participante es lo que Lofland [1971, 123] denominó el «drama de la selección». El proceso de análisis en observación participante tiene forma de embudo, con las delimitaciones sucesivas del centro de análisis. Esta estructura conlleva unas serie de decisiones ineludibles y cada vez más drásticas, que obligarán a abandonar gran parte del material recopilado. Con frecuencia se trata precisamente del material preferido por el investigador, las intuiciones que le parecen más afortunadas. Pero en este punto el investigador debe ser intransigente consigo mismo: es más importante asegurarse de que su disertación presenta argumentos sólidos y un estilo ágil, que aceptar divagaciones tal vez ricas en intuiciones pero que podrían hacer perder el hilo del análisis.

Pasando a los aspectos operativos, hay que decir que la primera fase consiste en la *descripción* de cada una de las cuestiones abordadas por el investigador. La descripción (que contrasta con la explicación y en particular la explicación causal, la única vía para el verdadero conocimiento científico según el enfoque positivista), se considera en observación participante un «producto netamente científico», que resulta fundamental para la comprensión sociológica, y no queda limitado a una función auxiliar de fondo.

La descripción en la observación participante no es sólo la exposición sensorial de lo que los ojos del investigador ven o sus oídos oyen: debe ser, tal y como lo expresó con acierto el antropólogo Clifford Geertz, una descripción «sustanciosa», en la que no sólo se presenta lo que se ve y lo que se oye (eso sería una descripción superficial), sino que esta percepción sensorial se enriquece con significados e interpretaciones, se enmarca en un contexto cultural e histórico y dentro de una red de relaciones sociales. «Imaginad a dos chicos que contraen rápidamente los párpados del ojo derecho. Para uno se trata de un tic involuntario; para el otro, una señal de entendimiento a un amigo. Los dos movimientos son idénticos [...]. Sin embargo, hay una gran diferencia entre un tic y un guiño» [Geertz, 1973; 6]. El significado que enriquece la descripción no siempre es unívoco, ya que «el etnógrafo se encuentra ante una multitud de estructuras conceptuales complejas, muchas de las cuales se solapan o entrecruzan entre sí» [*ibídem, 10*]. Por ello la descripción del antropólogo es un relato con una trama compleja que presenta distintos niveles de lectura, un relato «denso».

A ello hay que añadir el hecho de que la descripción del observador participante no está compuesta sólo por el relato de lo que ha visto. Se trata de una verdadera «construcción» que resulta de ensamblar materiales diversos: «narraciones» (procedentes de entrevistas en profundidad, historiales de vida, grabaciones en directo de los protagonistas), reportajes periodísticos, crónicas, textos de discursos, material documental de todo tipo, etc., junto con las reseñas visuales de los ambientes sociales, los hechos y los personajes observados. En este sentido, tampoco se trata de una mera descripción: cada elección sobre el material a presentar lleva implícita

una interpretación concreta, y hasta la más simple de las descripciones visuales implica una perspectiva cultural específica.

Después de la descripción viene la *clasificación*. Al observar el mundo social, la configuración física de los ambientes sociales y los grupos, los comportamientos de los individuos y los hechos colectivos, los actos individuales y las acciones más complejas, el investigador captará repeticiones y similitudes entre distintos objetos sociales, que le permitirán construir clasificaciones.

Tal vez el caso más sencillo de clasificación es la identificación de secuencias temporales. En su investigación sobre las cárceles en Italia, Ricci y Salerno [1971] descomponen el proceso de ingreso en la cárcel en varios pasos, una secuencia que se repite en casi todos los casos estudiados: la detención, la conducción de los detenidos a la comisaría, la toma de huellas digitales, la notificación de la orden de encarcelamiento, el traslado a la cárcel, el paso simbólico de la entrada, el trámite de asignación de un número al detenido, el cacheo, la entrega de mantas y plato, y el aislamiento en la celda. Los investigadores interpretan después todos estos actos como los aspectos de un proceso de despersonalización del hombre-detenido, que es despojado de todas sus posesiones personales, para transformarse de individuo con nombre y apellidos en recluso anónimo identificado por un número.

La identificación de secuencias temporales conlleva descomponer un flujo continuo de acontecimientos en bloques separados en el tiempo. Otra forma de clasificación es la que consiste en ordenar una masa de objetos sociales diversos, agrupándolos en clases, en función de sus similitudes/diferencias. Los objetos sociales que pueden clasificarse de este modo son muy dispares: pueden ser personas, comportamientos, instituciones, acontecimientos, etc.

Un ejemplo de clasificación de comportamientos sería la investigación realizada por Walum [1974] sobre el acto de «apertura de una puerta» que tiene lugar cada vez que dos extraños tienen que entrar en el mismo momento por la misma puerta. Walum estudió este acto para el caso concreto en que los dos extraños son de distinto sexo, por lo que entran en juego valores como masculinidad y feminidad, iniciativa y pasividad, anticonformismo, independencia e incluso, en momentos determinados y en ciertos ambientes, ideología política.

Walum observó el desarrollo de este acto en las puertas de la facultad en la que daba clases y clasificó a los sujetos observados en cinco tipos, según el comportamiento mostrado: los «confusos» (que no saben cómo comportarse y acaban chocando el uno contra el otro), los «desafiantes» (que son conscientes del significado del acto en la interacción entre sexos y lo utilizan para desafiar al otro), los «humanitarios» (que abren la puerta si ven que la otra persona lo necesita, por ejemplo, si va cargada con unos paquetes, o por educación, pero con independencia del sexo de la persona), los «tradicionales» (que saben que las costumbres al respecto han cambiado pero no quieren adaptarse) y los «rebeldes» (que se oponen deliberadamente a las normas vigentes).

El siguiente paso es la identificación de las *dimensiones de la tipología* (o clasificación). El observador suele distinguir los tipos en función de sus valoraciones personales sobre la similitud o la diferencia, mediante un proceso mental sintético e intuitivo. En cambio, si actuamos de forma analítica y conceptual, debemos reve-

lar la estructura conceptual de la clasificación, indicando las características que hacen que los tipos sean diferentes entre sí.

Como ejemplo podemos citar la investigación de Rothschild-Whitt [1979] sobre la organización del trabajo. Observando seis cooperativas del sur de California y comparando su forma de trabajar y de organizarse con la estructura tradicional de la organización del trabajo en la sociedad occidental, Rothschild-Whitt identificó ocho dimensiones que reflejan las diferencias entre ambos modelos: autoridad (que puede encontrarse en los individuos situados en la cúspide de la organización o en la colectividad), normas (que pueden estar formalizadas y definidas con detalle o ser mínimas y establecerse sobre la marcha), control social (que puede encargarse a controladores concretos o a la colectividad), relaciones sociales (que pueden ser impersonales o personalizadas), reclutamiento (que puede seguir unos cauces prefijados, como pruebas de acceso, o hacerse recurriendo a los conocidos), incentivos (monetarios o morales), estratificación social (jerárquica o igualitaria), división del trabajo (que puede ser máxima, con separación entre trabajo manual e intelectual, o mínima, asignándose tareas muy distintas a la misma persona).

Este proceso de identificación de las dimensiones de una tipología permite desentrañar la estructura conceptual de la clasificación y llegar a esas abstracciones teóricas denominadas tipos ideales, de las que hablamos al principio del libro.

Quizás la forma más general de interpretar el material recogido mediante observación participante es identificar lo que Spradley [1980, 140] llama los *temas culturales* que están presentes en la sociedad estudiada. Spradley recuerda que la investigación del observador participante tiene que abarcar siempre dos niveles: «analizar los pequeños detalles de la cultura que está estudiando, y al mismo tiempo intentar trazar las líneas generales del panorama cultural, [...] los principios que dan sentido al conjunto». Spradley emplea el concepto de «tema cultural» del antropólogo Morris Opler, que lo utilizó para describir la cultura de los Apaches. En este caso uno de los temas culturales era que «los hombres son física, mental y moralmente superiores a las mujeres», y este principio tácito era una de las bases de su concepto del mundo y determinaba muchos de sus comportamientos en el ámbito familiar y social.

Retomando un ejemplo citado anteriormente, uno de los temas culturales presentes en el mundo de las bandas estudiadas por Jankowski era lo que él denominó «visión darvinista» de las relaciones sociales, según la cual la vida es una lucha continua por la supervivencia: cada día tenemos que competir por los escasos recursos disponibles, el prójimo sólo puede considerarse un enemigo y sólo sobrevive el más fuerte.

La identificación de los temas culturales es uno de los objetivos finales de los estudios de inspiración antropológica, y con frecuencia suele haber un tema principal que ilumina toda la investigación. Éste es el caso de la investigación que Edward Banfield realizó en los años cincuenta en Montegrano, un pueblo de una región del sur de Italia. Banfield encontró en el *familismo amoral* un rasgo cultural que marcaba todos los aspectos de la vida de la población, y que consistía en una incapacidad total de sus habitantes «para actuar juntos por el bien común o para

conseguir cualquier fin que vaya más allá del interés material inmediato de la familia nuclear» [Banfield, 1958, 10].

Este tema cultural está presente en todo el estudio de Banfield, y el autor lo utiliza para interpretar una larga serie de hechos que observó: en Montegrano no existen organizaciones de beneficencia, la propia Iglesia es paupérrima (pocas personas asisten a la misa del domingo, y casi todas son mujeres; las ofrendas son modestas), los partidos políticos prácticamente no existen, faltan asociaciones voluntarias (la única que hay es un club recreativo, que sólo sirve para jugar a las cartas y nunca se ha ocupado de resolver los problemas del pueblo), no existen iniciativas de grupos de profesionales con fines colectivos (los comerciantes, aunque están interesados en tener buenas conexiones por carretera, nunca han hecho nada para reclamar mejoras en las desastrosas infraestructuras existentes), etc.

Para finalizar hablaremos del estilo preferido para la redacción de las conclusiones de una investigación realizada con la técnica de observación participante. Se suelen emplear dos adjetivos para describir este estilo de redacción: *reflexivo* y *narrativo*.

La «reflexividad» de la investigación hace referencia al hecho de que el investigador forma parte del mundo que está estudiando. Y, por tanto, así como el proceso de investigación del observador participante no es un proceso impersonal y neutral, también el estilo de escritura está estrechamente ligado a la personalidad y la cultura del investigador. En este aspecto es muy distinto del estilo de la encuesta por muestreo (por citar el prototipo de la investigación cuantitativa). En la encuesta por muestreo el estilo de presentación de los resultados es impersonal, formal y, por tanto, suele estar escrito en tercera persona; en la observación participante es más informal y está escrito en primera persona. También hay que decir que la escritura del informe etnográfico no se realiza al final de la investigación y separada de la misma, tras la recogida y el análisis de los datos (como sucede en la investigación cuantitativa), sino que, como hemos visto, parte de su redacción se realiza durante la observación misma, mediante la redacción de notas, apuntes, etc. De ahí que el investigador también se implique a la hora de redactar, aportando sus emociones y reacciones personales, que se convierten en parte viva del relato.

A menudo el informe concluyente de la investigación con observación participante va acompañado de un apéndice autobiográfico, en el que el investigador relata el itinerario recorrido para realizar el trabajo: el génesis de la investigación, la identificación del caso específico estudiado, el acceso y la aceptación en el grupo, los mediadores culturales, los informadores, los problemas que van surgiendo, los progresos de la reflexión teórica, los cambios de perspectiva, etc.; y, por otra parte, sus reacciones emocionales, los errores cometidos, las dificultades encontradas. Es una exposición que tiene utilidad informativa y metodológica. En un tipo de investigación tan marcado por las decisiones subjetivas del investigador, esta reconstrucción puede ayudar al lector a evaluar mejor la validez del trabajo, a interpretar el significado de ciertas conclusiones, a valorar los posibles efectos de perturbación provocados por la acción y la presencia del observador, y a comprender hasta qué punto los resultados obtenidos están condicionados por imposiciones externas que los convierten en un caso único, o si es posible hacer generalizaciones a partir de

ellos.

El estilo debe ser «narrativo», similar al relato o la crónica periodística, y distante *tanto* de la abstracción conceptual de la formulación teórica *como* de la abstracción aritmética de la investigación cuantitativa. El estilo narrativo es una forma de escribir concreta y directa, con descripciones detalladas, crónicas pormenorizadas de acontecimientos, textos que reproducen las intervenciones en directo de los protagonistas, con personajes comparables a los de un relato. Según Spradley: «el niño tendrá un nombre propio; lo tendrá en brazos una madre bien definida que le estará dando el pecho en un momento concreto y en un lugar específico» [1980, 162]. El observador participante no desea extraer la identidad general del sujeto estudiado, eliminando sus connotaciones locales y contingentes para construir un concepto abstracto, sino que precisamente basa su comprensión y su capacidad de comunicarla al lector en la especificidad y singularidad de esa situación concreta.

La redacción debe reflejar una forma «natural» de describir la realidad social, que viene a ser como «contar una historia» [Van Manen 1988]. El fin es transmitir al lector lo máximo posible de su experiencia observadora, ofreciéndole la oportunidad de hacer una lectura imaginativa de la cultura estudiada, mediante una descripción rica en sentimientos y matices, capaz de transmitir la riqueza de detalles y la vitalidad del mundo observado. Hammersley y Atkinson citan, por ejemplo, las afinidades literarias entre el trabajo de Malinowski y las novelas del escritor inglés Joseph Conrad, o entre los trabajos de quienes estudiaron la vida urbana inglesa de finales del siglo xix, como Booth y Rowntree, y los cuentos de Dickens.

Según Hammersley y Atkinson, «lo que podemos definir como «el relato de una historia» es sólo una parte del trabajo del etnógrafo. Éste también está ocupado, de un modo u otro, en la enunciación de una teoría» [1983, 210]. Aunque una parte importante del corpus del informe etnográfico la constituye esa «descripción sustanciosa» que utiliza los recursos pictóricos de la retórica literaria, no hay que olvidar que el objetivo final de la investigación social y, por tanto, también de la observación participante, es la elaboración de alguna forma de generalización o teoría. De ello se deriva la necesidad, al menos en la fase final del informe de investigación, de que el análisis supere el nivel de lo particular para pasar a formular tipologías, temas culturales, y, más en general, elaboraciones teóricas.

9.7. LA OBSERVACIÓN EN «MICROSOCIOLOGÍA»

9.7.1. Del estudio de «los otros» al estudio de «nosotros mismos»

Antes de concluir el tema de la observación participante queremos detenernos en ciertas evoluciones del método que no siempre pueden encuadrarse bajo el título general de observación participante de origen etnográfico, pero que, sin embargo, tienen una inspiración común con ésta y no pocos puntos de contacto con sus procedimientos empíricos.

La técnica de la observación participante, como ya hemos explicado, surge en

el marco de la antropología como instrumento para estudiar y describir una cultura «distinta» (de la del observador). En sus primeras aplicaciones en la sociología, mantiene sus características originales de técnica para el estudio de «otras» culturas. Del estudio de las tribus primitivas, pasó al estudio de las «tribus urbanas», pero siempre enfocado en grupos de individuos «distintos», es decir, segmentos de la sociedad separados de la corriente principal, con una subcultura propia y diferente de la cultura dominante, ya sean grupos sociales (bandas juveniles, sectas religiosas, minorías étnicas, grupos desviados, etc.) o comunidades territoriales (suburbios, barrios urbanos, ciudades de provincias, etc.).

Sin embargo, a medida que se desplazaban del estudio de tribus más lejanas a tribus cada vez más próximas, los investigadores advirtieron que el método podía usarse no sólo para estudiar a los «distintos», sino también para estudiarse a sí mismos. Y percibieron que con la observación participante y el enfoque del etnógrafo se podía analizar cualquier aspecto del ser humano que tuviera especificidad cultural. En efecto, así como existe una subcultura de la banda juvenil o de los aficionados al fútbol, también existe una subcultura dentro del hospital, de una organización empresarial, de un partido o de un colegio profesional.

Siguiendo con este proceso de aproximación al estudio de sí mismos y su *propia* cultura, se ha ido descubriendo que no sólo los aspectos organizados de la vida social tienen su cultura, sino que incluso los momentos más banales de la cotidianidad, los actos mecánicos de cada día, también son portadores de una cultura oculta y están cargados de significados implícitos que escapan a los propios actores sociales. De este modo, en un sector de la sociología fue surgiendo el desafío de «estudiar de manera sistemática los hechos y actividades aparentemente vulgares [...]. Haciendo de la banalidad un tema de investigación y reconociendo su importancia en la vida cotidiana, [estos] sociólogos [...] transforman el objeto que quieren estudiar. En otros términos, las banalidades ya no son banales, sino que llegan a adquirir importancia» [Schwartz y Jacobs, 1979, 183].

No entraremos en el análisis detallado de las corrientes que han creado lo que a veces se denomina *microsociología,* que estudia «pequeños segmentos de espacio y tiempo y pequeñas cantidades de personas: el individuo y la interacción, el comportamiento y la conciencia» [Collins, 1988; 3]. Diremos simplemente que las primeras contribuciones en este ámbito proceden de los escritos de George Herbert Mead, y más adelante, de Herbert Blumer, en los años treinta.

Blumer aplicó las enseñanzas de Mead a la investigación social, dando lugar a la escuela del *interaccionismo simbólico.* La clave de su planteamiento está en la importancia que atribuye al proceso de interpretación. El autor escribe que los seres humanos actúan con las cosas y las personas según los *significados* que éstas tienen para ellos (una mujer puede ser una madre o una dependienta, una silla puede ser una pieza de mobiliario para sentarse cómodamente o un objeto para contemplar, un colegio tiene un significado distinto para los profesores, los alumnos, los padres, los empleados o las personas sin hijos). Estos significados proceden de la forma en que los individuos reaccionan ante esa cosa o persona; los significados no son individuales, sino sociales, y el individuo los aprende mediante la *interacción* social con sus semejantes. El conjunto de significados compartidos es

lo que llamamos «cultura». El individuo hace suyos estos significados mediante un proceso de *interpretación* de los hechos sociales que lo rodean, que depende de las premisas culturales de cada uno.

Este planteamiento se encuentra en perfecta sintonía con los principios metodológicos de la observación participante, que nace precisamente de la exigencia de captar los significados atribuidos por los individuos a sus actos, de percibir el mundo con sus propios ojos. Al mismo tiempo, el hecho de que el análisis se centre en la comunicación verbal y no verbal, en la definición de situaciones y en los mecanismos de interpretación, implica una reconstrucción minuciosa de las interacciones sociales que permite desplazar la atención del investigador hacia los pequeños hechos, en lugar de centrarse en procesos sociales de mayor envergadura.

Este desplazamiento hacia el análisis de la vida cotidiana recibe un fuerte impulso en los años sesenta y setenta gracias a la obra de Erving Goffman. Goffman estudió sobre todo la interacción social y sus ritos, tratando de identificar sus modelos. Estudió lo que sucede en la oficina, en la cocina, en una cena formal, en una conversación; analizó el comportamiento en lugares públicos (lo que hacen las personas en una lavandería, cómo se cruzan dos transeúntes en la calle, la forma de vestir, etc.). En el probablemente sea su libro más famoso [Goffman, 1959], compara la vida con el teatro: la vida consiste en representaciones con sus actores y su público. Tenemos el escenario para la representación formal, y las bambalinas para las interacciones informales; y cita los ejemplos de los obreros que se comportan de una forma determinada en presencia del jefe y dejan de actuar cuando éste se va; lo mismo sucede con los jueces, los médicos, los estudiantes y todos nosotros, que tenemos nuestro escenario y nuestro papel «entre bambalinas». Según Goffman, toda la vida social está guiada por rituales, de la conversación informal a los comportamientos «educados», o las relaciones más íntimas y personales. La violación de estas normas sociales se interpreta como un síntoma de enfermedad mental.

9.7.2. La etnometodología

Esta atención por las pequeñas interacciones de la vida cotidiana encuentra su expresión teórica en la *etnometodología*. Este término fue acuñado por Garfinkel en los años cuarenta a partir de las palabras griegas *èthnos* (pueblo), *métodos* (método) y *lògos* (discurso), y hace referencia al «análisis de los métodos ordinarios que los individuos ordinarios emplean para realizar sus actos ordinarios» [Coulon, 1995, 2]. Así como la etnobotánica y la etnomedicina son el estudio de la forma en que las diversas culturas clasifican las plantas o practican la medicina, la etnometodología es el estudio de los métodos y prácticas que las personas normales emplean para descodificar el mundo, darle un significado, llevar a cabo cualquier acto: es el estudio del «razonamiento práctico de sentido común», la «ciencia del mundo cotidiano».

Según este planteamiento, la actuación cotidiana está regulada por normas bien definidas que rigen la interacción entre las personas. Se trata de normas implícitas, de las que el sujeto agente, el individuo, no es consciente, aunque se adapta a ellas continuamente. Con el fin de reconocer estas normas, los investigadores propusie-

ron incumplir las convenciones tácitas, a fin de sacarlas a la luz. De este modo empezaron a realizarse los denominados experimentos etnometodológicos, que podían consistir, por ejemplo, en acercarse más de la cuenta a una persona durante una conversación, en beber de los vasos de otros comensales en una cena, en hablar solos por la calle, en quitarse los zapatos en una reunión de trabajo, etc. Las reacciones de desorientación, cuando no de irritación, que provocan tales conductas son, según los investigadores, la confirmación explícita de que violan una norma implícita que rige la interacción social, y que queda desvelada mediante su violación.

Los planteamientos mencionados, el interaccionismo simbólico, la aportación de Goffman y la etnometodología, son muy diferentes entre sí, pero también presentan rasgos comunes. En particular, su *objeto*, los actos de la vida cotidiana, y su *objetivo*, la interpretación de estos actos mediante un proceso que podríamos llamar de «problematización de lo obvio».

Esta rama de estudios ha ofrecido nuevo material de investigación y ha abierto nuevos horizontes de investigación para la sociología y, en particular, para la investigación realizada con técnicas de observación. Aunque hay que decir que no siempre se puede hablar de investigaciones que utilizan plenamente esta técnica: como veremos en los ejemplos ofrecidos, con frecuencia se trata de pura observación sin participación.

Entre las investigaciones inspiradas más directamente en la teoría de Goffman podemos citar los trabajos de Adam Kendon [1990] sobre la organización del comportamiento en la interacción a corta distancia entre individuos. Kendon llevó a cabo estudios sobre las funciones de la mirada en la conversación entre dos personas, partiendo de la consideración de que el lenguaje de los ojos es un medio de comunicación importante entre los hombres. La investigación de Kendon, realizada hacia la mitad de los años sesenta, consistía en la grabación audiovisual de conversaciones, de una media hora de duración, entre pares de personas, sentadas una frente a otra. Obviamente el objetivo de la grabación era captar la dirección de las miradas durante la conversación, y su dinámica estaba correlacionada con el contenido de la conversación.

Kendon estudió también la forma de saludarse de las personas. La observación se realizó en fiestas al aire libre celebradas en casas privadas, y consistió en la grabación con cámara de vídeo tanto de los saludos de los dueños de la casa a los invitados conforme iban llegando a la fiesta, como de los saludos de los invitados entre sí cuando se encontraban con los demás. Como escribe Kendon, «con el ritual del saludo los actores se indican mutuamente el estatus social, el grado de familiaridad y simpatía que existe entre ellos y también, con frecuencia, los roles que van a desempeñar en el encuentro que se acaba de iniciar» [*ibídem*, 154].

En relación con la interacción entre dos personas también podemos citar los estudios sobre la pelea. La pelea es una dinámica interpersonal que está presente en todos los contextos sociales (tanto en la familia como en el trabajo), entre personas de todas las edades (ya sean niños o adultos), a todos los niveles de intimidad (entre los miembros de una pareja o entre dos conductores que no se conocen). En todas estas situaciones es fácil percibir unas estructuras comunes, donde las fases

de la disputa (iniciación del conflicto, argumentaciones, negaciones, acusaciones, amenazas, desafíos, excusas, compromisos, negociaciones, resolución del conflicto, etc.), pueden encuadrarse en unos modelos básicos que se repiten en todas ellas.

Allen Grimshaw [1990] ha recogido en una antología algunos estudios sobre las discusiones. Entre ellos podemos citar una investigación de Corsaro y Rizzo [1990] sobre las discusiones entre niños de entre tres y cinco años, realizada mediante observación sobre el terreno y grabación de las conversaciones en una escuela infantil estadounidense y una italiana, en la que hallaron similitudes y diferencias entre las dos culturas. En la misma antología se cita una investigación sobre discusiones familiares, realizada mediante la grabación de conversaciones durante 64 cenas de otras tantas familias estadounidenses de clases sociales y etnias distintas [Vuchinich, 1990]. En este caso el autor estudió la forma en que se resuelve el conflicto, identificando cinco tipos: la sumisión de uno de los contendientes, la intervención de una tercera persona desde una posición autoritaria (que suele ser uno de los padres), el acuerdo, la suspensión del conflicto sin resolución (por ejemplo, cuando una intervención externa desplaza la atención de los contendientes), y la retirada de uno de los contendientes (que dice no querer seguir discutiendo, abandona la mesa, etc.). Como puede verse, el objetivo de la investigación es identificar unos modelos recurrentes, y el procedimiento de análisis es el clásico de identificación de tipos generalizables.

Podemos citar otros ejemplos de investigación procedentes de estudios sobre comportamientos colectivos. Heritage y Greatbatch [1986] realizaron un estudio sobre las génesis del aplauso, analizando las grabaciones de 476 discursos políticos pronunciados en 1981 en Gran Bretaña durante congresos de partidos. Los autores seleccionaron las partes del discurso que iban seguidas de aplausos, y de esta forma lograron identificar los mecanismos retóricos (estructura de la argumentación, secuencias de afirmaciones y negociaciones, tonos de voz, pausas, etc.) que utilizan los oradores para provocar un aplauso. Los autores constataron que las tres cuartas partes de los aplausos siguen a siete esquemas retóricos bien definidos, al margen del partido y el contenido de las afirmaciones.

La investigación de Clayman [1993] que estudia las manifestaciones de desaprobación del público sobre una serie de discursos políticos pronunciados en Estados Unidos y Gran Bretaña trata un tema similar al del ejemplo anterior. El autor analizó las fases que preceden a la manifestación abierta de desaprobación (murmullos, silbidos, etc.), los mecanismos de inicio y evolución de la desaprobación (detonación por parte de un pequeño grupo, propagación entre los presentes, reacción de los incondicionales, etc.), así como las formas de reaccionar del orador (defensas explícitas, silencios, empleo de los tonos de voz, intentos de seguir el discurso como si no hubiera pasado nada, etc.).

Los estudios sobre el modo en que las personas se desplazan por las calles también están estrechamente ligados a la perspectiva etnometodológica de desvelar las reglas y la estructura del comportamiento cotidiano. Los movimientos que realizan las personas que caminan por las transitadas aceras de una ciudad para cruzarse sin chocar con los demás son comportamientos sólo aparentemente naturales, en realidad, responden a unas reglas muy precisas que aprendemos desde pequeños

(la antropóloga Lyn Lofland destaca que una de las formas clásicas que emplea el cine para caracterizar a un niño de campo es su incapacidad para moverse entre la multitud ciudadana sin chocar con los demás).

Entre los diversos trabajos sobre este tema podemos citar el de los investigadores ingleses Collet y Marsh [1974], que instalaron una cámara de televisión en un séptimo piso sobre Oxford Circus, en Londres, y grabaron los movimientos de los peatones al cruzar la calle. La observación consistía en escoger a una pareja de peatones y filmar todos sus movimientos. Los investigadores analizaron los distintos momentos de la dinámica: por ejemplo, codificaron analíticamente el instante del cruce, clasificando los movimientos del cuerpo (hombros, brazos, cabeza) y la mecánica de los desplazamientos (líneas directrices del acercamiento, tiempos empleados, paso a la derecha o a la izquierda, etc.), descubriendo entre otras cosas que los hombres y las mujeres actúan de manera distinta (la mayoría de los hombres giran el torso, mientras que las mujeres dan la espalda, se protegen más con un brazo, etc.). Collet y Marsh estudiaron también la comunicación corporal, es decir, todos esos mensajes involuntarios (la indicación inconsciente de intenciones, los movimientos de respuesta, la función de colaboración de las miradas, etc.), que generan ese movimiento coordinado (el «baile de los peatones», como lo llaman los autores) de la multitud que cruza la calle.

Otro tema clásico de este tipo de estudios son las colas. Para Mann las personas que esperan en una cola son un «sistema social en miniatura, [...] con normas tácitas que regulan los comportamientos para avanzar y conservar el puesto, para gestionar las ausencias temporales, para aplicar sanciones» [1969, 340]. Mann realizó investigaciones sobre ciertas colas muy largas que se forman en diversas ocasiones y en distintas culturas: las colas que forman los aficionados en Melbourne cada mes de agosto para comprar las entradas para las finales de fútbol (en 1965 se formó una cola de 25000 personas para comprar 12500 entradas), las colas que se forman cada semana en Nueva York ante el edificio de la Opera, para comprar entradas, las colas de Tokio ante las agencias de alquiler de viviendas, las de Londres en las paradas de autobuses, en México para reservar asiento en el autobús, las colas en las tiendas de alimentación en Rusia y Cuba, las de Nicaragua ante las rudimentarias lavanderías para hacer la colada.

Mann analizó el fenómeno de la cola bajo distintos puntos de vista: la autoorganización de las personas cuando la cola dura varios días, las sanciones aplicadas para castigar los comportamientos que violan las normas, el negocio que florece en torno a las colas (personas que cobran por hacer la cola por otros, niños que cuentan cuántas personas hay delante y por poco dinero se lo dicen a quienes esperan, etc.), las motivaciones por las que las personas hacen cola, aunque sepan que no conseguirán lo que quieren (por ejemplo, una interminable cola de niños ante un cine de Melbourne esperando recibir la camiseta de Batman, que se entregaría sólo a los primeros 25 niños), y así sucesivamente. Además, Mann percibió una conexión entre la forma de hacer la cola y el arraigo en la cultura en cuestión de los valores de igualitarismo y orden social, distinguiendo a Inglaterra como el país donde la «cola democrática» es, más que en ningún otro sitio, un verdadero «estilo de vida» [Mann, 1973, 56].

El análisis de la conversación es un campo en el que la investigación de inspiración etnometodológica ha tenido un gran desarrollo, hasta el punto de convertirse prácticamente en una disciplina independiente. Se parte de la consideración de que la conversación es una de las formas más comunes de interacción entre los individuos, y que, como todas las formas de interacción, no se desarrolla de manera casual, sino que obedece a una serie de reglas no escritas y se estructura según modelos estándar, de los que los interlocutores no son conscientes y que son una parte fundamental de la cultura correspondiente. La tarea del análisis de la conversación es desvelar esta estructura subyacente, mediante la descomposición de la conversación en unidades básicas y la identificación de secuencias típicas.

Para hacer esto se creó una técnica de transcripción de conversaciones en la cual se transcriben junto al texto y mediante una simbología adecuada, formas de comunicación como el énfasis, los tonos de voz, las pausas, las interrupciones, las palabras no acabadas, las superposiciones de frases, los suspiros, los murmullos, etc. Con esta descomposición y el análisis de las diversas fases de la conversación, estos estudios revelaron algunas estructuras recurrentes: las reglas que rigen la atribución de turnos en la conversación, las modalidades de inicio y cierre, los modos de corregirse a sí mismos y al interlocutor, de mostrar simpatía e interés o parecer poco interesados, las preguntas implícitas, las que no esperan respuesta, las diferencias entre la conversación informal, entre amigos, y la formal, en reuniones de trabajo, la relación asimétrica entre juez e imputado, médico y paciente, etc.

Para concluir este tema, el lector habrá percibido la distancia entre esta investigación de tipo «microsociológico» y la investigación clásica basada en la observación participante, no sólo desde el punto de vista del contenido (aquí limitado a los pequeños actos de la vida cotidiana), sino también desde el punto de vista de los procedimientos empleados. En efecto, el análisis de la interacción y el comportamiento ordinario se suele basar en la *pura observación,* en la reconstrucción de una mecánica social compuesta por de actos básicos, más que en la observación *participante* orientada a captar la percepción subjetiva del individuo. Sin embargo, el lector habrá captado también los puntos de conexión entre este enfoque, que pretende descubrir el significado oculto de los pequeñas actos cotidianos (como cruzar una calle), con el estudio de los significados, valores y temas culturales que subyacen a comportamientos más complejos (como la participación en un oficio religioso o la expresión del voto). Esta conexión justifica nuestra decisión de incluir la etnometodología dentro de este capítulo dedicado a la observación participante.

9.8. LIMITACIONES Y RECURSOS

La técnica de la observación participante también presenta ciertas dificultades. En este apartado discutiremos las tres más importantes: *a*) la subjetividad; *b*) la falta de generalización; y, *c*) la falta de estandarización.

- *Subjetividad (del investigador).* Hemos dicho que en la observación participante el investigador es el *instrumento* de la investigación, en el sentido de que

todo el proceso de recopilación de datos se filtra por sus ojos y sus sentidos, su sensibilidad y su capacidad de identificación. A estos condicionantes de carácter muy personal hay que añadir los condicionamientos culturales. Como escribió Spradley, la investigación etnográfica es «una cultura que estudia una cultura» [Spradley, 1980, 13]; y, al ser una interacción entre dos culturas, su resultado dependerá de la cultura estudiada, pero también de la que estudia. El antropólogo Clifford Geertz escribe: «Lo que llamamos nuestros datos son en realidad nuestras interpretaciones de las interpretaciones de otros sobre lo que ellos y sus compatriotas hacen [...]. Ya desde el momento de la exposición de los hechos [...] estamos dando explicaciones; y, lo que es peor, explicaciones de explicaciones» [Geertz, 1973, 9]. En otras palabras, al describir otras culturas, el etnógrafo también se describe a sí mismo: las preguntas que plantea, la selección de los hechos y las interpretaciones que da de éstos dependen de su punto de vista. No hay duda alguna de que ante el mismo objeto de estudio, por ejemplo, las obras de construcción de una autopista en la selva o en un poblado berebere, un investigador indio y otro noruego darían descripciones radicalmente distintas.

- *Falta de generalización (de los casos estudiados).* Las dificultades para generalizar los resultados de una investigación realizada mediante observación participante, se derivan, por una parte, de la subjetividad del investigador que estudia, y por otra, de la subjetividad de los casos estudiados. La observación participante suele consistir en el estudio de un caso o de unos cuantos casos; se trata de estudios en profundidad, pero a escala pequeña. Por otra parte, el requisito de profundidad que exige este planteamiento impide su aplicación a gran escala. De ahí que algunos afirmen que se trata de estudios sobre realidades que sólo se representan a sí mismas, «cuyos resultados describen fenómenos únicos en lugar de clases de fenómenos, y cuyas conclusiones no pueden aplicarse a ningún grupo distinto del estudiado» [Borman, LeCompte y Goetz, 1986, 48].

- *Falta de estandarización (de los procedimientos utilizados).* Al principio de este capítulo hemos hablado de la dificultad para describir y enseñar la técnica de la observación, debido a la ausencia de procedimientos estandarizados aplicables de forma universal, y hemos subrayado la especificidad de cada proceso de investigación. De estas características se deriva también la imposibilidad de reproducir dichos estudios: si el investigador cambia, cambian también los sujetos y los entornos observados, las formas de observar, la secuencia de la producción de datos sociales y, por tanto, las propias características del material empírico utilizado.

Sin embargo, conviene subrayar que la subjetividad del proceso observador, la singularidad de cada caso estudiado y imposibilidad de replicar el proceso de investigación, son una *limitación,* pero al mismo tiempo también una *ventaja* de la observación participante. Esta aproximación a la conciencia social no aspira a la objetividad. Al contrario, reivindica el valor de su propia subjetividad, afirmando que sólo mediante la implicación personal es posible llegar a comprender el objeto de estudio, y que para lograr el verdadero conocimiento se debe pasar por la experiencia directa. Hay que añadir que, aun sin estar de acuerdo con esta visión, para

quien se mueve en un plano epistemológico distinto es difícil no reconocer los estímulos y los conocimientos adquiridos gracias al método etnográfico.

Por último queremos hacer una consideración sobre un aspecto un tanto especial de la observación participante: las *dificultades de realización*. La observación participante requiere una gran inversión de tiempo y recursos psicológicos. Se necesita tiempo para llegar a entender la estructura de otras formas de pensar, para comprender los valores, las normas, los códigos de comunicación de otra cultura, un tiempo que no todo el mundo tiene. Como escribe John Madge, «nunca ha sido fácil convencer a los investigadores de que se conviertan en observadores participantes. Éstos deben dedicarse por completo a la investigación y deben estar libres de compromisos familiares y otras responsabilidades» [1962; 217]. No es casual que, entre las investigaciones sociológicas más interesantes realizadas con observación participante, encontramos un número considerable de tesis doctorales, con un trabajo de campo realizado cuando los autores eran jóvenes de poco más de veinte años, y tenían, por tanto, mucho tiempo (y entusiasmo) disponible.

La otra inversión que requiere la observación participante son recursos psicológicos. El trabajo de campo conlleva una inmersión en la realidad social estudiada que puede provocar situaciones difíciles de gestionar desde el punto de vista emocional: requiere capacidad de adaptación psicológica, capacidad para controlar los propios sentimientos y gestionar relaciones interpersonales difíciles.

A ello se añade el hecho de que en el curso de una investigación con observación participante, el investigador puede tener que hacer frente a situaciones de aislamiento. Hemos dicho que el investigador es el instrumento de la investigación, ya que ésta depende en gran medida de su capacidad personal para percibir y entender, pero eso significa también que él carga con todo el peso de la investigación. No existen procesos preestablecidos, procedimientos estadísticos, que puedan guiarle y ayudarle, y en último término, subsanar, con modelos estandarizados, sus posibles debilidades.

9.9. SÍNTESIS DEL CAPÍTULO 9

1. La observación participante es una técnica de investigación mediante la cual el investigador se integra, durante un periodo de tiempo relativamente largo, en un grupo social determinado, en su entorno natural, y establece relaciones de interacción personal con sus miembros, con el fin de observar sus actos y comprender sus motivaciones mediante un proceso de identificación. En la observación participante la implicación y la identificación no sólo no se evitan, sino que se buscan (y por tanto la objetividad y la distancia, que formaban la base del planteamiento neopositivista, no se consideran un objetivo deseable).
2. La observación participante se suele utilizar para estudiar dos tipos de grupos: pequeñas sociedades autónomas, ubicadas en una zona concreta y dotadas de un universo cultural cerrado que abarca todos los aspectos de la vida *(estudios de comunidad)*, y subculturas que se han desarrollado dentro de segmentos sociales, que pueden representar aspectos de la cultura dominante, o estar en

conflicto con ella *(estudios de subculturas).*

- En la observación participante el investigador puede declarar abiertamente desde el principio sus verdaderos objetivos, es decir, que desea entrar en el grupo sólo para estudiarlo *(observación declarada),* o puede infiltrarse en el grupo fingiendo formar parte del mismo como un miembro más *(observación encubierta).* Para entrar en la comunidad que quiere estudiar, el investigador utiliza con frecuencia un *mediador cultural,* una persona que goza de la confianza de la comunidad y que, por sus características culturales, también comprende las motivaciones y las necesidades del investigador. Otra figura de la observación participante son los *informadores,* personas pertenecientes a la comunidad estudiada que el observador utiliza para conseguir datos e interpretaciones desde el interior de la cultura estudiada.
- Para determinar «qué observar», los objetos de observación potenciales pueden clasificarse en las siguientes categorías: contexto físico, contexto social, interacciones formales, interacciones informales, e interpretaciones de los actores sociales.
- En la observación participante el proceso de registro de los datos observados tiene una gran importancia. ¿Cuándo debe realizarse ese registro? Lo antes posible. ¿Qué debe registrarse? Descripción de hechos, interpretaciones del investigador e interpretaciones de los sujetos estudiados. ¿Cómo debe registrarse? Aplicando los criterios de la *distinción* (entre descripción e interpretación del investigador y de los sujetos) y la *fidelidad* (al reproducir comentarios de los sujetos, puntos de vista, etc.).
- El análisis del material empírico suele pasar por las siguientes fases: descripción, clasificación, identificación de las dimensiones de las tipologías y construcción de las tipologías (o *tipos ideales*). La forma más general de interpretar el material recogido mediante observación participante consiste en la identificación de los *temas culturales* presentes en la sociedad estudiada.
- La observación participante nació como instrumento para estudiar las culturas «distintas». Del estudio de las tribus primitivas, pasó al estudio de las «tribus urbanas», aunque manteniendo su objetivo de estudiar a aquellos que de algún modo eran considerados «distintos». A continuación se comprendió que cada campo de la actividad humana tiene una cultura específica y propia que se puede estudiar con el criterio etnográfico y la observación participante. La microsociología es el estudio de las pequeñas acciones cotidianas, y la etnometodología es el estudio de los métodos y las prácticas empleadas por las personas normales para descodificar el mundo, darle significado y actuar. Los actos cotidianos están regulados por normas implícitas bien definidas que gobiernan las interacciones entre las personas. El objetivo de la etnometodología es objetivo desvelar y estudiar estas normas.

9.10. LECTURAS COMPLEMENTARIAS

Como presentaciones generales del planteamiento etnográfico en la investigación social, que contienen tanto reflexiones teóricas como indicaciones prácticas, sugerimos tres textos clásicos: P. Spradley, *Participant observation* (Holt, Rinehart y Winston, 1980, 195 pp.); J. Lofland y L. H. Lofland, *Analizing Social Settings* (Wadsworth, 1995, 2ª ed., 186 pp.); M. Hammersley y P. Atkinson, *Ethnography: Principles in Practice* (Routledge, 2ª ed., 1995, 312 pp.).

Para una reflexión sobre las tendencias más recientes y una propuesta inspirada en el enfoque llamado posmoderno, sugerimos la obra de N. Denzin *Interpretive Ethnography: Ethnographic Practices for the Twenty-first Century* (Sage, 1997, 326 pp.).

Una interesante reflexión procedente del punto de vista «positivista» sobre las orientaciones actuales del planteamiento etnográfico en investigación social, se encuentra en J. Goldthorpe, *Sociological Ethnography Today: Problems and Possibilities*, en Goldthorpe (2000).

Por último, sobre la etnometodología sugerimos una introducción en D. W. Maynard y S. E. Clayman, «The Diversity of Ethnomethodology» (en *Annual Review of Sociology, 1991*), y una exposición más completa, en J. Heritage, *Garfinkel and Ethnomethodology* (Polity Press, 1984, 336 pp.)

CAPÍTULO **10**

LA ENTREVISTA CUALITATIVA

- LA OBTENCIÓN DE DATOS MEDIANTE PREGUNTAS 344
- ENTREVISTA CUANTITATIVA Y ENTREVISTA CUALITATIVA 345
- TIPOS DE ENTREVISTA 349
  - 10.3.1. Entrevistas estructuradas, semiestructuradas y no estructuradas 350
  - 10.3.2. Casos especiales 357
- REALIZACIÓN DE LA ENTREVISTA 361
- ANÁLISIS DEL MATERIAL EMPÍRICO 366
  - 10.5.1. El análisis de tipo cualitativo 366
  - 10.5.2. El análisis de tipo mixto (cuantitativo y cualitativo) 367
- FALTA DE ESTANDARIZACIÓN E INTERACCIÓN 369
- SÍNTESIS DEL CAPÍTULO 10 372
- LECTURAS COMPLEMENTARIAS 373

La técnica de la entrevista cualitativa, como la encuesta por muestreo, pretende obtener datos preguntando a los sujetos, pero con el objetivo característico de la investigación cualitativa de conocer la individualidad de la persona entrevistada y ver el mundo con sus ojos. En este capítulo empezaremos planteando las diferencias entre la entrevista cualitativa y la cuantitativa, a continuación presentaremos los distintos tipos de entrevista cualitativa, las formas de realizar las entrevistas y las fases en que se realizan. Por último ofreceremos algunas sugerencias sobre la forma de analizar el material empírico obtenido.

10.1. LA OBTENCIÓN DE DATOS MEDIANTE PREGUNTAS

La entrevista cualitativa se puede considerar equivalente de la observación participante en cuanto a los objetivos del investigador. Obviamente, la inmersión en la realidad social que se realiza con la entrevista cualitativa no es tan profunda como con la observación participante. Con la entrevista cualitativa el investigador no pretende entrar en el mundo estudiado hasta el punto de identificarse con él para poder verlo con los ojos de sus protagonistas. Pero, en todo caso, su objetivo último es *conocer la perspectiva del sujeto estudiado*, comprender sus categorías mentales, sus interpretaciones, sus percepciones y sentimientos, los motivos de sus actos.

Podemos definir la entrevista cualitativa como *una conversación: a) provocada por el entrevistador; b) realizada a sujetos seleccionados a partir de un plan de investigación; c) en un número considerable; d) que tiene una finalidad de tipo cognitivo; e) guiada por el entrevistador;* y, *f) con un esquema de preguntas flexible y no estandarizado.*

Vamos a repasar los puntos de esta definición. En primer lugar, la entrevista la provoca el entrevistador, y en esto se distingue de la conversación ocasional. Ello no significa que la conversación ocasional no resulte útil para la investigación social: de hecho, el observador participante suele usarla con frecuencia para comprender mejor la realidad en la que está inmerso. Pero la entrevista cualitativa es una conversación solicitada de manera explícita por el entrevistador (el encuentro se suele programar con una cita).

En segundo lugar, la entrevista cualitativa se realiza a una serie de sujetos seleccionados según un plan sistemático de recogida de datos, lo cual quiere decir que ni siquiera el entrevistado es ocasional: éste es elegido por sus características (que haya vivido determinadas experiencias, que pertenezca a determinadas categorías sociales, etc.). Hay que contar con un número considerable de sujetos (por lo general, no menos de varias decenas [1]). Por tanto, no se trata de entrevistas esporádicas (como sucede en las entrevistas periodísticas a sujetos aislados, como un líder sindical, un policía, el protagonista de un suceso, etc.), sino de entrevistas a un número de sujetos tal que permita inferir informaciones generalizables a

[1] En las investigaciones que presentaremos aquí, el número de sujetos entrevistados nunca es inferior a 50; lo más común es que sean un centenar de entrevistas.

una población más amplia (aunque no se pretende llegar a muestras representativas en un verdadero sentido estadístico).

El objetivo de esta conversación es la finalidad cognitiva del entrevistador. Por tanto, no son ocasionales ni el hecho, ni la persona entrevistada, ni el tema de la conversación. Y su objetivo es estrictamente cognitivo, de modo que nuestro concepto de entrevista no incluye el diálogo del psicólogo con el paciente (que tiene una finalidad terapéutica) ni la entrevista de trabajo (que tiene la finalidad de evaluar al candidato).

Por último, no se trata de una conversación normal, un diálogo entre dos personas en el que los roles de los interlocutores están equilibrados y situados al mismo nivel, sino de una conversación guiada, en la que el entrevistador establece el tema y controla que su desarrollo responda a los fines cognitivos que él se ha marcado. Esta intervención del entrevistador puede ser más o menos marcada, pero en todo caso deberá respetar la libertad del entrevistado para estructurar su respuesta o la conversación como crea conveniente (de lo contrario estaríamos en el caso de la encuesta por muestreo y el cuestionario).

10.2. ENTREVISTA CUANTITATIVA Y ENTREVISTA CUALITATIVA

Para obtener información sobre la sociedad planteando preguntas a los sujetos podemos utilizar dos instrumentos: un cuestionario estandarizado (es decir, estructurado de manera rígida en las preguntas y respuestas), o una entrevista «libre» (es decir, sin una estructura rígida establecida de antemano)[2]. Estas técnicas parecen similares, pero en realidad son muy distintas, puesto que se basan en dos filosofías de investigación o paradigmas divergentes.

Ya hemos mencionado las diferencias de fondo entre estos dos modos de hacer investigación social. A continuación vamos a revisarlas a partir de las características de la entrevista cualitativa.

- *Falta de estandarización.* Ésta es la diferencia fundamental entre el cuestionario y la entrevista, diferencia que genera todas las demás. Mientras que el objetivo del cuestionario es enmarcar al entrevistado en una serie de esquemas preestablecidos por el investigador (las respuestas a una pregunta cerrada), el fin de la entrevista es comprender las categorías mentales del entrevistado, sin partir de ideas y conceptos predefinidos. Como afirma Michael Patton, «la finalidad de la entrevista cualitativa es entender cómo ven [el mundo] los sujetos estudiados, comprender *su* terminología y *su* modo de juzgar, captar la complejidad de *sus* percepciones y experiencias individuales [...]. El objetivo prioritario de la entrevista cualitativa es proporcionar un marco dentro del cual los entrevistados puedan expresar *su propio* modo de sentir

[2] Cuando hablamos de «entrevista» sin más, nos referimos a la entrevista cualitativa; para referirnos a la entrevista realizada según el criterio cuantitativo utilizaremos el término «cuestionario».

con sus propias palabras» [1990, 290, cursiva en el original].

El criterio cuantitativo obliga al entrevistado a limitar sus propias respuestas, lo enseña a autocontrolarse. Si el entrevistado quisiera exponer sus propias ideas, por ejemplo explicar por qué, según su experiencia personal, prefiere la gestión privada de los hospitales a la pública, no podría hacerlo. Estas reflexiones del entrevistado no se consideran relevantes y no son registradas por el entrevistador, que le ha preguntado si está «muy, bastante, poco o nada» a favor de la gestión privada de la sanidad. Al entrevistado no se le permitiría hablar ampliamente sobre un tema que le interesa mucho y sobre el que tiene sus propias convicciones, ni extenderse menos sobre otro que no le interesa y sobre el que no sabe qué responder, puesto que el esquema de interrogación es inflexible: el entrevistador debe plantear las mismas preguntas a todos los sujetos, no puede variarlas según los entrevistados ni éstos pueden desviar la conversación hacia temas que concuerden mejor con su personalidad y sus experiencias.

En todos estos casos la voz del entrevistador domina la del entrevistado. En cambio, en la entrevista cualitativa, aunque el entrevistador sigue planteando los temas de conversación, *la voz sobresaliente debe ser la del entrevistado:* en el caso extremo (y en algunos sentidos ideal), el entrevistador habla muy poco, se limita a estimular y animar al entrevistado, que podrá exponer con libertad absoluta su forma de ver las cosas, sus motivaciones, sus pensamientos.

Dado que el objetivo es *tener acceso a la perspectiva del sujeto estudiado,* la entrevista cualitativa debe ser flexible, adaptarse a las personalidades distintas de los entrevistados, poder cambiar de un sujeto a otro. Debe conceder al entrevistado plena libertad de expresión, para ponerlo en condiciones de destacar su propio punto de vista utilizando sus propias categorías mentales y su propio lenguaje. En definitiva, el instrumento de la entrevista no puede estar (o debe estar poco) estandarizado.

• *Comprensión frente a documentación (contexto de descubrimiento frente a contexto de justificación).* Las preguntas pueden ser un instrumento de recopilación de datos o un instrumento para comprender la realidad social. En el primer caso, dentro del planteamiento cuantitativo, las preguntas se utilizan para recopilar información sobre las personas en relación con sus opiniones, sus comportamientos y sus características sociales. Si pudiéramos obtener los datos que necesitamos de otra forma, por ejemplo fotografiando a los sujetos, podríamos sustituir las preguntas por fotografías. En cambio, en el planteamiento cualitativo, la entrevista *no* se emplea, en primera instancia, para recopilar datos sobre las personas, sino para hacerlas hablar y entender sus puntos de vista.

La diferencia entre estos dos criterios, que se puede considerar una diferencia entre cantidad y calidad, extensión y profundidad, condiciona también el número de sujetos a entrevistar. Supongamos, por ejemplo, que queremos estudiar el fenómeno de la delincuencia menor, tirones de bolso, hurtos, agresiones, vandalismo, en un barrio urbano, y deseamos conocer la incidencia cuantitativa, los grupos sociales afectados, los tipos de actos criminales más habituales. Dado que muchos de los pequeños casos de delincuencia (hurtos de bicicletas, intentos de

tirones de bolso, robos en comercios, etc.) no se denuncian, éstos sólo pueden conocerse preguntando a las personas. En este caso la pregunta se utiliza para obtener datos; se seleccionará una muestra de ciudadanos y se les preguntará: «¿En los últimos meses ha sido usted objeto de episodios como [...]?», etc.; y así sucesivamente.

Por otro lado, se puede decidir utilizar el instrumento de la pregunta, no para cuantificar el fenómeno de la delincuencia menor, sino para profundizar en sus efectos sobre la sensación de seguridad de los ciudadanos, la confianza en las instituciones, el grado de inquietud y miedo, los cambios en las costumbres de la población. Es decir, no usamos la pregunta para recopilar datos, sino para percibir el fenómeno social estudiado con los ojos y los sentimientos de los entrevistados.

En el primer caso se tratará de una recogida extensiva (se entrevistará por ejemplo a una muestra de varios cientos de personas), puesto que el objetivo es obtener una imagen representativa de la situación social y analizar después los datos recogidos con instrumentos estadísticos. En el segundo caso se tratará de una recogida intensiva: un único caso, con el cual el entrevistador ha sabido establecer una relación de empatía, puede resultar más útil que una decena de cuestionarios para entender e interpretar una determinada dinámica social.

La distinción entre los objetivos de recogida de datos y comprensión de la realidad social nos lleva a la distinción introducida por Reichenbach entre el *contexto del descubrimiento* (el momento de concepción de una idea nueva) y el *contexto de la justificación* (el momento de su comprobación empírica). En las preguntas estandarizadas prevalece el contexto de la justificación: se pregunta al entrevistado por qué está a favor o en contra de la pena de muerte, presentándole varias opciones de respuestas para elegir, porque el investigador considera que ya tiene un conocimiento del fenómeno estudiado suficiente como para poder reconocer todas las respuestas posibles, y sólo pretende «recopilar datos», con el fin de cuantificar el peso de las distintas alternativas y relacionar las respuestas a estas preguntas con otras variables.

Si, por el contrario, el investigador no conoce los motivos que llevan a las personas a estar a favor o en contra de la pena de muerte, y quiere utilizar el instrumento de la entrevista para «descubrir» este aspecto de la realidad, emplearía preguntas no estandarizadas: en este caso su objetivo principal es *comprender* el fenómeno estudiado. *Este objetivo no excluye el anterior:* más adelante veremos situaciones de investigación en las que la entrevista cualitativa se utiliza *también* para recoger datos sobre la realidad social. Pero su *objetivo prioritario* sigue siendo la comprensión, y los datos cuantitativos recogidos quedarán como un subproducto: descarnados y empobrecidos para poder compararlos entre sí y sacrificados por las modestas dimensiones de la muestra.

- Falta *de muestra representativa.* Otra diferencia entre cuestionario y entrevista, derivada de los dos puntos anteriores, está relacionada con la muestra. La encuesta por muestreo se realiza sobre una muestra construida de tal forma que se pueda considerar «representativa», es decir, que pueda reproducir a un tamaño reducido las características de la población.

La entrevista cualitativa no aspira a este objetivo. Aunque la elección de los sujetos a entrevistar se realice de manera sistemática, este plan nace más de la exigencia de cubrir una variedad de situaciones sociales que del objetivo de reproducir las características de la población a una escala reducida. Y no porque el objetivo no sea deseable, sino porque resultaría caprichoso y en cualquier caso inviable dadas las dimensiones de la muestra. Con frecuencia las muestras utilizadas en las investigaciones con entrevistas cualitativas rondan el centenar de casos, cantidad insuficiente para poder considerarlas representativas.

Si no es necesario que la muestra sea representativa, tampoco es necesario que la elección de los sujetos se haga de forma aleatoria. Por lo general, el plan de selección de las personas a entrevistar se realiza del siguiente modo: se identifican unas cuantas variables importantes en relación con el tema estudiado (entre dos y cuatro), se combinan estas variables (nominales), y se entrevista a un número determinado de sujetos que cumplan las características de la combinación de variables. Los sujetos se eligen con el criterio del muestreo por cuotas, dejando al entrevistador libertad para elegir a quién entrevistar, siempre que se cumpla la exigencia establecida por la cuota.

Por ejemplo, en la investigación sobre los militantes del Partido Comunista (PCI) y la Democracia Cristiana (DC) realizada en Italia en los años sesenta [Alberoni *et al.*, 1967] se entrevistó a 108 activistas, usando como base del diseño de muestreo las variables «partido» (dicotómica: DC y PCI) y «ciudad de residencia» (con seis valores distintos correspondientes a seis ciudades italianas). La combinación de estas dos variables genera doce celdas, para cada una de las cuales se eligió nueve sujetos, con un total de 108 entrevistas (54 por partido, 18 por ciudad). Una vez establecido que en la ciudad A había que entrevistar a nueve militantes de la DC y nueve del PCI, la elección de los mismos correspondía al entrevistador.

Obviamente, este procedimiento no genera una muestra que reproduzca las características de la población. Pero, como ya dijimos en el capítulo 2, el investigador cualitativo no sigue un criterio de representatividad estadística, sino más bien de representatividad sustantiva, para cubrir todas las situaciones sociales de interés para la investigación, más que para reproducir las características de la totalidad de la población.

De forma similar , en una investigación realizada en Estados Unidos en los años ochenta [Riessman 1990] sobre la reorganización de la vida personal y social tras el divorcio, se entrevistó a 52 hombres y 52 mujeres, un tercio de cada grupo debía llevar separado menos de un año, otro tercio entre uno y dos años y otro tercio entre dos y tres años.

Sin embargo, en la investigación cualitativa suele faltar un plan de muestreo previo. Aunque el investigador sigue siempre un criterio para la elección de los sujetos que entrevistará, éste se puede construir en el curso de la investigación. Como ya indicamos en el capítulo 2, el investigador elegirá los casos no por su presencia en la población, sino por el interés que tienen según él, que puede variar en el curso de la investigación. Otras veces la recogida de datos sistemática y predefinida resulta imposible debido a la naturaleza del problema estudiado, la singularidad del fenómeno o la dificultad de determinar a priori los sujetos de interés.

En la investigación llevada a cabo por Terry Arendell [1986] sobre la forma de reaccionar de las mujeres ante las crisis de divorcio, en un primer momento la autora trató de seguir un procedimiento clásico de muestreo, extrayendo 200 nombres de mujeres divorciadas de los registros de dos tribunales de Estados Unidos, y enviando a éstas una carta en la que les anunciaba que se pondría en contacto con ellas más adelante; sin embargo, con esta estrategia obtuvo sólo veinte nombres útiles, por lo que la investigadora tuvo que adoptar un procedimiento menos sistemático, poniendo anuncios en los lugares más diversos (periódicos locales, guarderías infantiles, iglesias, empresas con mano de obra femenina, etc.), y sólo de esta forma logró alcanzar su objetivo de 60 entrevistas.

Otras veces, sobre todo cuando se está estudiando una población «clandestina», se recurre a la técnica del muestreo «en alud». Barbagli [1995], en su investigación sobre la criminalidad, se propuso entrevistar a autores de delitos (en particular, ladrones y peristas), personas que por motivos obvios resulta difícil localizar (salvo que se recurra a los reclusos, que ya de por sí representan una selección de la población de referencia, con el problema añadido de que la situación de detención dificulta una entrevista sincera).

A pesar de estas dificultades, el autor consiguió entrevistar a 65 autores de delitos, localizados con un muestreo «de bola de nieve», a partir de conocidos personales de los dos entrevistadores que el autor utilizó, que en el pasado habían estado en contacto directo con el mundo de la delincuencia.

- *Criterio centrado en el sujeto frente a criterio centrado en las variables.* De lo expuesto en los puntos anteriores se infiere que el investigador está interesado en comprender los hechos sociales a partir de una lectura globlal de los fenómenos y, sobre todo, de los sujetos estudiados. Ya hemos considerado detenidamente esta cuestión en el capítulo 2, por lo que aquí simplemente lo citaremos. En el análisis de los datos procedentes de entrevistas cualitativas, los individuos no se fragmentan en variables para después estudiar dichas variables. No se generan frecuencias ni correlaciones, (puede hacerse, pero será un producto secundario del análisis), sino que se *reconstruyen historias.* Como ya hemos dicho, el objetivo no es entender las relaciones entre las variables, sino comprender las manifestaciones en su individualidad: la categoría del nexo causal se sustituye por la de la experiencia vivida. Se busca reconstruir modelos, tipologías, secuencias, a partir de casos analizados en su totalidad. Y, en consecuencia, el punto de partida (tanto para el análisis de los datos como para la reflexión teórica) lo constituye el individuo, no la variable.

10.3. TIPOS DE ENTREVISTA

Podemos clasificar las entrevistas según su grado de estandarización, es decir, el grado de libertad o restricción que se concede a los dos actores, el entrevistador y el entrevistado. Ya hemos hecho una distinción entre los instrumentos estandarizados (cuestionario) y los instrumentos no estandarizados (entrevista) en la recogida de datos mediante preguntas. Pero es necesario profundizar en esta distinción, en par-

ticular por lo que respecta a la entrevista. Ésta se caracteriza por ser un instrumento abierto, flexible, adaptable a los distintos contextos empíricos, que se puede moldear en el curso de la interacción. Pero la flexibilidad puede presentar varios grados. Por tanto, distinguiremos tres tipos básicos: entrevistas estructuradas, semiestructuradas y no estructuradas. Más adelante consideraremos también ciertos tipos especiales de entrevista.

10.3.1. Entrevistas estructuradas, semiestructuradas y no estructuradas

• *Entrevista estructurada.* En este tipo de entrevistas se hace las mismas preguntas a todos los entrevistados *con la misma formulación y en el mismo orden.* El «estímulo» es, por tanto, igual para todos los entrevistados. Éstos, sin embargo, tienen plena libertad para responder como deseen. En definitiva, se trata de un *cuestionario de preguntas abiertas.* Cabe decir que, si bien la pregunta no restringe la libre manifestación del entrevistado, y aunque el entrevistador procure «dejar hablar al entrevistado», el simple hecho de plantear las mismas preguntas en el mismo orden a todos los entrevistados introduce un elemento de rigidez en la dinámica de la entrevista.

Como el lector ya sabe, en este manual hemos distinguido claramente entre métodos cuantitativos y métodos cualitativos, subrayando su adhesión respectiva a dos paradigmas distintos, de lo que se deriva su escasa compatibilidad. A pesar de ello, se puede decir que la entrevista estructurada es una técnica que intenta mediar entre ambos métodos. La predeterminación de la pregunta plantea serias limitaciones al objetivo de flexibilidad y adaptabilidad a la situación específica analizada, uno de los presupuestos del criterio cualitativo. En cambio, la libertad en las respuestas responde a los principios de este paradigma. Sin embargo, como veremos a continuación, las respuestas abiertas se pueden codificar en una matriz de datos, hecho que hemos usado como criterio discriminante para decidir si una técnica se puede considerar cuantitativa o no. Al mismo tiempo estas respuestas se analizan en su totalidad como contenidos, sin enmarcarlas en categorías generales.

Por tanto se trata de una técnica más bien híbrida, que garantiza en parte esa estandarización de la información recopilada necesaria en el «contexto de la justificación», y en parte esa apertura hacia lo desconocido e imprevisto que se incluye en el «contexto del descubrimiento». El intento de cumplir ambos objetivos de manera simultánea hace que el instrumento tenga una utilidad moderada para cada uno de ellos; la entrevista estructurada tiene una capacidad de estandarización menor que el cuestionario (con preguntas cerradas), y al mismo tiempo no permite profundizar tanto como una entrevista no estructurada. Pero, precisamente por su carácter bivalente, la entrevista estructurada es el instrumento idóneo cuando se quiere mantener el objetivo de estandarización de los resultados, es decir, «recoger datos» para describir también cuantitativamente una determinada situación social, pero al mismo tiempo el fenómeno estudiado no se conoce lo suficiente como para poder utilizar un cuestionario clásico de preguntas cerradas.

de las situaciones, cada una distinta de la otra, que impide la determinación de

un exhaustivo arco de respuestas antes de la administración de la entrevista. Tómese como ejemplo una investigación llevada a cabo en Gran Bretaña en los años sesenta-setenta, que tenía por objeto las costumbres y modalidades de crianza y educación de los hijos por parte de una muestra de 700 madres en el área urbana de Nottingham, entrevistadas en tres momentos sucesivos: cuando el niño tenía un año, después cuando tenía 4 años y por último cuando tenía 7 años. En la Tabla 10.1 hemos reproducido un breve extracto de la tercera oleada de entrevistas, sobre el tema de la disciplina y de los castigos. Por ejemplo, la pregunta «¿Cuáles son las cosas sobre las que no estáis de acuerdo tú y tu hijo, las que crean tensiones entre vosotros?», dirigida a descubrir los principales puntos de fricción entre madre e hijo, no puede ser «cerrada» previamente, ya que la gran variedad de situaciones impide prever todos los posibles puntos de enfrentamiento entre madre e hijo.

Otras veces se escoge esta solución no tanto porque la problemática sea poco conocida para el investigador, sino porque evoca una *complejidad de dimensiones* tal que las respuestas resultan totalmente imprevisibles (además de difícilmente sintetizables en pocas palabras). Preguntas tales como «¿Cuáles eran los mayores defectos de su madre?», o bien, para volver a la investigación citada anteriormente, «¿Qué opina del hecho de dar alguna bofetada a un niño de esta edad?», presentan tal variedad y complejidad de respuestas posibles que resulta complicado elaborar una lista exhaustiva (salvo que se presente una retahíla infinita de alternativas).

Por último, en muchos casos la decisión de servirse de la entrevista estructurada se asienta en la *cultura del entrevistado,* poco acostumbrado al lenguaje abstracto. Piénsese, por ejemplo, en una investigación realizada sobre mujeres ancianas y campesinas, con el objetivo de reconstruir su vida en la época en que contrajeron matrimonio (encuentro con su futuro marido, noviazgo, dote, banquete y viaje de novios, convivencia con las familias de procedencia, etc.). O bien en una investigación sobre niños en edad de enseñanza primaria, para reconstruir su relación con la escuela (véase en la Tabla 10.1 un extracto de esta segunda entrevista).

Ante este especial tipo de entrevistados, no es difícil conducir una entrevista que adopte una forma parecida a la de una conversación natural, con preguntas seguidas de respuestas libres; mientras que resultaría mucho más difícil una entrevista que adopte la forma innatural de preguntas con respuestas a escoger de entre un abanico predefinido. Nótese que mientras que la pregunta puede ser fácilmente comprendida (por ejemplo, la pregunta «¿Sabrías decirme por qué se va al colegio?» de la Tabla 10.1), la comprensión de las alternativas de respuesta puede ser mucho más difícil, puesto que exige una alta capacidad de abstracción; además, puede resultar que las alternativas de respuestas predefinidas hayan sido formuladas conforme a categorías mentales alejadas de las del entrevistado. Y, por último, si la pregunta es cerrada, el entrevistado debe hacer un esfuerzo de comprensión mucho mayor, ya que debe comprender tanto la pregunta como todas las respuestas alternativas que se le proponen.

• *Entrevista semiestructurada.* En este caso el entrevistador dispone de un «guión», que recoge los temas que debe tratar a lo largo de la entrevista. Sin embar- Podemos distinguir tres motivos que inducen al investigador a hacer uso de las

entrevistas estructuradas. El primero sería la *singularidad extrema de las situaciones,* muy distintas entre sí, que impide la formulación de una serie exhaustiva de posibles respuestas antes de la realización de la entrevista. Como ejemplo consideremos una investigación realizada en el Reino Unido en los años sesenta-setenta, para estudiar las costumbres y formas de educación de los hijos en una muestra de 700 madres en el área urbana de Nottingham, que fueron entrevistadas en tres momentos distintos: cuando el hijo tenía un año, cuando tenía cuatro años y por último cuando tenía siete años. En la tabla 10.1 hemos reproducido un breve extracto de la tercera oleada de entrevistas, sobre el tema de la disciplina y los castigos. Por ejemplo, la pregunta «¿cuáles son los temas sobre los que no están de acuerdo, que crean tensión entre usted y su hijo?», planteada con el fin de descubrir los puntos de fricción principales entre madre e hijo, no se puede «cerrar» previamente, ya que la gran variedad de situaciones impide prever todos los posibles temas que generan enfrentamiento entre ambos.

Otras veces se escoge esta solución, no porque la problemática sea poco conocida para el investigador, sino porque evoca tal *complejidad de dimensiones* que las respuestas resultan totalmente imprevisibles (y difíciles de sintetizar en pocas palabras). Las posibles respuestas a preguntas como «¿cuáles eran los mayores defectos de su madre?», o, volviendo a la investigación citada anteriormente, «¿le daría una bofetada a un niño de esta edad?», son tan variadas y complejas que resulta complicado elaborar una lista exhaustiva, salvo que se presente una serie interminable de opciones.

Por último, en muchos casos la decisión de utilizar la entrevista estructurada viene determinada por la *cultura del entrevistado,* poco acostumbrado al lenguaje abstracto. Pensemos, por ejemplo, en una investigación sobre mujeres ancianas y campesinas, con el objetivo de reconstruir su vida en la época en que contrajeron matrimonio (encuentro con su futuro esposo, noviazgo, dote, boda y viaje de novios, convivencia con las familias respectivas, etc.). O bien en una investigación sobre niños en edad escolar, para reconstruir su relación con la escuela (véase un extracto de esta segunda entrevista en la tabla 10.1).

Ante este tipo especial de entrevistados, no es difícil realizar una entrevista similar a una conversación natural, con preguntas seguidas de respuestas libres, y sin embargo resultaría mucho más difícil realizar una entrevista poco natural que constara de preguntas con respuestas a elegir entre una serie de opciones predeterminadas. Notemos que, mientras que la pregunta se puede comprender fácilmente (por ejemplo, la pregunta «¿sabrías decirme por qué se va al colegio?» de la tabla 10.1), la comprensión de las opciones de respuesta puede ser mucho más difícil, puesto que exige una elevada capacidad de abstracción; además, puede resultar que las opciones de respuestas predefinidas se hayan formulado según categorías mentales muy distintas de las que maneja el entrevistado. Por último, si la pregunta es cerrada, el entrevistado debe hacer un esfuerzo de comprensión mucho mayor, ya que debe comprender tanto la pregunta como todas las opciones de respuesta que se le proponen.

• *Entrevista semiestructurada.* En este caso el entrevistador dispone de un «guión», con los temas que debe tratar en la entrevista. Sin embargo, el entrevista-

dor puede decidir libremente sobre el orden de presentación de los diversos temas y el modo de formular las preguntas. En el ámbito de un tema determinado, el entrevistador puede plantear la conversación de la forma que desee, plantear las preguntas que considere oportunas y hacerlo en los términos que le parezcan convenientes, explicar su significado, pedir al entrevistado que le aclare algo que no entiende o que profundice sobre algún aspecto cuando lo estime necesario, y establecer un estilo propio y personal de conversación.

El guión del entrevistador puede ser más o menos detallado. Puede ser una lista de temas a tratar, o puede formularse de manera más analítica en forma de preguntas, aunque de carácter más general. En la tabla 10.2 hemos reproducido algunos extractos de dos guiones de entrevista estructurada. El primero se ha utilizado en la investigación ya citada *El activista de partido* (Alberoni *et al.*, 1967); esta parte de la entrevista trata sobre el compromiso político del activista, y el guión para el entrevistador contiene sólo los títulos de los temas que deben tratarse en el curso de la entrevista. El segundo guión ha sido extraído de la investigación *Getting a Job* (Buscando trabajo), que tenía por objeto estudiar las estrategias utilizadas por los sujetos para buscar trabajo, para destacar la importancia de la red de contactos personales y, en particular, los «débiles» y menos estructurados (Granovetter, 1995).

Esta forma de realizar la entrevista concede amplia libertad tanto al entrevistado como al entrevistador, y garantiza al mismo tiempo que se van a discutir todos los temas relevantes y se va a recopilar toda la información necesaria. El guión de la entrevista establece un perímetro dentro del cual el entrevistador decide no sólo el orden y la formulación de las preguntas, sino también si se va a profundizar en algún tema y, en su caso, en cuál de ellos. En general, el entrevistador no abordará temas que no estén previstos en el guión, pero tiene libertad para desarrollar temas que vayan surgiendo en el curso de la entrevista y que considere importantes para comprender al sujeto entrevistado, aunque no las incluya en el resto de las entrevistas.

- *Entrevista no estructurada.* En la entrevista estructurada las preguntas planteadas por el entrevistador se establecen de antemano, tanto en la forma como en el contenido; en la entrevista semiestructurada se establece de antemano el contenido, pero no la forma de las preguntas; por último, en la entrevista no estructurada, ni siquiera el contenido de las preguntas se fija previamente, y éste puede variar en función del sujeto.

La especificidad de la entrevista no estructurada está en la singularidad de los temas y el proceso de la entrevista. El único objetivo del entrevistador es plantear los temas que desea abordar, en el curso de la conversación; se hablará de la educación de los niños más que de consumo, de la participación en la vida política de la ciudad más que del sentido de lo sagrado; a continuación, el entrevistador dejará que el sujeto desarrolle su visión del asunto y mantenga la iniciativa de la conversación, limitándose a animarlo o incitarlo a que profundice cuando toque temas que parezcan interesantes. Además de esta función de estímulo, el entrevistador desempeña una función de control, limitando las divagaciones excesivas, procurando que la entrevista no se desvíe hacia cuestiones que no guardan conexión con el tema

estudiado. En una entrevista sobre el trabajo podremos hablar ampliamente sobre aspectos técnicos, interaccionales, salariales, profesionales, según por dónde lleve el entrevistado la conversación y en qué aspectos ponga el énfasis, pero en cualquier caso se hablará de trabajo y no, por ejemplo, de mujeres: si el entrevistado llevara la conversación por esos derroteros, el entrevistador deberá reconducirla hacia el tema principal.

Si bien el tema central de la conversación se establece de antemano, en el curso de la propia entrevista pueden ir aflorando temas complementarios no previstos, que resultan ser oportunos y relevantes. Por lo tanto, una entrevista puede extenderse más sobre un tema determinado, y otra sobre un tema distinto. No todas las entrevistas tienen por qué tratar los mismos temas. Algunos entrevistados tienen más cosas que contar, otros, menos. Algunos son más comunicativos y abiertos, y otros, más recelosos y cerrados. Además, la relación empática que se establece en el curso de la entrevista varía según los casos: algunos sujetos congenian mejor con el entrevistador, éste les inspira confianza y se atreven a hacer confidencias y reflexiones personales, pero otras veces esta confianza no está presente. Por tanto, este tipo de entrevistas varía mucho entre los distintos sujetos, y al final todas resultarán muy distintas entre sí, desde los temas abordados y hasta la extensión de las mismas.

Un ejemplo interesante de utilización de entrevistas no estructuradas es la investigación de Judith Wallerstein y Joan Kelly [1980] sobre los procesos de adaptación de padres e hijos tras un divorcio. Se trata de coloquios casi clínicos, durante los cuales las entrevistadoras tenían que abordar una larga serie de temas de fondo, como la situación económica y laboral durante el matrimonio, el origen del fracaso de la relación, la dinámica de la decisión de divorciarse, la preparación de la familia para el divorcio, las relaciones de los progenitores con los hijos durante el matrimonio y después del divorcio, la situación económica y laboral actual. Pero el modo de examinarlos y la atención dedicada a cada uno de ellos son elementos que se definen durante la propia entrevista, y derivan del curso natural de la misma más que de las exigencias establecidas por el entrevistador.

Conviene añadir que, en estas entrevistas, con el fin de dejar al entrevistado la máxima libertad para expresarse y extenderse sin demasiadas restricciones, los tiempos de la conversación los suele marcar el entrevistado: la entrevista puede llegar a durar varias horas, realizarse en varias sesiones e incluso podría tener las características de un verdadero relato autobiográfico.

La elección de uno de estos tres instrumentos depende de los *objetivos de la investigación* y las *características del fenómeno estudiado*. En el primer ejemplo presentado (sobre la relación madre-hijo), en lugar de la entrevista estructurada, los investigadores podrían haber elegido el instrumento de la entrevista no estructurada, dejando que las madres contaran libremente los diversos aspectos de su relación con el hijo. Pero como los investigadores también querían cuantificar ciertos fenómenos sociales, relacionar las variables, por ejemplo, analizar las diferencias existentes según la clase social de la madre, o el comportamiento de las madres según el género del hijo en cuestión, se prefirió la técnica de la entrevista estructurada, que permite abordar un mayor número de casos (en la investigación citada se entrevistó hasta 700 madres) y realizar comparaciones entre ellos gracias al hecho de que las

Tabla 10.1. Ejemplos de entrevistas estructuradas y posterior codificación

a) *Extracto de entrevistas a 700 madres sobre la relación con sus hijos de siete años* [Newson y Newson, 1978]

156. ¿Qué sucede cuando no están de acuerdo? ¿Cuáles son los temas que crean tensión entre usted y N?
157. ¿Su hijo/a N no le hace caso cuando usted le pide que haga algo que no le apetece hacer?
158. ¿Cómo reacciona usted cuando se comporta de esta forma?
159. ¿Qué sucede si él/ella se niega a hacer algo que usted quiere que haga?
160. ¿Le daría una bofetada a un niño de esta edad?
161. ¿Usted opina que los padres deberían evitar dar bofetadas, o piensa que dar bofetadas también puede servir para educar a los hijos?
162. En general, ¿piensa usted que una bofetada de vez en cuando da buenos resultados con los niños de esta edad?
163. ¿No cree que puede tener algún efecto negativo?
164. ¿Qué efecto le causa dar alguna bofetada a N? ¿Se siente desahogada o alterada? (¿O es algo que no le produce ninguna emoción?)

Codificación de las preguntas 160-163
1. En líneas generales aprueba el uso de las bofetadas.
2. Piensa que es poco deseable pero necesario.
3. En líneas generales desaprueba el uso de las bofetadas.

Codificación de la pregunta 164
1. Ninguna emoción, es un acto rutinario.
2. Se siente desahogada.
3. Se siente culpable y alterada.

b) *Extracto de entrevistas a niños* [Capecchi y Livolsi, 1973]
1. ¿Te gusta ir al colegio?
2. ¿Cuáles son las cosas que más te gustan del colegio?
3. ¿Cuáles son las cosas que menos te gustan del colegio?
4. ¿Sabrías decirme el peor defecto y la mejor cualidad de tu maestro?
5. ¿Sabrías decirme por qué se va al colegio? (averiguar si es por enseñanza-cultura, para aprender un oficio, por obligación)
6. ¿Hay algo que te gustaría hacer en el colegio y que ahora no puedes hacer?
7. ¿Hay cosas que te obligan a hacer en el colegio y que si pudieras no harías?

Codificación
Actitud hacia el colegio
1. Rechazo al colegio (con motivaciones concretas).
2. Escasa participación o ninguna.
3. Aceptación pasiva del colegio.
4. Aceptación motivada del colegio (especifica los elementos por los que le gusta ir al colegio).

Tabla 10.2. Extractos de los «guiones» de entrevistas semiestructuradas

a) *Entrevistas a activistas de partido* [Alberoni *et al.*, 1967]

Parte sobre «el compromiso político del activista».

1. La afiliación al partido:
   — Imagen del partido.
   — Cómo conoció el activista el partido.
   — Tipo de adhesión (progresiva o por crisis).
   — Las sucesivas fases del compromiso.
2. El adoctrinamiento:
   — Fuentes de formación (cursos, congresos, obras fundamentales, documentos oficiales, etc.).
   — Necesidad de preparación doctrinal.
   — La elección de la tendencia.
3. El proselitismo:
   — Campo de acción.
   — Ocasiones.
   — Instrumentos y modalidades.
   — Necesidad y significado del compromiso de proselitismo.
4. Tiempos, lugares y modalidades del compromiso político:
   — Durante el día.
   — Durante la semana.
   — Durante el año.
   — Durante el periodo electoral.
5. Características del buen activista.
6. La integración en el partido:
   — Significado funcional y recompensas al compromiso político.
   — Integración con las metas y los procedimientos del partido.
   — Las relaciones interpersonales en el partido.

Nota: Este fragmento equivale aproximadamente a 1/5 del guión completo; las otras partes se referían al ambiente familiar de origen, las experiencias personales del sujeto, la estructura del comportamiento habitual del activista, el concepto del partido y el concepto de sociedad.

b) *Entrevista sobre el acceso al mercado laboral y la carrera profesional* [Granovetter, 1995]

¿Mientras desempeñaba su trabajo anterior, decidió en algún momento buscar otro trabajo, o simplemente «sucedió»?
¿Cómo ha encontrado su trabajo actual?
Hablar de los siguientes temas: Si ha buscado trabajo de forma activa, ¿hubo métodos de búsqueda que no funcionaron?
En el periodo en que aceptó una oferta de trabajo, ¿le propusieron otros trabajos que rechazó? Si la respuesta es afirmativa, ¿cómo surgieron estas propuestas?

Cuando se incorporó al nuevo puesto de trabajo, ¿notó que estaba ocupando el puesto de otra persona que había antes? (Si la respuesta es negativa, tratar de

entender exactamente cómo y por qué se creó un nuevo puesto de trabajo).

¿Recuerda también cómo encontró el trabajo que tenía antes del actual? (Reconstruir toda la vida laboral desde el primer trabajo a tiempo completo hasta el trabajo actual. Si hay demasiadas experiencias laborales, indicar el primer trabajo a tiempo completo y uno de los más duraderos entre los intermedios; indicar también el número total de experiencias laborales).

entrevistas tienen una base común.

En este caso la elección estuvo determinada por los objetivos de la investigación. Un caso en el que la elección estuvo determinada por la naturaleza del problema estudiado es la ya citada investigación de Terry Arendell sobre mujeres divorciadas. La investigadora comenzó su trabajo de recopilación de datos con un tipo de entrevista estructurada, en la que, como ella misma explica, «el instrumento de recopilación de datos consistía en una lista de todas las preguntas que quería plantear en entrevistas formales»; pero pronto advirtió que «el uso de esta estructura tendía a interferir en el hilo del pensamiento de las mujeres entrevistadas», por lo que, partiendo de un replanteamiento de la entrevista como «proceso social con un itinerario propio no previsible», decidió adoptar un criterio no estructurado, «asegurándose sólo de que todos los problemas relevantes se trataran en algún momento de la entrevista» [1986, 184].

Es preciso recurrir a la entrevista estructurada cuando el diseño de la investigación implique entrevistar a un número elevado de personas, lo que exige un equipo de entrevistadores grande. Como veremos más adelante, la función del entrevistador en la entrevista cualitativa es muy importante y determinante, ya que el curso de la conversación depende mucho de la relación que el entrevistador establece con el entrevistado. Si hay muchos entrevistadores, hay que limitar la subjetividad de su relación obligándolos a ceñirse a un esquema común a todos, o, en última instancia, estableciendo de antemano el formato de todas las preguntas con una entrevista estructurada.

Para concluir, diremos que la diferencia entre la entrevista semiestructurada y la entrevista no estructurada es más bien difusa, y que el verdadero contraste está entre estas dos y la entrevista estructurada. Ante un guión de entrevista más bien esquemático puede resultar difícil decir si se configura una relación semiestructurada o no estructurada, mientras que la existencia de una secuencia de preguntas preestablecidas define con claridad un determinado tipo de relación de entrevista, muy controlada por el entrevistador.

10.3.2. Casos especiales

• *Entrevista no dirigida y entrevista clínica.* En los tres tipos de entrevista que hemos presentados hasta ahora, la interacción siempre estaba de algún modo controlada por el entrevistador, quien como mínimo determinaba los temas y los límites de la conversación. En la *entrevista no dirigida* no está establecido de antemano

ni siquiera el tema de conversación; el entrevistador se deja guiar por el entrevistado, que es libre para dirigir la conversación hacia donde prefiera, y el hecho de llevarla hacia un terreno concreto y no hacia otros se considera un elemento de diagnóstico. Por lo tanto, el entrevistador no sabe por dónde discurrirá la conversación, podría empezar hablando de trabajo y terminar haciéndolo de mujeres. Este tipo de entrevista tiene sus orígenes en la psicodinámica y la psicoterapia, cuyo objetivo es precisamente ayudar al paciente a sacar a la luz sentimientos profundos y en última instancia inconscientes.

La *entrevista clínica* es distinta de la anterior, dado que está guiada por el entrevistador (el psicólogo, el médico, el asistente social): el objetivo consiste en releer, mediante una entrevista en profundidad no muy distinta de la entrevista no estructurada, la biografía personal del entrevistado, reconstruyendo, en una especie de lectura a cámara lenta, el camino que lo ha llevado a una situación determinada, por ejemplo una conducta desviada (drogadicción, delincuencia, etc.).

El objetivo principal de estos dos tipos de entrevista es más terapéutico que cognitivo. Las dos se utilizan más para explorar en la personalidad del paciente que para obtener información sobre un determinado fenómeno social. Por este motivo las entrevistas no dirigidas y las entrevistas clínicas no tienen mucho interés para el sociólogo, cuyo objetivo es comprender los hechos sociales más que profundizar en la estructura de la personalidad del individuo entrevistado. Como escribe McCracken, la entrevista sociológica «se diferencia de la entrevista en profundidad realizada por el psicólogo en que aquélla se interesa más por las categorías culturales y los significados compartidos que por los estados afectivos individuales» [McCracken, 1988,7].

• *Entrevista a observadores privilegiados.* En los tipos de entrevista presentados hasta ahora el objeto de estudio son los sujetos entrevistados: si queremos estudiar a militantes de partido, entrevistamos a algunos militantes de partido, si queremos estudiar a los delincuentes, entrevistamos a algunos delincuentes. Sin embargo, podemos decidir entrevistar a sujetos, no como parte del fenómeno estudiado, sino como conocedores expertos del fenómeno, con una visión directa y profunda del mismo que los sitúa en una posición de observación privilegiada. Por ello se denominan «observadores privilegiados».

Por ejemplo, en su investigación sobre autores de delitos, Barbagli entrevistó a 65 delincuentes y a otras 28 personas más, entre las cuales abogados, magistrados, policías, empleados de entidades bancarias, etc., dado que, por sus experiencias profesionales, estaban en contacto directo con los delitos y los autores de los mismos.

De forma similar, en un estudio sobre una cárcel terapéutica del Reino Unido, una vez finalizado el tratamiento terapéutico de delincuentes con trastornos mentales, los autores entrevistaron a los 640 reclusos (la unidad de análisis principal de su estudio), y a varios miembros del personal de la prisión, como asistentes sociales, funcionarios, médicos y psicólogos, etc., que por su cargo tenían un conocimiento profundo del objeto de la investigación (Genders y Player, 1995).

Otras veces el observador privilegiado forma parte de la población objeto de

estudio, pero tiene una posición especial en la misma. Un ejemplo podría ser un líder de opinión o un líder de comunidad, al que se entrevista porque refleja las opiniones del grupo al que pertenece. También puede ser alguien que, por sus experiencias personales, tenga un conocimiento especialmente profundo del objeto de estudio: en la citada investigación sobre las cárceles se entrevistó con especial atención a los reclusos que llevaban mucho tiempo en la cárcel y habían estado en varios centros penitenciarios, por lo que tenían un conocimiento especialmente rico y variado del régimen carcelario.

Es frecuente recurrir a los observadores privilegiados en la fase preliminar de la investigación, cuando se están definiendo los límites del objeto de estudio. También se suele acudir a ellos para este mismo fin cuando el planteamiento de la investigación es de tipo cuantitativo, por ejemplo, si los datos se van a recoger mediante una encuesta por muestreo, y antes de plantear el cuestionario se quiere analizar el fenómeno estudiado preguntando a sujetos que lo conocen bien.

En la investigación de tipo etnográfico la entrevista a observadores privilegiados tiene una importancia documental fundamental. Como hemos visto en el capítulo anterior, este tipo de investigación se basa sobre todo en la observación participante, con entrevistas a los denominados «informadores», personajes del ambiente social estudiado, líderes de los grupos, etc., es decir, lo que hemos definido como «observadores privilegiados». Para este tipo de entrevista se ha acuñado incluso el término «entrevista etnográfica», pero está claro que el modelo es el de la entrevista no estructurada: el entrevistador deja hablar libremente al entrevistado, pero al mismo tiempo debe estimularlo en su función de testigo, pidiéndole explicaciones cuando algo no esté lo suficientemente claro, detalles cuando dé por supuesto algo que en realidad no lo está, para que cuente todo lo que sabe sobre el fenómeno estudiado.

Las entrevistas a observadores privilegiados son totalmente heterogéneas y no se pueden comparar entre sí: cada una es un caso en sí misma, tiene un desarrollo y un enfoque propios. Esto se debe a que los personajes entrevistados son muy distintos entre sí, y cada uno de ellos tiene una historia distinta que contar: basta considerar la lista de sujetos entrevistados de la investigación sobre delitos citada anteriormente (desde abogados a empleados de entidades bancarias), o el estudio sobre las cárceles (desde asistentes sociales al juez de vigilancia penitenciaria, etc.).

• *Entrevista de grupo*[3]. Hasta ahora hemos hablado de entrevistas individuales. Si, por ejemplo, estamos entrevistando a un activista de partido y queremos que nos cuente cómo reparte el tiempo entre la familia, el trabajo y el partido en su vida cotidiana, la presencia de su pareja en la entrevista puede constituir un elemento de perturbación. Pero en ciertos casos, la interacción, y en particular la

[3] Conviene indicar que la entrevista de grupo de la que se habla en el texto corresponde a lo que en la tradición anglosajona se denomina *focus group*. En cambio, el denominado «grupo de discusión» que se enmarca dentro de la tradición sociológica española, presenta una mayor espontaneidad y un menor grado de directividad en la conversación en comparación con la entrevista de grupo. Para una reflexión sobre el grupo de discusión, véase Callejo, 2001.

interacción de un grupo, puede permitir profundizar y favorecer la comprensión del fenómeno estudiado.

Por ejemplo, para estudiar la violencia en los estadios de fútbol puede resultar útil reunir en una habitación a diez integrantes de grupos ultras y mantener una discusión con ellos sobre el porqué de los actos de violencia. De la discusión pueden surgir motivos que no habrían salido a la luz en conversaciones individuales; además, esta discusión puede hacer aflorar la intensidad de los sentimientos, facilitando las comparaciones entre distintas posiciones. Por ejemplo, los motivos vinculados a la competición entre los distintos grupos que apoyan al mismo equipo podrían ser mucho más fuertes que los derivados de la hostilidad hacia equipos contrarios, pero este aspecto difícilmente podría aflorar en las entrevistas individuales, que permiten deducir si un motivo está más extendido que otro, pero no el grado de intensidad de cada uno de ellos, mientras que de la discusión colectiva, de los enfados de los participantes y su debate, puede emerger la fuerte carga emotiva de la oposición entre grupos que actúan en el mismo territorio.

Con frecuencia, como afirma Herbert Blumer, uno de los teóricos de la investigación cualitativa, «un número limitado de sujetos, siempre que estén bien informados y sean observadores, reunidos para discutir en grupo resulta mucho más útil que una muestra representativa. Un grupo de tales características, en el que cada sujeto comente de forma colectiva su parte de vida y se vea obligado a profundizar en ella en el contexto de una discusión y con el disentimiento de los demás, contribuirá a alzar los velos que ocultan la realidad social más que cualquier otra técnica que yo conozca» [Blumer, 1969, 41].

Las entrevistas de grupo se propusieron por vez primera en la investigación social de Merton, Fiske y Kendall [1956], quienes acuñaron el término *focused interview* (del que más tarde derivó *focus group)*. Como estos autores escriben, «en primer lugar las personas entrevistadas deben haber participado en una determinada situación: han visto una película, han escuchado un programa de radio, [...] han participado directamente en un determinado hecho social [...] (por ejemplo, una manifestación política, un rito, una revuelta)» [*ibídem*, 3]. A partir de esta experiencia común (que incluso puede ser programada por el investigador, por ejemplo, mostrando una película a un grupo de sujetos), las personas son entrevistadas en grupo por un entrevistador que ha estudiado previamente el hecho objeto de estudio y que estimula y dirige la discusión de tal forma que haga aflorar las distintas interpretaciones, las reacciones emotivas, las valoraciones críticas. Se trata, por tanto, de un debate centrado en un hecho bien delimitado.

El uso de las entrevistas de grupo en las ciencias sociales tiene una larga historia: según parece, Malinowski ya las utilizó para sus investigaciones antropológicas. Merton empleó esta técnica para estudiar los efectos de la propaganda política, y hoy día se utiliza mucho en el campo del marketing, donde se recurre a estos «grupos dirigidos» para conocer la opinión de los consumidores sobre un nuevo producto, una campaña publicitaria, etc.

Por lo general, estos grupos están formados por una decena de personas, cifra suficiente para incluir distintas posiciones, pero al mismo tiempo limitada para permitir la interacción entre todos los participantes. Los dirige un entrevistador

profesional que ha estudiado el tema objeto de discusión, lo ha analizado en detalle y modera el debate procurando desentrañar el problema en todos sus aspectos y controlar la dinámica de la interacción (impedir que un individuo o un grupo reducido monopolicen la discusión, estimular la participación de los más reservados, y no hacerlo de forma tan impositiva que impida la libre expresión de los participantes). Por lo general, las personas que forman parte del grupo deben ser más bien homogéneas (en el ejemplo de la violencia en los estadios, se tratará de grupos de *hooligans;* en una discusión sobre el modo en que los ciudadanos han percibido un mensaje del presidente del gobierno, se formarán grupos separados de trabajadores de distintas categorías, amas de casa, estudiantes, etc.); además, es imprescindible que todos los participantes conozcan el tema objeto de la discusión.

Por último, un aspecto importante de la entrevista de grupo es su rentabilidad: en el tiempo que se necesita para entrevistar a una sola persona se puede entrevistar a diez. Obviamente, el resultado es distinto: con la entrevista de grupo no se pueden obtener datos individuales. Por tanto, podemos afirmar que la herramienta de la entrevista de grupo resulta útil para comprender el fenómeno estudiado, pero no para documentarnos sobre el mismo.

10.4. REALIZACIÓN DE LA ENTREVISTA

No es fácil realizar una entrevista cualitativa buena. No se trata únicamente de conseguir que el sujeto consienta en ser entrevistado y responda con sinceridad a una serie de preguntas. La parte más difícil consiste en «hacer hablar» al entrevistado, provocar un relato fluido en el que el entrevistador se limite a escuchar y a plantear de vez en cuando alguna pregunta de aclaración o a hacer alguna intervención cauta para reconducir la conversación a los temas centrales. La entrevista cualitativa debe acceder, si no a la psique profunda, al menos al mundo de las emociones y los sentimientos del sujeto entrevistado, a sus pensamientos más genuinos.

Conviene indicar que, mientras que en el coloquio psicoterapéutico este acceso a la psique profunda del sujeto suele ser más fácil, dado que ha sido el propio paciente quien lo ha solicitado, en el caso de la entrevista con fines cognitivos, el acceso a una simple confidencia del entrevistado puede resultar muy difícil. Para obtener la plena colaboración del sujeto, el entrevistador debe lograr establecer una relación de confianza con el mismo, no como profesional, sino como persona, y esto no resulta sencillo en una relación muy limitada en el tiempo, que con frecuencia consiste en un solo encuentro, a veces sin preparación previa (por contraste con la interacción psicoterapéutica, que se desarrolla a lo largo de un largo período y en múltiples encuentros).

La entrevista cualitativa no es un procedimiento de obtención de información, sino un proceso de interacción social entre dos individuos. Por ello no resulta fácil dar reglas de actuación generales, dado que el desarrollo de la entrevista depende en gran parte de aspectos subjetivos. Por este motivo nos limitaremos a indicar algunas sugerencias.

• *Explicaciones previas.* En el capítulo 5 hablamos del problema del contacto con el entrevistado y la autorización a realizar la entrevista. Ya señalamos que el problema principal es la desconfianza inicial del entrevistado y su inseguridad ante una situación nueva y desconocida. En el caso de la entrevista cualitativa este problema es menor, ya que la entrevista se suele preparar mejor: el entrevistador no es un desconocido que llama a la puerta, sino que ha habido un proceso de acercamiento previo (carta de presentación, llamada telefónica solicitando una cita, presentación por parte de un conocido, etc.). Este acercamiento es posible porque el número de contactos necesarios para la entrevista cualitativa es menor que en el caso del cuestionario, y por el hecho de que al tratarse de una entrevista en profundidad, es imprescindible resolver todas las posibles dudas y motivos de desconfianza, ya que de lo contrario se corre un elevado riesgo de no cumplir los objetivos perseguidos.

Por tanto, el problema con el que nos enfrentamos en el contacto inicial no estriba tanto en superar la desconfianza del entrevistado como en *hacerle entender qué es lo que queremos de él.* Será necesario describir con detalle el objetivo de la investigación, explicarle por qué ha sido seleccionado para la entrevista y por qué le haremos determinadas preguntas que podrían ser personales, justificar en su caso la grabación de la conversación, etc.

• *Preguntas primarias.* Según la distinción que Kahn y Cannel [1967: 205] hacen entre «preguntas primarias» y «preguntas secundarias», las primeras sirven para introducir un tema nuevo o para abrir un nuevo interrogante, y las segundas tienen como finalidad estructurar y profundizar en el tema de la pregunta primaria. Utilizando una clasificación de Spradley [1979, 59], referida a lo que él llama entrevista etnográfica pero que puede aplicarse a cualquier entrevista cualitativa, podemos distinguir tres tipos de preguntas primarias. En primer lugar, las *preguntas descriptivas,* como, «¿me puede hablar de su trabajo?», «¿cómo se desarrolla su jornada laboral?». En segundo lugar están las *preguntas estructurales,* cuyo objetivo es descubrir la forma en que el entrevistado estructura su conocimiento, por ejemplo, «¿me puede decir qué tipo de personas trabajan para usted?», «¿de qué manera se puede ascender profesionalmente en su empresa?». Por último, tenemos lo que Spradley denomina *preguntas de contraste,* basadas en una comparación, por ejemplo, «¿qué diferencia hay entre un tractorista y un motorista?», «¿en qué sentido su trabajo actual es mejor que el anterior?».

• *Preguntas exploratorias.* Una entrevista no consiste simplemente en plantear una pregunta y registrar la respuesta correspondiente. La parte fundamental del bagaje técnico del entrevistador está formada por una serie de instrumentos que le permiten descubrir y poner de relieve las opiniones reales del entrevistado, incluso las más ocultas y menos explícitas. Un ejemplo típico de ello sería el uso de las denominadas preguntas exploratorias: éstas no son preguntas propiamente dichas, sino estímulos que tratan de ser neutrales y cuya función es animar al entrevistado a seguir con su relato, a no estar a la defensiva, a profundizar en la cuestión, a dar más detalles. Se trata de intervenciones no impositivas, mediante las cuales el entrevistador intenta hacer hablar al sujeto y estimular su iniciativa, sin influir en

él, dejando que sea el propio entrevistado quien escoja los temas en los que desea profundizar y la forma de exponerlos, siguiendo el hilo de su discurso, más que respondiendo a las preguntas. Podemos distinguir varias maneras de formular estos estímulos.

- *Repetición de la pregunta.* A veces basta con repetir la pregunta, formulándola de forma distinta, o haciendo hincapié en determinadas palabras.

- *Repetición de la respuesta o de una síntesis de las últimas respuestas (acción «eco»).* En este caso se trata de retomar las últimas respuestas del entrevistado, con sus palabras o con un resumen realizado por el entrevistador, para invitarlo a profundizar en ellas o aclararlas. Por ejemplo: «entonces, me ha dicho que...» (sigue la repetición literal de la frase), «¿pero qué entiende usted realmente por...?». O bien, «veamos, si lo he entendido bien usted me ha dicho que...» (seguido de un resumen de las últimas respuestas), «¿hay algún otro motivo?», «¿tiene algo más que decir al respecto?».

- *Estímulo, expresión de interés.* En este caso el entrevistador se limita a manifestar interés, con expresiones verbales o incluso con simples gestos, o con interjecciones de aprobación, en relación con todo lo que el entrevistado está relatando. Para indicar su interés el entrevistador puede utilizar también un «¿sí...?», que puede ser suficiente para animar al entrevistado a que prosiga con su narración. Otras expresiones neutras para obtener el mismo efecto son, por ejemplo: «esto es muy interesante», «continúe», «entiendo», «¿de verdad?».

- *Pausas.* A veces dejar pasar unos segundos de silencio, sin preocuparse de llenarlo enseguida por temor a crear malestar, puede favorecer la confianza por parte del entrevistado; otras veces esa incomodidad creada por la pausa hace que el entrevistado se sienta obligado a romper el silencio retomando su narración y dando más detalles.

- *Pedir al entrevistado que profundice.* A veces es necesario pedir explícitamente al entrevistado que interrumpa la narración para aclarar o profundizar sobre alguna cuestión. Ejemplos de intervenciones de este tipo serían, «¿puede explicármelo más detenidamente», «me gustaría conocer mejor su opinión al respecto», «no estoy seguro de haber entendido bien lo que piensa usted sobre este tema», «¿por qué reaccionó usted de ese modo?».

- *Lenguaje.* En el cuestionario el lenguaje tenía que ser claro y accesible, ya que la estandarización del instrumento exigía una selección cuidadosa de las formulaciones que resultaran comprensibles para todos los sujetos. En la entrevista cualitativa el problema del lenguaje no se plantea en estos términos, ya que el estilo lingüístico de las preguntas puede variar en función del sujeto, adaptándose a las características del entrevistado. El problema del lenguaje reside en que éste es el instrumento fundamental para establecer un clima de empatía, y por tanto de comunicación, entre el entrevistador y el entrevistado. Según Kahn y Cannel:

> El lenguaje proporciona al entrevistador y al entrevistado indicios para entender qué tipo de persona es su interlocutor. Las primeras indicaciones que tiene el entrevistado para saber si el entrevistador es una persona muy parecida o muy distinta de él mismo, las obtendrá precisamente del lenguaje que el entrevistador utilice para presentarse y para presentar el tema de la entrevista [...]. Si el entrevistador y el entrevistado «hablan el mismo lenguaje», es muy probable que tengan experiencias parecidas, y por tanto es mucho más probable que puedan entenderse mutuamente.
>
> [Kahn y Cannel, 1967; 111].

Obviamente, no siempre es necesario (ni posible) que el entrevistado y el entrevistador hablen el mismo lenguaje. A menudo el entrevistador se presenta con la apariencia y el lenguaje propios de un investigador y es aceptado como tal, siempre que se muestre realmente interesado y el entrevistado pueda percibirlo como alguien que sabe comprenderlo. Aun así pueden darse situaciones en las que reine un clima de desconfianza o de clara diferencia con el mundo del entrevistado, que sólo se puede superar si el entrevistador se percibe como una persona que puede entender ese mundo porque de alguna forma lo ha vivido personalmente o al menos ha estado cerca del mismo. Por ejemplo, en una investigación sobre drogadicción juvenil, un entrevistador que de alguna manera haya estado en contacto con esos ambientes, conozca su jerga y entienda fácilmente las situaciones que se producen en los mismos, tiene ventaja con respecto a un entrevistador que se presente con el estilo propio de un psicólogo.

- *Papel del entrevistador.* De todo lo dicho hasta ahora se deduce que en la entrevista cualitativa el entrevistador desempeña un papel no sólo fundamental, sino también *determinante* (en el sentido literal de que «determina» el resultado de la conversación). Conviene recordar brevemente lo dicho en el capítulo 5 sobre el enfoque *constructivista* en la investigación social, es decir, que en el paradigma interpretativo la entrevista se entiende *no* como una operación de «recopilación de datos», en la que el entrevistador registra las opiniones del entrevistado, sino como una relación dinámica en la que la entrevista es «construida» de manera conjunta por el entrevistador y el entrevistado, y su resultado depende en gran medida del vínculo empático que se haya instaurado entre los dos interlocutores.

  En este contexto la función del entrevistador resulta compleja y presenta ciertas contradicciones. Por una parte, debe desempeñar una función activa de orientación de la entrevista, procurando que la comunicación cumpla sus objetivos: el entrevistador «incentiva de forma sistemática las respuestas completas y exhaustivas [...] que son relevantes en relación con los objetivos de la entrevista, y tiende a desalentar las comunicaciones no relevantes para dichos objetivos» [Kahn y Cannel, 59]. Pero, al mismo tiempo, debe procurar que estas actitudes no alteren el pensamiento del entrevistado: al fomentar el tratamiento de temas más afines a los objetivos de la investigación, el entrevistador puede provocar que el entrevistado no manifieste sus sus verdaderos sentimientos, llegando a orientar sus respuestas.

  De todo lo expuesto se desprende claramente que la labor del entrevistador

cualitativo es mucho más laboriosa que la del entrevistador que utiliza un cuestionario. Al primero se le exige mucho más que cumplir fielmente una serie de instrucciones detalladas. Se le pide sensibilidad, intuición, capacidad de identificarse con la personalidad del interlocutor, experiencia con las relaciones humanas y, no menos importante, un conocimiento profundo del problema objeto de estudio. Cuanto más abierto y desestructurado sea el instrumento de obtención de material empírico, más capacidades deberá tener el entrevistador. En última instancia, en la entrevista no estructurada será el propio investigador (o los investigadores del equipo de investigación) quien realice las entrevistas directamente, sin utilizar entrevistadores, o se servirá de ellos de forma muy limitada, trabajando en estrecha colaboración con ellos durante toda la fase de realización de las entrevistas. Veamos algunos ejemplos extraídos de las investigaciones que hemos presentado anteriormente.

En la investigación sobre los activistas de partido [Alberoni *et al.*, 1967], las entrevistas (108 en total) fueron realizadas por doce investigadores en el sentido más estricto de la palabra, es decir, miembros del equipo de investigación (algunos de ellos fueron los redactores del libro en el que se presentaron los datos de la investigación, otros elaboraron ensayos a partir del estudio para diversas publicaciones). En su investigación sobre la criminalidad [Barbagli, 1995], el autor realizó personalmente las entrevistas a los 28 observadores privilegiados (magistrados, abogados, policías, etc.), mientras que para las otras 65 entrevistas realizadas a sujetos autores de delitos, dado que se trataba de un tema delicado, utilizó a dos entrevistadores que en alguna ocasión habían estado en contacto con ambientes delictivos. No obstante, los entrevistadores trabajaron en estrecha colaboración con el responsable de la investigación (que escuchaba la grabación de las entrevistas inmediatamente después de su realización, las comentaba con el entrevistador y juntos decidían si continuar la entrevista en un contacto posterior, determinando los puntos que quedaban pendientes). En la investigación que realizaron Wallerstein y Kelly [1980] sobre los procesos de adaptación de 60 parejas que se habían divorciado, las autoras llevaron a cabo las 120 entrevistas a los cónyuges y las 136 a los hijos, con ayuda de otras cinco entrevistadoras. Arendell [1986], en su investigación sobre mujeres divorciadas, entrevistó personalmente a las 60 mujeres estudiadas.

Esta norma de realizar la entrevista personalmente, o con ayuda de un número limitado de colaboradores que pueden considerarse coinvestigadores más que entrevistadores, se aplica sobre todo en las entrevistas semiestructuradas o no estructuradas. El criterio es menos vinculante en el caso de las entrevistas estructuradas, cuando la estandarización de las preguntas reduce la libertad del entrevistador y permite encargar las entrevistas a personal ajeno al equipo de investigación, lo que (a su vez) implica incluso que las muestras obtenidas tienen dimensiones considerables.

Por ejemplo, en la investigación sobre la relación madre-hijo citada anteriormente, los autores utilizaron un número bastante numeroso de entrevistadoras, que en su mayoría fueron visitantes médicos del City Health Department de Nottingham. Otras veces, en cambio, incluso en los casos de entrevistas estructuradas, los investigadores prefieren trabajar con un número muy reducido de entrevistadores, que trabajan en estrecha colaboración. En la investigación mencionada previamente sobre los niños y

el colegio (Capecchi y Livolsi, 1973), en la cual se entrevistó a 144 chicos de edades comprendidas entre los siete y los trece años, sus padres y su profesor, todas las entrevistas fueron realizadas por sólo seis entrevistadores. Un caso bastante excepcional en comparación con el resto de ejemplos propuestos es la investigación de Granovetter (1995) sobre las estrategias utilizadas para encontrar trabajo: aunque la muestra de entrevistados fue bastante numerosa ($N = 282$), todas las entrevistas las realizó personalmente el propio investigador.

De todas estas premisas, en concreto, la condición de que haya un contacto personal y profundo entre el entrevistado y el entrevistador, se deduce que las entrevistas cualitativas sólo pueden realizarse mediante una relación cara a cara (las entrevistas telefónicas, postales, telemáticas y similares sólo pueden servir para realizar cuestionarios). Además, como regla general, éstas deben grabarse siempre. La alternativa de transcribir un resumen de la respuesta durante la entrevista deberá evitarse, al menos en la entrevista estructurada, pues da lugar a síntesis incompletas, poco profundas y a veces incomprensibles. La grabación, en cambio, permite conservar la narración del entrevistado en su forma original y completa. A la hora de transcribir las cintas se recomienda seguir siempre un criterio de *transcripción integral,* incluyendo incluso las variaciones dialectales, los errores de sintaxis, las frases entrecortadas, etc., ya que precisamente estos aspectos del lenguaje hablado hacen que la entrevista sea más viva y comunicativa.

Por último debemos añadir que la utilización de una grabadora permite al entrevistador concentrarse en la conversación y mantener una relación más natural con el entrevistado, lo que resultaría prácticamente imposible si tuviera que tomar apuntes. Es cierto que en algunos casos la presencia de la grabadora podría inhibir la libre expresión del entrevistado, sobre todo cuando se trata de una persona que desempeña algún cargo oficial, pero esto no es frecuente, y aunque la presencia de la grabadora puede generar cierta incomodidad al principio, se va olvidando a medida que avanza la entrevista.

10.5. ANÁLISIS DEL MATERIAL EMPÍRICO

10.5.1. El análisis de tipo cualitativo

Después de realizar, grabar y transcribir las entrevistas llega el momento de analizarlas, interpretarlas y redactar el informe de investigación. En el capítulo 2 hablamos del análisis de los datos procedentes de entrevistas y comparamos los enfoques cuantitativo y cualitativo en las distintas fases de la investigación social. En esa exposición contrastamos repetidamente las técnicas de la entrevista cuantitativa (cuestionario) y la entrevista cualitativa para ayudar al lector a entender mejor las diferencias entre ambas. En este apartado ampliamos el análisis *cualitativo* de las entrevistas.

- *El análisis cualitativo de los datos* se centra en los sujetos y no en las variables, como sucede en la investigación cuantitativa. El criterio debe ser, por tanto, de

tipo *holístico,* es decir, el individuo es observado y estudiado en su totalidad, con la convicción de que cada ser humano (igual que cada hecho social) es algo más que la suma de sus partes (en este caso, las variables). El objetivo del análisis es *comprender a las personas,* más que analizar las relaciones entre variables.

• *La presentación de los resultados* se realiza en forma de narración, mediante relatos de episodios, descripción de casos, utilizando a menudo las mismas palabras de los entrevistados, para no alterar el material recopilado y transmitir al lector lo más fielmente posible las situaciones estudiadas. El procedimiento estándar para la presentación de los resultados es el siguiente: se expone un razonamiento, y para apoyarlo y ilustrarlo se reproduce un fragmento de entrevista. Por ejemplo, en la investigación sobre los activistas de partido se dice: la visión del mundo de los activistas del Partido Comunista de edad avanzada es estrictamente dicotómica, es decir, está basada en una distinción entre «nosotros» y «ellos», los amigos y los enemigos, el proletariado y los capitalistas. Y para ilustrarlo los autores transcriben fragmentos de entrevistas que traten, por ejemplo, sobre política internacional (EEUU y la URSS), sobre los partidos (PCI y DC), el deporte, etc.

En la investigación de Arendell, uno de los capítulos está dedicado a la situación económica de las mujeres divorciadas. La autora escribe: «Incluso las mujeres que antes de divorciarse ya habían considerado su situación económica en caso de no contar con el sueldo de su marido, ninguna habría imaginado jamás lo difícil que sería la vida después del divorcio [...]; el 90 % de ellas (56 de 60) se encontró en una situación de pobreza, o cuasipobreza. [...] Algo que no les sucedió a sus maridos [...]; el divorcio desplazó a las mujeres de la clase social a la que pertenecían cuando estaban casadas» [1986, 36-37]. Este análisis va seguido por una serie de fragmentos de entrevistas que ilustran con las palabras de las mujeres entrevistadas sus condiciones de vida, la desilusión, el desplazamiento social imprevisto, las dificultades cotidianas, el desinterés de los ex maridos, etc.

De este modo, el texto consiste en una combinación continua de análisis y explicaciones, ilustrados con ejemplos de fragmentos de las entrevistas.

• *Las síntesis y generalizaciones* suelen presentarse en forma de clasificaciones y tipologías. Ya hablamos de ello en el capítulo 2, cuando comparamos los procedimientos cuantitativo y cualitativo, y en el capítulo 9, cuando hablamos de la clasificación y las dimensiones utilizadas para la misma. A este respecto conviene hacer hincapié en que la forma de utilizar y analizar el material empírico obtenido de una investigación de tipo cualitativo es fundamentalmente unitario, y no varía según la procedencia del material (observación participante, entrevistas o material documental de otro tipo). Por este motivo no creemos necesario volver sobre estas cuestiones, y remitimos a las partes del libro en los que se trataron anteriormente.

10.5.2. El análisis de tipo mixto (cuantitativo y cualitativo)

Hay un aspecto del tratamiento del material procedente de las entrevistas que debemos considerar detenidamente. Para ello vamos a volver brevemente al resumen 5.1,

en el que afirmamos que el elemento que distingue realmente a la investigación cuantitativa de la cualitativa, la diferencia clave entre los dos criterios, es la presencia o no de la matriz de datos. En el capítulo 5 dijimos que la matriz de datos, como organización estandarizada de las respuestas, conlleva una posibilidad de clasificar y uniformizar las actitudes y los comportamientos humanos que es incompatible con el planteamiento individualista.

Sin embargo, existe una forma de recopilación de datos en la que coexisten los dos criterios: se trata de la *entrevista estructurada*. En la entrevista estructurada las preguntas se fijan de antemano y las respuestas son abiertas. Ello implica que a todos los entrevistados se les pide la misma información, pero se les deja libertad para exponerla como consideren más adecuado y para organizarla a su gusto. Esta base informativa común puede uniformizarse mediante un procedimiento de codificación.

Para ilustrarlo, volvamos a la investigación sobre las madres de Nottingham (Newson y Newson, 1978), y en particular a las preguntas incluidas en la tabla 10.1 sobre el tema de la disciplina y los castigos. La entrevista contenía varias preguntas sobre la práctica de dar bofetadas a los hijos. Como es natural, las respuestas fueron muy variadas (desde exposiciones problemáticas y analíticas a respuestas claras y concisas). Como de costumbre, estas respuestas se grabaron y se transcribieron en su totalidad; después se clasificaron según las categorías recurrentes. En este caso se recodificaron juntas las respuestas y las preguntas 160-163, dando lugar a una única variable de tres categorías (de la aprobación a la desaprobación de las bofetadas; ver la tabla 10.1). Las respuestas a la pregunta sobre la reacción emotiva de la propia madre se recodificaron también en tres categorías.

En la investigación sobre la actitud de los niños con respecto al colegio, también incluida en la tabla 10.1, se hizo algo parecido. En la entrevista había diversas preguntas orientadas a profundizar en la relación del niño con el colegio (por ejemplo: «¿te gusta ir al colegio?, ¿cuáles son las cosas que más te gustan del colegio?, etc.). La conversación se grababa. Posteriormente se codificaban las preguntas, no considerando las respuestas a cada una de las preguntas, sino asignando una valoración global a un bloque de preguntas. Por ejemplo, con respecto a las preguntas referidas en la tabla 10.1, se clasificaba al niño en una escala de cuatro puntos de «aceptación/rechazo del colegio», que abarcaba desde el rechazo total hasta la aceptación motivada. De esta forma se obtiene una matriz de datos clásica, en la que todos los sujetos estudiados se clasifican utilizando las mismas variables.

Notemos que lo que se clasifica y codifica es sólo una parte de la entrevista, que podría considerarse la parte común a todos los casos, una especie de mínimo común denominador, que podría ser muy reducido. Cada entrevista tiene además una parte específica propia, que es distinta en cada uno de los casos y que no debe perderse. De hecho, el análisis del material empírico se realiza en dos vertientes: por un lado, los datos estandarizados se analizan con los instrumentos clásicos de la estadística (distribución de frecuencias, tablas para analizar las relaciones entre variables, análisis multivariado), y por otro, los textos de las transcripciones se analizan según las modalidades características de la investigación cualitativa.

Conviene decir que este modo de cuantificar los datos cualitativos no se limita

al caso de la entrevista estructurada. Cada entrevista cualitativa puede tener una parte de temas comunes, de los que puede extraerse información relativa a todos los entrevistados. Por ejemplo, en su investigación sobre los activistas de partido [Alberoni *et al.*, 1967], los autores, al final de la recopilación de datos, pidieron a una «comisión» formada por tres jueces que evaluara los textos de las entrevistas (entrevistas semiestructuradas con un guión más bien genérico), y éstos releyeron los textos y, para cada sujeto, cumplimentaron un cuestionario de preguntas cerradas sobre los temas que habían tratado todos los entrevistadores.

Obviamente esta intervención de cuantificación tiene sentido sólo cuando se trabaja con muestras de gran tamaño, lo cual es poco frecuente en las entrevistas no estructuradas. Por ello hemos relacionado este tratamiento del material recopilado sobre todo con la entrevista estructurada, cuando la posibilidad de utilizar entrevistadores permite trabajar con muestras de tamaño aceptable para el análisis cuantitativo.

La codificación de las repuestas es un paso crucial en este procedimiento. Está claro que esta tarea sólo la puede realizar un grupo reducido de personas, que puedan trabajar en estrecha colaboración y establecer criterios de clasificación y valoración comunes. Las codificaciones casi siempre las suelen hacer los investigadores del equipo. Incluso cuando esto no es posible (por ejemplo, en una investigación sobre una muestra muy numerosa), es indispensable que haya una buena coordinación entre los investigadores (o el investigador). Además, para evitar encontrarse al final con entrevistas que no contengan los datos que se ha decidido codificar en la matriz de datos, se aconseja establecer el esquema de codificación *antes* de empezar las entrevistas, de manera que todos los entrevistadores sepan cuáles son los datos que deben solicitar a los entrevistados.

10.6. FALTA DE ESTANDARIZACIÓN E INTERACCIÓN

La falta de estandarización es al mismo tiempo el punto fuerte y el punto débil de las entrevistas cualitativas. Al imponer pocas obligaciones al *entrevistador,* le permite abordar caminos imprevistos, elaborar y explorar nuevas hipótesis sobre el fenómeno estudiado, yendo más allá de la formulación inicial del problema. Al conceder al *entrevistado* plena libertad de expresión, le permite manifestar su punto de vista utilizando sus propias categorías mentales y su propia forma de expresarse. Es, por tanto, un instrumento especialmente adecuado para el *descubrimiento* y, como ya hemos dicho, para la *comprensión.* Al mismo tiempo, la falta de estandarización hace muy difícil la valoración cuantitativa de los fenómenos estudiados y la comparación.

Por ejemplo, aunque con la entrevista cualitativa conseguimos identificar las razones, incluso las más profundas, por las que las personas están a favor o en contra de la expulsión de los inmigrantes ilegales de su país, es difícil cuantificar el peso de los distintos motivos (prejuicios raciales, conveniencia económica, legalidad, etc.), y resulta también difícil hacer comparaciones (¿quién está más a favor de la expulsión de inmigrantes, los hombres o las mujeres, los jóvenes o los adultos,

quienes tienen un nivel de estudios inferior o superior, etc.?).

Por tanto, la falta de estandarización conlleva una dificultad para hacer comparaciones (¿cómo podemos comparar a dos sujetos si no les hemos planteado la misma pregunta?), y una dificultad de síntesis (¿cómo podemos contar cuántos están en contra de la inmigración por prejuicios raciales si no hemos planteado la misma pregunta a todos ellos?).

Pero incluso en el supuesto de que se consiguiera contar y comparar, seguiría siendo difícil hacer deducciones sobre una población más amplia, debido a que número de casos estudiados es muy reducido. Como ya hemos dicho, la entrevista cualitativa, precisamente por la condición de profundidad que impone, no puede realizarse sobre muestras demasiado grandes. Una cosa es preguntar «¿está usted a favor o en contra de la expulsión inmediata de los inmigrantes ilegales?», o «¿de los problemas que le enumero a continuación, cuál ha sido el que ha provocado la crisis de su matrimonio?» (con preguntas tan concisas y con un abanico de respuestas predeterminado, se podría interrogar en poco tiempo, incluso por teléfono, a decenas de personas), y otra cosa muy distinta es profundizar sobre el problema de los inmigrantes, sobre las tendencias racistas arraigadas y quizás inconscientes (o en todo caso etnocéntricas) presentes en cada uno de nosotros, o entablar una conversación larga y estructurada sobre la crisis matrimonial del entrevistado. Estas conversaciones requieren mucho tiempo, lo que nos obliga a reducir el número de sujetos entrevistados.

A ello se añade el tiempo de preparación de las entrevistas: es necesario identificar y localizar a las personas, explicarles el objetivo de la entrevista, superar las barreras de desconfianza mediante contactos previos o una presentación por parte de intermediarios; hay que fijar las citas en lugares tranquilos donde los entrevistados puedan hablar libremente, en momentos en los que tengan tiempo suficiente para ello. Todos estos aspectos son mucho más laboriosos en el caso de la entrevista cualitativa que en el caso del cuestionario. Por ello, la primera requiere un trabajo de organización y una disponibilidad de tiempo mucho mayores. Por no hablar de los casos en los que la entrevista es tan delicada que no la puede realizar un entrevistador, sino que debe ser realizada por el propio investigador, con las consiguientes limitaciones en el número de sujetos que podrá entrevistar.

Por consiguiente, si no se estandariza *es más difícil comparar, contar* e *inferir* de la muestra estudiada generalizaciones sobre una población más amplia. Obviamente, estos inconvenientes quedan compensados de sobra por el hecho de que la entrevista cualitativa permite «llegar hasta el fondo», descubrir aspectos recónditos e imprevistos.

Pocos investigadores niegan la utilidad de la entrevista en profundidad para la investigación social. Para defender esta técnica se suelen alegar dos argumentos, muy diferentes, según se inspiren en el paradigma positivista o interpretativo. Un argumento que alegan incluso los defensores más fervientes del planteamiento cuantitativo, es que la entrevista cualitativa puede ofrecer contribuciones valiosas en la fase exploratoria de la investigación. Esta opinión ha predominado en los textos de metodología de la investigación social al menos hasta los años ochenta. En un manual de los más difundidos en los años sesenta y setenta, se puede leer:

«las técnicas informales constituyen un instrumento inigualable en el estadio piloto incluso de las encuestas por muestreo formalizadas, puesto que dan indicaciones sobre las cuestiones relevantes y sobre cómo deberían formularse las preguntas [...]. A veces éstas pueden resultar útiles, junto a las encuestas formales, para el estudio de problemas complejos» [Moser y Kalton, 19741, 301].

También se recurre a la entrevista cualitativa en la investigación social de inspiración cuantitativa para profundizar, *tras* la recogida de los datos con técnicas cuantitativas, sobre un determinado tema surgido, para darle trascendencia, color y visibilidad ante el lector, con ayuda de fragmentos de entrevista, o para profundizar en un tema que no ha sido explorado suficientemente en el cuestionario. Por ejemplo, si de una encuesta sobre el abstencionismo electoral resulta que es una actitud especialmente frecuente entre los jóvenes residentes en áreas urbanas con un nivel de estudios alto, se puede volver al terreno para entrevistar de nuevo a los sujetos que según el análisis de los datos presenten las características citadas, para entender mejor las motivaciones de su comportamiento con una entrevista no estructurada. O bien, si a partir del análisis de los cuestionarios se constata que existe un tema importante que se había descuidado, se puede decidir profundizar en el mismo con una serie de entrevistas *ad hoc*.

En todos estos casos la entrevista cualitativa cumple una *función auxiliar* en la investigación cuantitativa. La base empírica principal está constituida por los cuestionarios, y la fase cualitativa sólo sirve para preparar la investigación cuantitativa, para apoyar sus conclusiones, o para aclarar algunos aspectos que los datos cuantitativos han dejado en la sombra.

Para los investigadores que se inspiran en el paradigma interpretativo, la entrevista cualitativa es, de todas las técnicas basadas en el planteamiento de preguntas, la única vía para conseguir una verdadera comprensión de la realidad social. Como ya sabemos, el paradigma interpretativo desecha el mito de la objetividad de los hechos sociales, defendiendo una visión «constructivista» de la investigación, en la cual los datos no se «recopilan», sino que son «generados» por la interacción entre el sujeto que estudia y el sujeto estudiado. Por lo tanto, la entrevista, para poder generar conocimiento efectivo, debe ser el resultado de esa interacción social particular y única que se desarrolla entre el entrevistador y el entrevistado. Más que un *acto de observación*, constituye un *acto de interacción*, mediante el cual el investigador accede directamente, de forma no muy distinta a la del observador participante, al mundo del entrevistado. Y esta participación es lo que le permite reconstruir su realidad, logrando esa «visión desde dentro» que constituye la única forma verdadera de conocimiento social para el investigador inspirado en el paradigma interpretativo.

Como afirman Schwartz y Jacobs, en el paradigma interpretativo «la única realidad social "verdadera" es la que se "percibe desde el interior" [...]. Para comprender los fenómenos sociales, el investigador debe descubrir "la definición de la situación" que da el actor, es decir, su percepción y su interpretación de la realidad, y la relación de ambas con su comportamiento [...]. En esta perspectiva, los significados sociales (que rigen el comportamiento) no se limitan a las actividades, instituciones u objetos sociales como tales» [Schwartz y Jacobs, 1979, 7-8].

10.7. SÍNTESIS DEL CAPÍTULO 10

1. Podemos definir la entrevista cualitativa como una conversación provocada por el entrevistador, dirigida a sujetos elegidos sobre la base de un plan de investigación y en un número considerable, que tiene finalidades de tipo cognitivo y es guiada por el entrevistador según un esquema flexible y no estandarizado de planteamiento de preguntas.
2. El planteamiento de preguntas a los seres humanos para conocer la sociedad puede realizarse mediante un planteamiento cualitativo (con entrevistas cualitativas) y mediante un planteamiento cuantitativo (con cuestionarios estandarizados). La primera opción se diferencia de la segunda en los siguientes aspectos: ausencia de estandarización, comprensión del fenómeno más que documentación del mismo (contexto del descubrimiento frente a contexto de la justificación), ausencia de una muestra representativa, y punto de vista centrado en el sujeto y no en las variables.
3. Las entrevistas cualitativas pueden ser de tres tipos. En las *entrevistas estructuradas,* a todos los entrevistados se les plantean las mismas preguntas con la misma formulación (las preguntas están definidas de antemano tanto en el contenido como en la forma, como en un cuestionario de preguntas abiertas). En las entrevistas *semiestructuradas* el entrevistador no plantea preguntas formuladas de antemano, sino que utiliza un «guión» que contiene los temas que deberá tratar en el curso de la entrevista (por tanto, las preguntas están predeterminadas en cuanto al contenido, pero no en cuanto a la forma). En las entrevistas *no estructuradas* el entrevistador sólo tiene que abordar determinados temas en el curso de la conversación de la forma que considere más adecuada (es decir, no están prefijadas ni la forma ni el contenido de las preguntas). Casos particulares de entrevistas cualitativas son: las *entrevistas no dirigidas y clínicas*, las *entrevistas a testigos privilegiados* y las *entrevistas de grupo.*
4. La entrevista cualitativa no consiste únicamente en el acto de recopilación de información, sino que conlleva un proceso de interacción entre dos individuos. Realizar una entrevista cualitativa es una tarea difícil. La dinámica de la entrevista puede describirse considerando los siguientes aspectos: las explicaciones preliminares, las preguntas primarias y secundarias, las preguntas exploratorias, el lenguaje y el papel del entrevistador.
5. Una vez que se han realizado las entrevistas, éstas se deben analizar. El *análisis de los datos* se centra en los sujetos (a diferencia del criterio cuantitativo, en el que el análisis se centra en las variables), lo que quiere decir que se observa y se estudia el individuo en su totalidad. Los *resultados* se presentan en forma de narración. Las síntesis y las generalizaciones suelen adoptar la forma de clasificaciones y tipologías. En el caso de las entrevistas estructuradas, los datos se pueden analizar utilizando un criterio mixto, que combina técnicas cualitativas y cuantitativas (las técnicas cuantitativas conllevan la construcción de la matriz de datos).

10.8. LECTURAS COMPLEMENTARIAS

Una reseña muy completa sobre los aspectos conceptuales, metodológicos y prácticos de la entrevista, que incluye referencias a la historia de su utilización, sus diversas formas, los campos de aplicación, los aspectos técnicos y los procedimientos de análisis del material empírico, es la recopilación de ensayos de J. F. Gubrium y J. A. Holstein (eds.) *Handbook of Interview research: Context and Method* (Sage, 2002, 981 pp.).

En cuanto a las monografías dedicadas a la entrevista y sus formas, citamos, para el enfoque tradicional, el volumen clásico de R. L. Kahn y C. F. Cannell *Dynamics of Interviewing: Theory, Technique and Cases* (Krieger, 1983, 368); una crítica a la tradición basada en el estímulo-respuesta, según la cual la entrevista es un producto del entrevistador y el entrevistado, se encuentra en E. G. Mishler, *Research Interviewing: Context and Narrative,* Harvard U. P., 189 pp.

Sobre formas particulares de entrevistas no estructuradas, remitimos a J. T. Chirban, *Interviewing in Depth: The Interactive-Relational Approach* (Sage, 1996, 144 pp.), y a G. McCracken, *The Long Interview* (Sage, 1988, 88 pp.).

En cuanto a las entrevistas de grupo, una primera y breve reseña introductoria puede encontrarse en D. L. Morgan *«Focus Groups»,* en *Annual Review of Sociology, 1996.* Para una descripción completa más detallada y aplicada, remitimos a D. W. Stewart y P. N. Shamdasani, *Focus Groups: Theory and Practice* (Sage, 1990, 153 pp.).

Sobre la técnica de los grupos de discusión, véase el libro monográfico de J. Callejo *El grupo de discusión: introducción a una práctica de investigación* (Barcelona: Ariel Practicum, 2001). El apéndice del libro de García Ferrando *et al. El análisis de la realidad social: métodos y técnicas de investigación* (Madrid: Alianza Editorial, 1986), de Ibáñez, resultará útil para el lector poco iniciado en esta técnica. En ese mismo libro, el capítulo de A. Ortí resulta también muy útil para entender los supuestos ontológicos y metodológicos en los que se basa la técnica.

CAPÍTULO **11**

EL USO DE LOS DOCUMENTOS

- DEFINICIÓN DEL OBJETO 376

- DOCUMENTOS PERSONALES 377
  - 11.2.1. Autobiografías 378
  - 11.2.2. Diarios 379
  - 11.2.3. Cartas 380
  - 11.2.4. Testimonios orales 382

- DOCUMENTOS INSTITUCIONALES 388
  - 11.3.1. Medios de comunicación 390
  - 11.3.2. Narrativa, textos pedagógicos, cuentos populares 392
  - 11.3.3. Material judicial 394
  - 11.3.4. Documentos de la política 395
  - 11.3.5. Documentos empresariales y administrativos 397
  - 11.3.6. Huellas físicas 399

- LEER LOS DOCUMENTOS HUMANOS 401
  - 11.4.1. Leer los documentos personales 401
  - 11.4.2. Leer los documentos institucionales 402

- SÍNTESIS DEL CAPÍTULO 11 403

- LECTURAS COMPLEMENTARIAS 404

Toda sociedad produce una gran cantidad de documentos, y en este capítulo veremos cómo pueden utilizarse estos documentos como bases de datos para la investigación social. El capítulo tiene dos partes. En la primera parte trataremos los documentos personales, es decir, documentos generados por los individuos para su uso privado (diarios, cartas, autobiografías), así como testimonios orales y autobiografías. En la segunda parte consideraremos los documentos institucionales, es decir, documentos generados por las instituciones y que se pueden considerar «huellas» de su forma de organización y la cultura que las ha generado.

11.1. DEFINICIÓN DEL OBJETO

Al introducir esta tercera parte del libro hemos dicho que *observar, preguntar* y *leer* son las tres acciones fundamentales en las técnicas de investigación cualitativa. En los dos capítulos anteriores hemos tratado las dos primeras. Ahora hablaremos de la tercera, leer, y en particular leer los documentos que el hombre y las instituciones han generado y siguen generando. Veamos en qué consisten estos documentos.

Un *documento* es un material informativo sobre un determinado fenómeno social que existe con independencia de la acción del investigador. Por tanto, el documento es generado por los individuos o las instituciones para fines distintos de los de la investigación social, aunque ésta puede utilizarlo para sus propios fines cognitivos. Estos documentos pueden ser de diversos tipos: cartas, artículos de periódico, diarios, autobiografías, organigramas de empresas, actas de consejos de administración, actas parlamentarias, sentencias judiciales, balances de empresa, certificados de matrimonio, reglamentos, contratos comerciales, etc. Normalmente se trata de documentos en forma escrita, pero también entran en la categoría de los documentos sociales las «huellas materiales». Así mismo, incluiremos en esta categoría los testimonios y los recuerdos de los individuos sobre su pasado.

El hecho de que estos documentos se generen independientemente de la acción del investigador [1] comporta dos ventajas claras con respecto a las técnicas de investigación presentadas anteriormente. En primer lugar se trata de información «no reactiva», es decir, que no se ve afectada por la interacción estudioso-estudiado y sus posibles efectos de distorsión. Mientras que en una entrevista el entrevistado puede no ser sincero en sus respuestas o dejarse guiar por las convenciones sociales para «quedar bien» ante el entrevistador, la misma persona será probablemente más sincera y directa en una carta a un familiar o en las páginas de su diario. De este modo, el balance de una empresa, un historial clínico, una declaración de la renta, el acta de una reunión, etc., son documentos que, aunque ofrecen una visión particular de la realidad social, están libres de alteraciones debidas al acto de la investigación.

La parte negativa es que, dado que el documento ya existe previamente, el investigador no puede preguntarle y tiene que conformarse con su contenido, aun-

[1] Con excepción de los testimonios orales, pero volveremos sobre esta cuestión más adelante.

que éste resulte incompleto para los fines cognitivos que persigue.

La segunda ventaja derivada de la existencia de los documentos al margen de la acción del investigador es que estos documentos permiten estudiar también el pasado. En este caso el trabajo del sociólogo se confunde con el del historiador y, de hecho, las técnicas de investigación de ambos se solapan (desde este punto de vista es difícil distinguir entre sociología e historia social). Para separar ambos campos, aquí nos limitaremos al estudio de la sociedad contemporánea.

En este capítulo hablaremos de dos tipos de documentos totalmente distintos entre sí. El primer tipo son los *documentos personales*. Se trata de documentos de carácter *privado,* que han sido elaborados por los individuos en primera persona y para un uso estrictamente «personal»; también se llaman documentos «expresivos», ya que contienen expresiones de sentimientos, acontecimientos y, más en general, de la personalidad de quien los ha escrito. Dentro de esta categoría se incluyen las cartas, las autobiografías y los diarios.

El segundo tipo lo forman los que llamaremos *documentos institucionales.* Se trata de documentos producidos por instituciones o por individuos, en el contexto de la parte «institucionalizada» de su vida. Pueden ser, por ejemplo, discursos, documentos empresariales, artículos de periódicos, actas de juicios, etc., en general, documentos de carácter *público,* por contraste con el carácter privado de los anteriores.

11.2. DOCUMENTOS PERSONALES

Los diarios, cartas, memorias personales, etc., han tenido siempre un papel destacado entre los documentos históricos, pero la investigación histórica ha utilizado tradicionalmente este material cuando se trataba de testimonios de personalidades relevantes, de protagonistas de la historia. Las ciencias sociales se interesan por los documentos de este tipo producidos por «individuos normales», a fin de reconstruir las dinámicas y las relaciones sociales a partir de las vivencias de los protagonistas en sus vidas cotidianas.

Según Plummer, «el mundo está repleto de documentos personales. Las personas tienen diarios, mandan cartas, hacen fotografías, toman apuntes, cuentan sus recuerdos, escriben en las paredes, publican sus memorias, escriben cartas a los periódicos [...]» [1983, 13]. Todos estos documentos tienen en común la característica de ser expresiones genuinas de la personalidad de sus autores. El aspecto peculiar de los documentos personales es que, al proceder de una experiencia interior del individuo y no estar destinados a uso público, presentan esa característica de «visiones desde dentro» que permite incluirlos por pleno derecho dentro del enfoque interpretativo.

La historia de este tipo de investigación sociológica discurre paralela a los avatares de la investigación cualitativa. En los años veinte y treinta, a partir de la primera propuesta de utilización de documentos personales en la notable investigación de Thomas y Znaniecki, *El campesino polaco en Europa y América,* publicada entre 1918 y 1920, este tipo de material se uso mucho, sobre todo entre los investigadores de la Escuela de Chicago. La popularidad de estos instrumentos decayó

pronto, como le sucedió en general a toda la corriente de investigación cualitativa. Sin embargo, a partir de los años ochenta, el punto de vista biográfico ha vuelto a tomar fuerza en la investigación social, fomentado por el desarrollo de la denominada «historia oral» en el campo de la investigación histórica, que utiliza técnicas de investigación y material empírico muy parecidos.

11.2.1. Autobiografías

La autobiografía es el relato *escrito* de *toda la vida* de una persona, *elaborado por ella misma,* en un periodo de tiempo más bien limitado y, por tanto, con una *visión retrospectiva* puesto que si la escritura se produce durante el curso de los acontecimientos, estaríamos ante un *diario*. En su forma ideal, se trata de obras redactadas *espontáneamente* por personas *no seleccionadas*.

Las biografías de este tipo son poco frecuentes en las ciencias sociales, ya que encontrarse con alguna de ellas en el proceso de búsqueda de documentos relativos al fenómeno estudiado depende totalmente de la suerte. Por ello en investigación social es frecuente provocar la autobiografía, es decir, solicitar expresamente a sujetos que se consideran interesantes para los fines de la investigación que escriban su autobiografía. Por ejemplo, Thomas y Znaniecki, en la investigación citada anteriormente, pidieron a un inmigrante polaco llamado Wladek, que escribiera su autobiografía, a cambio de una compensación económica, aunque, como escriben los autores, «probablemente la ambición, el interés literario y el interés por su propia vida se convirtieron pronto en los motivos principales» [Thomas y Znaniecki, 1918-20; vol. 3, 86].

En estos casos, cuando el investigador encarga la autobiografía, éste puede guiar el relato planteando preguntas o proporcionando una lista de los temas que quiere que se traten en la misma. Un clásico del uso del relato autobiográfico en investigación social es la obra *The Professional Thief,* de Edwin Sutherland, de 1937, que consiste en la autobiografía de un ladrón profesional: fue escrita por el delincuente en sus dos terceras partes, siguiendo una lista de temas elaborada por el investigador (Sutherland), y luego se completó con conversaciones entre el sujeto estudiado y el investigador, que transcribía el contenido de las mismas inmediatamente después de que se produjeran. La redacción final fue la síntesis de estos dos materiales. El libro consta de 200 páginas de relato autobiográfico, seguidas de 30 páginas que contienen la interpretación del investigador. En el caso del campesino polaco, el texto final también estaba constituido por unas 250 páginas de autobiografía y una introducción de 70 páginas escrita por los autores, aunque en este caso la autobiografía fue uno de los muchos tipos de documentos personales que utilizaron los autores, en su amplio trabajo de investigación.

No obstante, la autobiografía más clásica es la que escribe el propio autor y se publica como testimonio personal, sin comentarios ni interpretaciones. Por ejemplo, *Pimp: the Story of my Life* (Iceberg, 1967). Tras haber pasado un largo periodo de tiempo en prisión condenado por proxenetismo, el autor escribió la historia de su vida, en la que describió el mundo de la delincuencia en Chicago durante los años cuarenta y cincuenta, con la prostitución, el tráfico de drogas, la corrupción y la

violencia; esta autobiografía ofrece una descripción desde dentro, lúcida y con una gran cantidad de información.

Otro ejemplo es la obra titulada *La isla de la cuarentena*. En 1891, con 16 años de edad, su autor dejó su pueblo en los Alpes italianos para ir a buscar trabajo, recorriendo Austria, Hungría, Bohemia, Silesia, Galicia y Sajonia; con 30 años cruzó el Atlántico, trabajó en las minas canadienses para llegar a Nueva York en 1922, y fue precisamente durante la parada obligatoria de la cuarentena a la que se sometía a los inmigrantes antes de poder entrar en Estados Unidos, cuando escribió su autobiografía [Del Piero, 1994]. Ésta constituye un retrato de notable interés sociológico sobre las grandes migraciones que afectaron a las zonas más pobres de Europa entre los siglos XIX y XX.

Estas narraciones autobiográficas «puras», es decir, escritas de forma directa sin la mediación ni la interpretación del investigador, no pueden encuadrarse fácilmente dentro de la investigación social y se consideran más bien un género literario. Conviene decir que quien escribe una autobiografía no es una persona corriente: el propio hecho de decidir escribir la propia historia o aceptar la invitación de relatarla ya denota una personalidad muy particular. También conviene señalar que la autobiografía a menudo comprende sólo lo que el autor considera interesante y con valor narrativo, y excluye todo lo que a sus ojos podría suponer un desprestigio para él. Por último, la autobiografía, como todos los documentos basados en la memoria, presenta el riesgo de que su autor racionalice a posteriori los hechos del pasado. El condicionamiento de los roles sociales habituales o la exigencia de considerar experiencias pasadas desde el punto de vista actual, podrían dar una imagen falseada de las actitudes reales con las que se afrontaron ciertas situaciones en el pasado.

11.2.2. Diarios

Los diarios se consideran «el documento personal por excelencia». Su peculiaridad reside en el hecho de que los diarios se escriben para uso estrictamente personal e inmediatamente después de que se produzcan los hechos descritos. Esto hace que este tipo de documentos estén exentos de dos de los defectos principales de las autobiografías. Sobre todo el hecho de que los actos, opiniones, formas de pensar y sentir se registren en el mismo momento en que se están viviendo, hace de este tipo de documentos un testimonio muy valioso y único de la vida interior de la persona que escribe, así como una descripción precisa del desarrollo de los hechos en los que participa. Pese a ello, hay que decir que, mientras que en la literatura los diarios están muy presentes, no existe ninguna investigación social de importancia que se haya basado exclusivamente en diarios, probablemente por un motivo similar al que hemos dado para explicar la escasa presencia de las autobiografías: el sociólogo sólo dispone de este tipo de documentos en casos excepcionales, y con frecuencia se trata de material particular, difícilmente generalizable.

En la tradición de la investigación sociológica los diarios se utilizaron sobre todo en los años veinte y treinta, cuando predominaba el modelo de investigación social de la Escuela de Chicago. Pero también en estos casos su uso fue siempre limitado y tuvo una función complementaria o auxiliar. En su manual sobre meto-

dología de la «investigación de campo» Palmer [1928] cita dos investigaciones que utilizaron diarios: una sobre el suicidio de Cavan [1928], que analiza los diarios de dos suicidas, y una sobre la desorganización de la familia de Mower [1927], que reproduce el diario de una mujer joven cuyo matrimonio desembocó en una crisis. También el matrimonio Lynd, en su investigación sobre Middletown, afirma haber utilizado dos diarios.

Un caso particular de diario que se usa en investigación social son los relatos diarios elaborados a petición del investigador. Ya hemos mencionado el uso de los «balances de tiempo»: se trata de fichas que se entregan a los sujetos estudiados en las que la jornada se divide en fracciones de tiempo, y para cada una de ellas se pide a dichos sujetos que anoten lo que han hecho durante un determinado periodo (por ejemplo, un día o una semana). La técnica, típicamente cuantitativa en el caso descrito, también se puede usar en investigación cualitativa. Un ejemplo de ello es la investigación de Maas y Kuypers [1974], que realizaron un estudio longitudinal sobre los mecanismos de adaptación a la vejez de una muestra de 142 sujetos de clase social alta, para el cual pidieron a un cierto número de ellos que escribieran un diario durante una semana, invitándolos a ir más allá de la simple cronología de sus actividades, anotando también pensamientos, sentimientos y comentarios.

Zimmerman y Wieder [1977] hicieron un uso similar de los diarios: ofrecieron una pequeña compensación económica a los sujetos para que escribieran un diario durante una semana y les pidieron que describieran todas las actividades respondiendo a las preguntas ¿quién?, ¿qué?, ¿cuándo?, ¿dónde?, ¿cómo? La originalidad de la técnica radica en el hecho de que posteriormente el diario se utilizaba también como base para una entrevista en profundidad sobre los hechos expuestos en el mismo.

Para finalizar debemos decir que, obviamente, estas formas de diarios «encargados» por el investigador presentan características muy distintas a las de los diarios «espontáneos» (entre otras cosas, no son documentos en sentido estricto, si limitamos el término al material que existe al margen de la investigación). Sin embargo, son una forma interesante y original de obtener material empírico.

11.2.3. Cartas

Una de las primeras y más conocidas investigaciones sociológicas, *El Campesino polaco en Europa y América,* de William Thomas y Florian Znaniecki [1918-20], que ya hemos mencionado anteriormente, se basa, en gran parte, en cartas. Se trata de correspondencia cruzada entre emigrantes polacos en Estados Unidos y sus familiares en Polonia, que los autores consiguieron poniendo un anuncio en el periódico de la comunidad polaca residente en Chicago, en el que ofrecían una compensación económica por cada carta. De esta forma, lograron reunir una gran cantidad de cartas, que utilizaron como material empírico para su investigación (junto con artículos de periódicos, la autobiografía que hemos citado antes y otros documentos). Una parte del contenido de estas cartas se publicó, ocupando más de 600 páginas, en los volúmenes que recogían la investigación.

Las cartas se emplearon para describir las formas tradicionales de vida en el campo polaco, las primeras transformaciones sociales tras el proceso de industrialización y de concentración urbana de fin de siglo, la desorganización de la sociedad campesina tradicional y su reorganización en nuevas estructuras, y el proceso de integración de los inmigrantes polacos en Estados Unidos. Con el uso de las cartas los autores intentaron sobre todo centrar la atención en los aspectos subjetivos de los procesos sociales, y en particular en los mecanismos de adaptación de los individuos en una sociedad en vías de transformación.

A continuación citaremos muy brevemente otras investigaciones que utilizaron material epistolar. Strauss realizó un estudio bastante singular a partir de las cartas que el propio autor recibió durante 25 años de un amigo alcohólico. En este estudio se relatan las vicisitudes de las hospitalizaciones, la relación irregular con el trabajo, los tratamientos de desintoxicación y en general las experiencias subjetivas en el contexto de esta dramática forma de vida. Como ejemplos de investigaciones sobre múltiples sujetos con una experiencia común, podemos citar un estudio de Salzinger [1958] a partir de las cartas escritas por niños con trastornos emocionales a sus amigos y parientes, y una investigación de Janowitz [1958] basada en cartas y diarios requisados a los soldados alemanes durante la guerra.

Las cartas enviadas a periódicos o semanarios y las cartas enviadas a personajes de la vida pública constituyen un caso especial de estos documentos. Son documentos muy distintos de los anteriores, puesto que el componente de espontaneidad y manifestación de la propia subjetividad puede quedar ensombrecido por el hecho de tratarse de un documento destinado al público (o a un personaje público). Sin embargo, dado que puede contener documentación homogénea extensa, este material puede resultar muy interesante para una investigación. Entre los estudios basados en material epistolar enviado a cargos políticos cabe citar la investigación de Sussman [1963] sobre las cartas recibidas por el presidente estadounidense Franklin Delano Roosevelt. En cuanto a las cartas enviadas a los periódicos, cabe citar un caso interesante, la investigación de Boltanski [1990], que estudió la «denuncia pública», es decir, el acto del ciudadano que reacciona ante una injusticia de la que ha sido testigo denunciándola públicamente. El autor basó su estudio en el análisis de 275 cartas recibidas por el diario *Le Monde* en los años 1979, 1980 y 1981, de las que sólo se publicaría una pequeña parte en el periódico.

Las investigaciones basadas exclusivamente o casi exclusivamente en material epistolar son poco frecuentes en investigación social, y actualmente tienden a desaparecer. Esto se debe a que, con la difusión del teléfono, la comunicación por carta se ha hecho cada vez menos habitual en las sociedades modernas. Ésta se convierte en un instrumento de comunicación importante sólo en situaciones de separación forzosa, como en el caso de guerras o de emigraciones.

Las cartas son un instrumento importante para expresar las vivencias internas de los individuos, su definición subjetiva de las situaciones, y desde este punto de vista ofrecen un material de gran pureza. No obstante, hay que puntualizar que la carta «no expresa sólo el mundo de la persona que escribe, sino también la percepción de la persona que escribe que tiene el destinatario» [Plummer, 1983, 23]; en este sentido la carta no siempre es una manifestación inalterada de la personalidad del

sujeto que escribe (como sucede con los diarios), sino que se debe interpretar como el producto de la interacción entre dos personas.

11.2.4. Testimonios orales

La utilización de los documentos personales expuestos hasta ahora para los fines de la investigación presenta varios problemas. En primer lugar, *a*) se trata de *documentos infrecuentes,* producidos por situaciones o por personas excepcionales; por otra parte, resulta difícil obtenerlos, puesto que los propietarios de cartas, diarios, u otros documentos personales pueden oponerse a su publicación. En segundo lugar, al no ser producidos para los objetivos de la investigación, los documentos pueden ser *b*) *fragmentarios e incompletos,* privando al investigador de datos fundamentales para la comprensión de la totalidad de la realidad estudiada. Por último, estos documentos presentan el grave problema de *c*) la *escasa representatividad* de los sujetos estudiados. El investigador no puede escoger los casos a estudiar, sino que debe aceptar los que la situación le ofrece. Además, serán casos esporádicos, un conjunto sin orden ni regularidad. Otra limitación es que los sujetos que producen los documentos casi nunca son representativos de la población general relevante para el problema estudiado. En primer lugar, por una cuestión de alfabetización, dado que se trata de documentos escritos. Por ejemplo, en el caso de los campesinos polacos, quienes sabían escribir cartas formaban un subconjunto con características muy específicas dentro de la población campesina de esos tiempos, que en su mayoría era analfabeta. En segundo lugar, por una cuestión de estructura de la personalidad. Como ya hemos apuntado, quienes tienen un diario o escriben su propia autobiografía de manera espontánea son una minoría y se trata de personas con ciertas peculiaridades que las hace distintas del resto.

Por todo ello, se ha desarrollado un nuevo enfoque para la recopilación de documentos personales en la investigación social que consiste en recoger de manera sistemática testimonios orales de los individuos sobre su propia vida. Se trata de *provocar* lo que podríamos llamar «autobiografías orales»: *relatos de vida* que no deberían diferenciarse en nada de las autobiografías, excepto en el hecho de que son orales y no escritas.

Este enfoque se usa en varias disciplinas sociales, como la antropología, donde el estudio de sociedades iletradas ha exigido desde el principio recurrir a las fuentes orales, la psicología, que nos ha proporcionado célebres estudios de caso, en los que el componente biográfico ha tenido una función explicativa fundamental en la definición de psicopatologías, la historia, en la que recientemente se ha desarrollado un movimiento de «historia oral» que ha aportado notables contribuciones en diversos sectores de historia contemporánea (historia del movimiento obrero, historia de la condición de la mujer, historia urbana, historia de la industrialización, etc.).

Estas narraciones del pasado ante un entrevistador se parecen mucho a la técnica de la entrevista no estructurada que hemos visto en el capítulo anterior, aunque existen al menos tres diferencias importantes entre ambas.

En primer lugar, se trata de contar la propia vida (generalmente, toda la vida transcurrida hasta el momento de la narración), por tanto, estamos en la dimensión

autobiográfica, algo que no se da en las entrevistas, o al menos no está presente necesariamente.

En segundo lugar, en las historias de vida el papel del entrevistador es mucho más limitado. Su función está limitada al registro de los datos, como ocurre en una biografía dictada a un magnetófono o en una carta dictada por el campesino analfabeto al escribano. Con frecuencia la entrevista a una persona sobre su vida es un relato autobiográfico que se produce de manera autónoma casi sin preguntas, y si las hay, no son preguntas propiamente dichas, sino más bien estímulos para estructurar mejor el recuerdo o ayudar al narrador a enlazar la secuencia de los hechos. Por lo tanto, las conversaciones, no siguen las líneas de una entrevista, sino las de una *narración* guiada por el hilo de la memoria, una especie de monólogo del sujeto que habla y relata su propia vida.

Por último, el método biográfico no se preocupa particularmente por la cantidad de casos estudiados. Mientras que las entrevistas cualitativas se suelen realizar sobre varias decenas de casos, como hemos visto en el capítulo anterior, las investigaciones que usan el método biográfico se suelen conformar con unas cuantas unidades, y este criterio no excluye la posibilidad de usar incluso un solo caso (sería el «estudio del caso», del que veremos algunos ejemplos), partiendo del supuesto de que la comprensión de los mecanismos individuales, incluso de una sola persona, puede arrojar luz sobre el fenómeno estudiado, igual que el análisis de un único caso clínico en la investigación psicológica.

Obviamente, esta técnica no presenta una de las características específicas del «documento» tal como lo hemos definido aquí: el hecho de que el documento existe al margen de la investigación y se produce con un objetivo distinto de los fines cognitivos de la propia investigación. Los relatos orales, en cambio, son solicitados por el investigador, y ello nos obliga a revisar nuestra definición, aunque esta revisión no afecta a su sentido general. Muchas veces, aunque no siempre, el uso de los testimonios orales consiste en la grabación de los recuerdos de las generaciones más ancianas, esas generaciones que, como escribe Evans en su investigación sobre los pueblos de East Anglia, son «por así decirlo, historia viviente, y sus miembros [son] "libros ambulantes"» [Evans 1975; 16].

Por tanto, en cierto sentido, entrevistar a las personas sobre su vida sigue siendo «leer documentos», sólo que en este caso se trata de un tipo particular de documentos, una especie de «documentos vivientes». Siguiendo con las palabras de Evans, «estas personas ancianas habían ejercido su oficio, conservado las tradiciones y usado las herramientas exactamente como lo habían hecho sus antepasados durante siglos, y podían describir sus actividades de forma precisa y con todo detalle, en un lenguaje que poseía en sí mismo innumerables ecos del pasado» [*ibídem*].

El hecho de que el investigador solicite los relatos orales conlleva el riesgo de la «reactividad»: la presencia del entrevistador puede provocar una reacción por parte del entrevistado que altere la naturalidad de su relato. Por ello es muy importante que el entrevistador limite al máximo las interrupciones en la narración, para mantener intacta la connotación *naturalista*, es decir, la espontaneidad del relato.

Otro de los inconvenientes de esta técnica se deriva de la memoria y sus posibles fallos. Pero, como escribe Paul Thompson, uno de los fundadores de la historia oral

moderna, «parece que los fallos de memoria afectan sobre todo a la memoria reciente [...]. [Por otra parte, en la edad anciana] se encuentra lo que los psicólogos llaman el fenómeno del «reexamen de la vida»: de repente surgen recuerdos acompañados del deseo de comunicarlos y una franqueza característica, determinada por la conciencia de haber llegado al fin de la vida activa», por lo que, en palabras de Thompson, «entrevistar a los ancianos no añade problemas metodológicos distintos de los problemas normales de la entrevista» [1972, 13-14].

La utilización de las fuentes orales ofrece numerosas ventajas (si las comparamos con los documentos escritos). Al poder elegir a las personas a entrevistar, el investigador puede: *a*) evitar la irregularidad de los casos estudiados, seleccionando una muestra que presente una cierta uniformidad, cuando no directamente representatividad. Al plantear las preguntas puede: *b*) evitar los datos fragmentarios e incompletos, animando al narrador para que cubra eventuales lagunas en su narración.

Por último, al poder preguntar a las personas y servirse de relatos orales en lugar de material escrito, que por definición está circunscrito a los individuos letrados, el investigador puede: *c*) evitar que su base informativa se limite a los estratos sociales más altos, logrando llegar así a una reconstrucción de la historia «desde abajo», centrada en las condiciones de vida de las clases inferiores, la cultura material y las expresiones religiosas populares.

Esta aproximación a la obtención de material documental a través de testimonios orales presenta diversas variantes. Aquí distinguiremos las que en nuestra opinión son las dos formas principales de esta técnica: los historiales de vida y la historia oral.

El «historial de vida» es el relato de la vida de una persona realizado por ella misma a un entrevistador mediante conversaciones o entrevistas. La «historia oral» es el relato de una persona sobre acontecimientos en los que ha participado. El historial de vida tiene por objeto *un individuo,* su biografía, sus emociones, su forma de pensar, la sucesión de sus vicisitudes personales. La historia oral tiene por objeto *la sociedad,* las costumbres, los acontecimientos sociales. En el primer caso, el narrador es el centro de la narración, en el segundo, lo son el mundo y las vicisitudes que ha vivido. Podemos añadir que los historiales de vida se encuadran en su mayoría en la tradición sociológica de la investigación social, mientras que la historia oral se sitúan dentro de la tradición histórica.

• *Historiales de vida.* Los historiales de vida son, como hemos dicho, relatos autobiográficos centrados en las vivencias personales de la narrador. Por lo general, las investigaciones sociológicas se basan en un número limitado de estas «autobiografías narradas». Para hacernos una idea, lo habitual es que las investigaciones se basen en una decena de relatos. Pero también abundan las investigaciones basadas en una única biografía. No obstante, en nuestra opinión, el criterio más importante para clasificar las investigaciones realizadas con esta técnica no es el número de casos estudiados, sino más bien el modo en que el investigador trata el material biográfico. Tras analizar las investigaciones disponibles con este criterio, podemos distinguir tres maneras de tratar el material estudiado.

La primera forma de tratar este material es lo que podríamos denominar un uso *precientífico* de los documentos, que consiste en el registro puro y simple de autobiografías: los relatos son recopilados y publicados en un volumen, con algunas páginas de presentación escritas por el investigador (cuyo rol en este caso se asemeja más al de periodista que al de investigador social). Sin negar la fascinación que pueden despertar estos fragmentos de vida, que suelen ser sugerentes y tener un cierto valor literario, los historiales de vida, como cualquier otro documento social, no hablan por sí solos: es necesario preguntarles.

Esta forma de utilizar el material biográfico podría estar justificada en alguna ocasión por la motivación «naturalista» de respetar las voces narradoras: reproducir los relatos autobiográficos en su forma original, sin selecciones, comentarios ni interpretaciones, en otras palabras «sin manipulación», significa dar voz a quien normalmente no la tiene, un modo de dejar que los oprimidos, los marginados y los excluidos se hagan oír, y permitir que el punto de vista de las clases inferiores salga a la luz, sin la mediación del investigador, que inevitablemente lo alteraría, por tener una cultura distinta.

Pero la vivacidad de estas narraciones va acompañada de su carácter fragmentario. El hecho de reunir varios relatos de vida sin una síntesis común, o al menos un intento de encuadrar sus temas fundamentales en un esquema explicativo, impide que estas recopilaciones puedan convertirse en verdaderas investigaciones sociales, productos científicos propiamente dichos.

El *segundo modo* de tratar el material autobiográfico, en cambio, sí contiene este esquema explicativo. Consiste en presentar un amplio ensayo interpretativo junto con las autobiografías, siempre completas y con autonomía propia. Éste es el caso de la investigación ya mencionada de Thomas y Znaniecki sobre el campesino polaco. Consideremos la primera parte de esta investigación, que se corresponde con los dos primeros volúmenes de los cinco que forman la edición original. Esta parte contiene un total de 754 cartas [2] que ocupan más de 600 páginas, precedidas por un ensayo de casi 200 páginas en el que se hace un análisis de la sociedad campesina polaca (la familia, el sistema de clases, la economía, las creencias religiosas). Todos los analistas han valorado siempre esta interpretación, como todas las partes teóricas de la investigación de Thomas y Znaniecki, por su análisis inteligente y profundo de gran alcance teórico, que realmente ilustra los efectos de las transformaciones sociales en las vidas individuales. Pero al mismo tiempo también se ha destacado su distanciamiento de la parte empírica de la investigación, de la que debería haberse alimentado.

La crítica más conocida en este sentido es la que expresó Herbert Blumer en 1938, con motivo del debate que propuso el Social Science Research Council estadounidense sobre la investigación de Thomas y Znaniecki. A continuación reproducimos textualmente las palabras de Blumer:

---

[2] El hecho de que, en este caso, el material empírico no esté compuesto por historiales de vida sino por cartas es irrelevante para los objetivos de nuestro planteamiento (las consideraciones sobre el tratamiento de material biográfico se pueden aplicar a todos los documentos personales).

> El problema al que debemos enfrentarnos es la relación entre el material empírico y el análisis teórico [...]. Se presume que los análisis teóricos partieron de este material o se comprobaron empíricamente con el mismo [...]. Resulta evidente que no todos los avances teóricos de Thomas y Znaniecki derivan del material empírico contenido en sus obras [...]; quizás ni siquiera sus concepciones teóricas más importantes derivan de dicho material.
>
> [Blumer, 1939, 73-74].

Blumer prosigue afirmando que probablemente los autores iniciaron la investigación teniendo ya claros unos esquemas interpretativos, derivados de su familiaridad con el mundo campesino polaco. El modo en que lo hicieron no es relevante, lo que importa es el hecho de que «las consideraciones teóricas del *El campesino polaco en Europa y América* van mucho más allá de su material empírico [...]. Algunas interpretaciones nacieron al margen de los documentos, mientras que otras veces las interpretaciones no están comprobadas adecuadamente» [*ibídem*, 75].

Esta discordancia entre las interpretaciones teóricas y el material empírico que resulta tan evidente en esta investigación, es un riesgo constante en las investigaciones basadas en documentos personales (y, más en general, en la investigación cualitativa); y el riesgo es mayor cuando el investigador decide no realizar el análisis mediante una comparación continua, afirmación tras afirmación, con el material empírico, sino que toma la decisión de tratar las narraciones autobiográficas (u otros documentos personales) como un corpus unitario que debe interpretarse de forma global y autónoma.

El *tercer modo* de analizar los documentos biográficos consiste en contrastar de manera sistemática el material empírico y la interpretación teórica, en una comparación continua que adquiere también una visualización propia por la forma en que se construye el texto. La exposición se guía por los problemas e interrogantes teóricos, éstos constituyen el hilo conductor que da cohesión al material empírico. Las consideraciones socio-antropológicas ya no se formulan al margen de la presentación de los documentos, sino que los comentarios interpretativos y los textos autobiográficos se integran y relacionan entre sí: el comentario sirve de guía y el texto narrativo sirve de ilustración y confirmación. En otras palabras, el testimonio oral no queda aislado, sino que es interpretado, los episodios se relacionan entre sí, se altera su orden cronológico natural y se sitúan en un contexto más amplio; el investigador se impone la tarea de organizar el material, encontrar nexos lógicos y establecer conexiones entre los hechos (por ejemplo, entre una forma determinada de socialización y ciertos comportamientos adultos).

Un ejemplo de investigación que aplica este procedimiento es *The Fence,* del investigador estadounidense Darlel Steffensmeier [1986]. Esta investigación se basa en la experiencia de vida de un individuo llamado Sam Goodman, un perista de sesenta años que el autor conoció en una cárcel americana cuando estaba cumpliendo tres años de prisión por tráfico de mercancía robada. La obra es una descripción muy gráfica e ilustradora del submundo de la ilegalidad, que discurre paralelo a la sociedad convencional. El relato narra la forma en que el perista inició su actividad, el

modo en que estableció los contactos con los ladrones y los clientes, el funcionamiento del mercado de mercancías robadas, la forma de fijar los precios, la realización de las transacciones, la cuantía de los beneficios que reporta el negocio, las habilidades que requiere, entre ellas, las destrezas necesarias para evitar ser descubiertos y detenidos por la policía, y en caso de detención, la forma de gestionar las relaciones con la justicia para limitar al máximo las penas impuestas.

La exposición no sigue una cronología autobiográfica, sino que se divide en distintos temas (la «profesión» del perista, el acceso a la actividad, la forma de comprar, vender y gestionar las relaciones con la policía, etc.). Y cada tema se aborda relacionando el análisis y los documentos disponibles. Por ejemplo, el capítulo titulado «Cómo se compra la mercancía robada» contiene una cuidada exposición sobre las normas tácitas que regulan el precio (el criterio básico de un tercio de su precio en tiendas minoristas, las variaciones en función del riesgo y la demanda del mercado, la competencia con otros peristas, las modalidades de pago, la negociación, etc.). Y cada argumento se trata de la siguiente forma: presentación y discusión de la problemática planteada por el autor, afirmación seguida por el fragmento de entrevista que la ilustra. Por ejemplo, la afirmación «algunos ladrones están más informados que otros sobre el valor de la mercancía que ofrecen, y por tanto son más hábiles en la negociación con el perista» [*ibídem*, 89] va seguida de un fragmento de entrevista a Sam que habla de los ladrones con experiencia, que saben el precio de todos los artículos, y de los jóvenes, que no saben negociar, de la habilidad del perista para reconocer a un «buen» ladrón y tratarlo bien para asegurarse de que contará con sus servicios en el futuro, etc.

• *Historia oral.* Ya hemos dicho que en este caso el recurso al relato de los testimonios está más motivado por la intención de recopilar información sobre la sociedad en la que vive el sujeto que por el objetivo de conocer su vivencia personal. Tradicionalmente se ha considerado que la historia era una disciplina que utilizaba como fuente el material escrito, mientras que la antropología se servía de fuentes orales. Recientemente este planteamiento ha sido criticado, y es cada vez más frecuente utilizar testimonios orales en la base documental de la reconstrucción histórica.

La que más tarde se denominaría «historia oral» tiene su origen a finales de los años cuarenta en la Columbia University y se desarrolla rápidamente dando lugar a un movimiento que en los años sesenta contaba con varios centenares de miembros en Estados Unidos, comprometidos en proyectos de grabación en directo de testimonios oculares de hechos y acontecimientos de relevancia histórica (incluso con el único fin documental de reunir testimonios para poder disponer de ellos en el futuro). Posteriormente el movimiento adquirió un carácter más académico, y la escuela inglesa se fue haciendo cada vez más importante (en 1972 se
Unido la revista *Oral History*).

En general, puede decirse que uno de los objetivos de est
contraponer la historia del pueblo y de la vida material a la histor
de los tratados, las vicisitudes de las clases inferiores y las min
güísticas a la historia de las elites y las clases dominantes, la his
mientos de liberación nacional y racial a la de los imperialismo

«historia desde abajo», a veces cargada de populismo, pero que ha aportado importantes contribuciones a la reconstrucción de las características de las sociedades que nos han precedido y que han sabido valorar ese filón de información que representa la memoria de las generaciones mayores.

Un ejemplo clásico de historia social basada en fuentes orales es *The Edwardians,* de Paul Thompson [1975], que traza un amplio retrato de la sociedad británica de finales del siglo XIX y principios del XX (el periodo del reinado de Eduardo VII, hijo y sucesor de la reina Victoria), describiendo la estructura de sus desigualdades sociales y económicas, las relaciones entre padres e hijos, las diferencias entre el campo y la ciudad, el uso del tiempo libre, el papel de la religión, los movimientos políticos, etc. Todo ello tomando como base documental una muestra bastante sistemática de individuos: se entrevistaron unos 500 individuos de ambos sexos nacidos entre 1872 y 1906, de diversas clases sociales, con el objetivo de cubrir a grandes rasgos las estructuras sociales y geográficas de la sociedad británica de aquellos años. Además, dentro del material se incluyeron los historiales de vida de doce personas de la misma generación.

Hay que precisar que este tipo de estudios raras veces utilizan muestras tan sistemáticas y tan grandes: lo habitual es utilizar varias decenas de testimonios en lugar de varios cientos. En cuanto a la integración en el texto de estos relatos personales, la exposición presenta una parte dominante que contiene el análisis descriptivo e interpretativo del autor, con inclusiones esporádicas de los fragmentos de entrevista a modo de ilustración y apoyo de las afirmaciones.

11.3. DOCUMENTOS INSTITUCIONALES

Hasta ahora hemos hablado de documentos que proceden del ámbito privado de la vida de las personas. Pero, como ya hemos dicho, también en la esfera pública de todas las sociedades se producen una infinidad de documentos, que permanecen como «huellas» de esa cultura y que pueden servir de documentación para el investigador social. Aparte de las personas o las instituciones, las «cosas» también pueden hablar al investigador social, presentándose como «huellas físicas» de la cultura que las ha producido.

Por ejemplo, Michael Foucault reconstruye el nacimiento y el desarrollo de lo que él define como la «sociedad disciplinaria», que se produce entre los siglos XVI y XIX con la puesta en marcha de una serie de procedimientos orientados a «encasillar, controlar, medir, formar a los individuos, hacerlos dóciles y útiles al mismo tiempo», y para describir este proceso utiliza documentos escritos (por ejemplo, reglamentos de las escuelas y las fábricas) y documentos materiales (por ejemplo, la arquitectura de los edificios).

Foucault destacó las transformaciones en los edificios públicos, que ya no se construyen para ser vistos o vigilar el espacio exterior, sino para permitir un control interno, estructurado y detallado. Por lo tanto, las escuelas, los hospitales, las fábricas, así como las prisiones y los cuarteles, adquieren el carácter de estructuras de vigilancia, mediante construcciones arquitectónicas circulares que permiten

ver sin obstáculos y reconocerlo todo inmediatamente. La misma finalidad se desprende de lo que podríamos llamar «documentos empresariales y administrativos» de la época, por ejemplo los reglamentos de organización de las fábricas. El autor cita, entre otros, el reglamento de una fábrica francesa de 1809, que especifica con el máximo detalle el sistema disciplinar en el trabajo. A partir de estos documentos, de los reglamentos de las fábricas, las escuelas, el ejército y los hospitales, el autor reconstruye el clima de disciplina de la época, que controla las vidas de los alumnos, enfermos, obreros o soldados, hasta el más mínimo detalle.

De las disciplinas humanísticas la historia es la que más confía su conocimiento al estudio de los documentos que han dejado las sociedades a lo largo de los siglos. Pero la sociología también puede encontrar en estos documentos una base empírica válida para el estudio de la realidad contemporánea. Las sociedades modernas producen cada día una enorme cantidad de documentos muy diversos: contratos, demandas y ofertas de trabajo, horarios de trenes, recibos de teléfono, fichas de la policía, programación televisiva, billetes, documentos de identidad, periódicos, anuncios publicitarios, facturas comerciales, certificados de registros civiles, fichas electorales, balances y organigramas empresariales, expedientes asistenciales, registros y redacciones escolares, testamentos, actas de reuniones, sentencias judiciales, señales de tráfico, y otros muchos ejemplos.

Estos documentos son el producto de la vida institucionalizada de nuestra sociedad, y por este motivo los hemos encuadrado bajo el título general de «documentos institucionales». Suelen ser textos escritos, aunque también abarcan otro tipo de material documental. No recogen sólo los hechos «memorables» que se producen en una sociedad o cultura (como tratados de paz entre dos países, leyes o textos literarios), sino que proceden también y sobre todo de la vida cotidiana de las personas comunes. Así, los documentos institucionales pueden convertirse en «datos» para la investigación, material empírico para estudiar los fenómenos sociales más variados: desde los listines de teléfono usados para reconstruir la distribución territorial de los apellidos y mediante ella las raíces geográficas y culturales de las distintas poblaciones regionales, hasta las ventas de entradas de teatro o de cine para estudiar los hábitos de consumo de cultura de la población.

Como ya vimos en el capítulo 7, el Estado registra y sistematiza numerosos documentos sobre actos administrativos para producir las estadísticas oficiales. En las estadísticas oficiales estos documentos se organizan de tal forma que permitan un tratamiento cuantitativo de la información que contienen.

En este capítulo trataremos otro tipo de documentos, menos uniformes y organizados, más adecuados para un análisis cualitativo que cuantitativo. En este sentido conviene precisar que no existe una distinción clara entre los documentos adecuados para el análisis cuantitativo y los adecuados para el análisis cualitativo: muchas veces el mismo tipo de documento se puede analizar con ambos tipos de técnicas. Esto se aplica sobre todo a los documentos escritos o «transcritos» de comunicaciones orales, que constituyen la mayor parte del material documental que presentaremos aquí. En efecto, un texto se puede analizar de forma cualitativa, «interpretándolo» en su globalidad y desde el punto de vista de sus «significados», y de forma cuantitativa, subdividiéndolo en elementos homogéneos para

relacionarlos después entre sí. En particular, el criterio cuantitativo ha dado lugar a una rama de la investigación social, el denominado *análisis de contenido,* que aplica procedimientos de descomposición de los textos para codificarlos en una «matriz de datos» que se analizará después con técnicas estadísticas.

Podemos decir que la diferencia entre el análisis del contenido y el análisis textual de tipo cualitativo es paralela a la diferencia que existe entre los cuestionarios y las entrevistas cualitativas. El análisis del contenido se ha perfilado como un cuerpo de técnicas con entidad propia que habría merecido un apartado específico en este manual. Sin embargo, las limitaciones de espacio nos han impedido dedicarle más atención.

En la parte que sigue no estudiaremos los documentos textuales desde el punto de vista de las técnicas de análisis, sino desde el punto de vista de las fuentes, para ofrecer al lector una idea de dónde puede encontrarse el material documental y cómo puede utilizarse para estudiar los fenómenos sociales más diversos. Por limitaciones de espacio, presentaremos sobre todo ejemplos de análisis cualitativo de documentos.

Empezaremos tratando los textos, para pasar después a los documentos burocráticos y concluir con el material más heterogéneo.

11.3.1. Medios de comunicación

Ya hemos dicho que todas las culturas generan una gran cantidad de documentos; podemos añadir que en las sociedades modernas el principal productor de material documental sobre la propia sociedad son probablemente los medios de comunicación de masas. Todos conocemos la cobertura documental que ofrece la prensa diaria sobre todos los aspectos fácticos y culturales de nuestras sociedades, desde las crónicas de hechos insignificantes hasta la representación de las ideologías y los valores dominantes; también sabemos que la prensa constituye una fuente documental incomparable para el estudio de las sociedades del siglo XX y las últimas décadas del siglo XIX. Del mismo modo, la televisión representa hoy en día un espejo fiel de la sociedad, añadiendo el poder de las imágenes a la ya rica documentación de la prensa escrita.

Se han realizado muchas investigaciones a partir de los medios de comunicación de masas, sobre la producción, el uso y el contenido de sus textos. En este contexto nos limitaremos a hacer algún apunte sobre análisis cualitativos de artículos de periódicos y programas de televisión.

Sobre el empleo de crónicas periodísticas en la descripción de acontecimientos sociales para analizarlos después con los métodos de la antropología cultural, cabe citar el estudio de David Kertzer sobre los ritos de la política. Kertzer afirma que la política se expresa a través del simbolismo y que ésta, sobre todo en las sociedades contemporáneas, se sirve mucho más de instrumentos simbólicos que de la fuerza, para consolidar su poder y conquistar el consenso de las masas. «Por tanto, para entender el proceso político es necesario comprender cómo se introduce la dimensión simbólica en la política» [Kertzer, 1988; 2].

Para hacerlo, el autor analiza distintos ritos de la política: algunos tomados de

las sociedades primitivas (las ceremonias de sucesión de los jefes en tribus de África, la sublimación ritual del conflicto en tribus de Brasil), otros de las sociedades modernas (los funerales de Kennedy en 1963, la visita del presidente Ronald Reagan al cementerio militar alemán de Bitburg en 1985, una manifestación pacifista en Manhattan en 1967, etc.). Para la documentación de estos episodios de crónica moderna utiliza material escrito procedente de periódicos. A continuación reproducimos algunos extractos de la descripción de los funerales de Kennedy:

> Un afuste de cañón, arrastrado por seis caballos grises, avanza hacia la Casa Blanca para recibir el féretro; es el mismo afuste sobre el que se transportó el cuerpo del presidente Roosevelt, menos de veinte años antes [...]. Los tambores, cubiertos por una tela negra y tocados en lenta cadencia [...]. Detrás de la bandera nacional desfila en parada una compañía de la Marina [...]. Desfilan también un sacerdote católico, un ministro protestante, un rabino [...]. Por último, con paso lento, un caballo sin caballero: es el símbolo del líder caído. Está cubierto por una gualdrapa negra; de la silla, dentro de una guaina de plata, cuelga una espada con empuñadura negra [...]. En ese mismo instante en todo el país los ciudadanos suspenden toda actividad [...]. Times Square se sume en el silencio. Los taxistas paran sus vehículos y se bajan, para detenerse en silencio y cabizbajos junto a los coches; los trenes se detienen en pleno campo; los autobuses se acercan a las aceras [...]. [Y por último] en la iglesia de St. Mathew [...] el pequeño John Kennedy junior, de tres años, se une a los soldados y a las fuerzas policiales para dar el último adiós a su padre [...]. En el cementerio, cincuenta aviones, uno por Estado, zumban en las alturas con el fragor de un trueno, formando una uve invertida. A la uve le falta la punta final, el vértice: el líder, el vértice de la nación, se ha ido.
>
> [*ibídem*, 58-61].

Kertzer analiza paso a paso todos los momentos del ritual, explica sus elementos simbólicos y los interpreta. El autor defiende el valor del rito por su capacidad para promover la solidaridad social, analiza el problema de la religión civil estadounidense y la necesidad de lo que él llama la adoración del Estado: «en las sociedades caracterizadas por el pluralismo religioso, los ritos religiosos pueden convertirse en un factor de división social, por ello la adoración al Estado puede proporcionar al pueblo una manera de dar expresión a la propia unidad» [*ibídem*, 65].

La investigación de Kertzer es un buen ejemplo de utilización de los periódicos para los fines de la investigación social. El lector ya habrá percibido la riqueza informativa de la crónica periodística, la abundancia de detalles que, guiado por hipótesis teóricas, el analista experimentado sabe valorar y plasmar en un persuasivo discurso.

Incluso al margen de la parte propiamente descriptiva de la crónica, la prensa es una fuente inagotable de información sobre la sociedad. Pensemos en la publicidad (para estudiar las transformaciones de los tabúes sociales en forma de reclamo), los anuncios de oferta y demanda de empleo (para estudiar la dinámica del mercado laboral), las necrológicas (para estudiar las transformaciones de la imagen social de la muerte), las cartas a los periódicos (para estudiar los cambios en las preocupaciones de los ciudadanos).

Nuestro segundo ejemplo se refiere a los documentos televisivos, que proporcionan también una gran cantidad de información sobre las sociedades, sus acontecimientos o su cultura. Se han realizado diversas investigaciones sobre los distintos tipos de programas o «géneros» televisivos: noticiarios, deportes, variedades, acción, aventura, programas de entrevistas, telenovelas, comedias, concursos, acontecimientos musicales, documentales, dibujos animados, publicidad, etc. Vamos a considerar brevemente una investigación realizada por dos investigadoras estadounidenses sobre las telenovelas, relatos televisivos por capítulos que se suelen emitir en horario de tarde, que se centran básicamente en el mundo de las relaciones sentimentales, «el único ejemplo de entretenimiento televisivo de ámbito nacional dirigido [...] al público femenino» [Cantor y Pingree, 1983] y, como tales, capaces de reflejar el estereotipo femenino de la sociedad y la imagen de la «mujer media» dominante en el entorno de la publicidad y el consumo.

El aspecto más interesante de la investigación es un análisis de la serie *The Guiding Light,* la única telenovela que se emitió por la radio antes de convertirse en serie televisiva y que fue el programa más largo de todos los realizados en Estados Unidos, con una duración total de 45 años (el primer capítulo radiofónico se retransmitió en 1947). Las dos investigadoras analizaron los personajes, el estilo y los argumentos de los diálogos, la ambientación de las escenas, las relaciones interpersonales y los problemas sociales, los patrones morales, los roles de género. El material estudiado, al abarcar un periodo de tiempo tan largo, es una fiel representación de los cambios que fue experimentando con el paso de los años la condición y la imagen de la mujer en la sociedad estadounidense.

Para las autoras la telenovela es una forma de arte popular muy cercana a la realidad y que refleja todas sus transformaciones: como ellas escriben, «la telenovela es un ejemplo ideal de cómo una forma de arte popular puede evolucionar para reflejar las condiciones sociales cambiantes de los propios destinatarios» [*ibídem*, 16].

11.3.2. Narrativa, textos pedagógicos, cuentos populares

Ante todo conviene señalar que la narrativa puede constituir una extraordinaria fuente de conocimiento social, aunque los investigadores sociales no la utilicen mucho. Recordamos la brillante antología de Lewis Coser: *La sociología a través de la literatura* (1963) en la que el autor recoge una serie de pasajes de autores clásicos (Dostoievski, Proust, Mark Twain, etc.) sobre temas puramente sociológicos, como cultura, socialización, estratificación, poder y autoridad, burocracia, comportamientos de masas, desviación, etc., y muestra que la literatura puede ofrecer grandes retratos y análisis sociales agudos. Obviamente se trata de material útil sobre todo para la fase interpretativa de la investigación, que puede ofrecer puntos de partida, ideas, intuiciones e hipótesis para el investigador. Pero tampoco hay que pasar por alto su valor documental potencial: por ejemplo, las descripciones realistas de Émile Zola sobre las condiciones de vida de los mineros en la Francia de finales del siglo XIX en su novela *Germinal* son un documento social de gran valor descriptivo.

En España existen muchos autores cuya obra es una fuente valiosa de conocimiento de la sociedad. Por ejemplo, Benito Pérez Galdós, cuyas novelas describen el mundo de la clase media en el último tercio del siglo XIX. Otro referente de la misma época es Leopoldo Alas, *Clarín,* cuya novela *La Regenta* constituye un cuadro completo de la sociedad del momento. Así mismo, destaca la corriente literaria de los años cincuenta conocida como «novela social». Autores como Alfonso Grosso o Caballero Bonald describen en sus novelas la vida de los campesinos de los latifundios andaluces, así como las vicisitudes de la clase obrera en distintas ciudades de España. Por su parte, autores como Miguel Delibes describen con mucho detalle la realidad vivida por la clase media durante los años sesenta, por ejemplo, en su famosa obra *Cinco horas con Mario.*

Los textos pedagógicos, los libros de preceptos religiosos o morales y los libros de texto escolares son una fuente documental importante sobre la cultura de una sociedad. Dada su finalidad, transmitir a las nuevas generaciones los valores y las normas del mundo de los adultos, son probablemente el espejo más fiel (y a veces más ingenuamente transparente) de la cultura y la ideología de cada época. Se han realizado muchas investigaciones sobre este tipo de material, sobre todo en el marco de la investigación de los procesos educativos. Recordemos una investigación realizada en Italia sobre los libros de lectura de las escuelas primarias, en la que se compararon los textos utilizados en tres épocas distintas: el periodo fascista, la posguerra y los años setenta. Esta investigación mostró claramente las transformaciones de los valores que se produjeron en el periodo estudiado, por ejemplo, la pérdida de importancia de los ideales de la «patria» y el «valor», que en los años treinta ocupaban una posición predominante y en los dos periodos siguientes desaparecen casi por completo, el surgimiento de la «religión» en la Italia democristiana de los años cincuenta, la presencia invariable de los valores ligados a la «familia», etc. [Livolsi *et al.,* 1974, 265].

Para concluir haremos una mención a los documentos de la cultura popular. En todas las sociedades abundan los documentos procedentes de la tradición oral. Una investigación muy conocida sobre este material es *Morfología del cuento,* del investigador ruso Vladimir Propp [1928], quien analizó la estructura de los cuentos de la tradición popular rusa. Su análisis se basa en la idea de que las unidades constitutivas de un cuento no son los personajes y los acontecimientos, sino las «funciones» de estos personajes y acontecimientos, es decir, el significado que cobran para el desarrollo de la historia. Por lo tanto, cada cuento puede reducirse a una sucesión de funciones que se presentan según modelos recurrentes (a pesar de la diversidad de historias, personajes y ambientaciones), y ésta es la aportación más importante del estudio. Por ejemplo, muchos cuentos comienzan con la siguiente sucesión de funciones: alejamiento (uno de los miembros de la familia se aleja de casa), obstáculo (al héroe se le impone una prohibición), infracción (la prohibición se infringe), denuncia (el antagonista recibe información de su víctima), trampa (el antagonista trata de engañar a la víctima), etc. El planteamiento de Propp tiene una gran relevancia porque representó el inicio del análisis estructural del que nacería después el estructuralismo moderno.

11.3.3. Material judicial

El ámbito judicial ofrece abundante material documental para la investigación social. Sentencias judiciales, actas de juicios, transcripciones de interrogatorios, denuncias, etc. constituyen una amplia base documental para el estudio de diversos fenómenos sociales y en especial para el estudio de las conductas desviadas.

Una investigación que se sirve de este tipo de documentos es el estudio de Diego Gambetta [1992] sobre la mafia siciliana. El autor desarrolla el argumento de que, a pesar de presentarse de forma brutal y oscura, la mafia no pertenece a un mundo tan distinto al nuestro, sino que constituye un caso particular de actividad económica específica: es una industria que produce, promueve y vende protección privada. A partir de esta definición, el autor desarrolla un análisis de las actividades mafiosas, desvelando sus mecanismos internos, los contratos entre protector y protegido, las garantías que ofrecen, los servicios que prestan, las retribuciones pactadas, el control del territorio, las relaciones con la autoridad constituida y con la política. Gambetta analiza también la organización de la mafia, los ritos de iniciación, los equilibrios entre los grupos, las «carreras» mafiosas, las alianzas, las normas internas y los códigos de conducta.

El material documental que el autor utiliza para sus reconstrucciones procede en su gran mayoría de transcripciones de testimonios de «arrepentidos» realizados ante los jueces instructores o durante las declaraciones en los juicios que se celebraron en dos ciudades sicilianas, Palermo y Agrigento, en la segunda mitad de los años ochenta. Se trata de documentos llenos de episodios detallados, análisis de los motivos de las conductas de los propios mafiosos, esbozos de interpretación de sus experiencias, que proporcionan al investigador información muy valiosa sobre el fenómeno mafioso. Conviene indicar que la producción de material documental a partir de las declaraciones de «arrepentidos» es un hecho bastante excepcional en el ámbito de los estudios sobre la desviación. No obstante, el caso en cuestión es un buen ejemplo de utilización de estos documentos (y en particular los testimonios de los imputados) en investigación social.

Otra fuente de información de origen judicial son las sentencias dictadas por los tribunales. Éstas suelen ser muy detalladas, reproducen los hechos, las distintas posturas surgidas durante el juicio y las motivaciones de la sentencia. Por ello constituyen una documentación excelente sobre la sociedad y los valores dominantes. Estas sentencias suelen publicarse periódicamente. En España existen una infinidad de bases de datos que recopilan sentencias de los tribunales de justicia: el Tribunal Constitucional, el Tribunal Supremo, los Tribunales Superiores de Justicia, las Audiencias Provinciales y otros tribunales de justicia menores. Normalmente, las bases de datos disponibles se especializan en tipos concretos de jurisprudencia. Por ejemplo, la base de datos *Atlas Fiscal* contiene sentencias de jurisprudencia fiscal, la base de datos *Atlas Laboral* contiene sentencias de jurisprudencia laboral y de la seguridad social, la base de datos *Colex Data Administrativo* contiene sentencias de jurisprudencia contencioso-administrativa, la base de datos *Colex Data Civil* contiene sentencias de jurisprudencia civil. Por último, cabe destacar dos bases de datos que contienen sentencias de todas las materias: la *Base de datos Aranzadi* y *El*

derecho: base de datos de jurisprudencia. Ambos ordenan las sentencias según el tribunal de procedencia. En general, todas las bases de datos citadas se actualizan cada dos o tres meses.

Un ejemplo de utilización de este material para la investigación social sería el estudio que Paul Robertshaw (1985) llevó a cabo en el Reino Unido sobre la discriminación sexual en la cultura judicial. El autor analizó 76 juicios celebrados en el país en la segunda mitad de los años setenta relativos a discriminación sexual en el trabajo. Analizando sobre todo el lenguaje empleado en el juicio y en la documentación escrita, el autor puso en evidencia el prejuicio antifemenino que impregnaba el sistema judicial británico en todos los ámbitos, desde los tribunales locales hasta las cortes de apelación y la cámara de los lores.

En este caso, los documentos judiciales no se utilizaron para estudiar conductas ilegítimas, sino para arrojar luz sobre la ideología de la sociedad. En otros casos se han empleado para analizar el fenómeno de la delincuencia. Por ejemplo, Bourget, Gagne y Moamai (2000) analizaron 387 casos de homicidio entre cónyuges registrados en Québec entre 1991 y 1998. Los expedientes judiciales se emplearon como material empírico para el estudio de los casos individuales, y revelaron que la mayoría de los crímenes son cometidos por hombres que están separados de sus mujeres y que a menudo presentan cuadros clínicos de depresión.

11.3.4. Documentos de la política

El material documental que ofrece la política al investigador social es muy rico: basta pensar en los actos parlamentarios, los programas de los partidos, los discursos de los políticos, la propaganda electoral. Los documentos de la política fueron de los primeros que interesaron a los investigadores y que se utilizaron para la investigación social. Por ejemplo, las primeras investigaciones de análisis del contenido tuvieron como objeto la propaganda política.

Uno de los primeros estudios fue el realizado por Lasswell y Leites [1949] sobre los eslóganes utilizados en la Unión Soviética para la celebración del 1 de mayo entre 1918 y 1943, eslóganes que los autores clasificaron en función del contenido en 11 categorías: símbolos «revolucionarios», como internacional comunista, socialismo, revolución mundial; símbolos «nacionales», como madre patria, patriotismo, defensa, enemigo, agresión; símbolos «de personas», como Lenin y leninismo, Marx, Engels, Luxemburg, Hitler, etc. A continuación analizaron la evolución en el tiempo de la importancia de estos símbolos, por ejemplo, constataron que con el paso de los años y la progresiva institucionalización de la república disminuía la presencia de los símbolos revolucionarios y aumentaba la de los nacionales, disminuían los símbolos «morales», como la solidaridad, la disciplina y el honor, y aumentaban los «liberales», como la fraternidad, la libertad y la ciudadanía. Los autores utilizaron así los eslóganes del partido como indicador de la línea del mismo y del clima político de la época.

Entre los estudios más recientes podemos mencionar la investigación realizada por Jeffrey Tulis [1987] sobre la evolución, a lo largo de la historia de Estados Unidos, de la «retórica presidencial». El autor estudió los «llamamientos a la na-

ción» de los presidentes, es decir, esa forma de comunicación que consiste en hablar directamente al pueblo, sin ninguna mediación institucional (como el Parlamento). Tulis analiza los discursos de los presidentes que se han sucedido en la historia de Estados Unidos y revela la profunda transformación que se produjo entre los siglos XIX y XX, demostrando que la retórica de la apelación presidencial es un aspecto típico de la política estadounidense del siglo XX.

En el siglo XIX los discursos del presidente eran de carácter predominantemente ceremonial, no patriótico, y sobre todo no pretendían movilizar a la opinión pública para influir en el proceso legislativo. El autor constata una marcada transformación entre los siglos XIX y XX en lo que respecta al número de apelaciones presidenciales, la forma (paso de la comunicación escrita a la oral), su destinatario (el Congreso en el siglo XIX, y el pueblo en el siglo XX), etc., y afirma que las nuevas tecnologías de comunicación y en particular la llegada de la televisión explican solo una parte de esta transformación (si bien la televisión favorece notablemente la eficacia de la comunicación directa), y que ésta se puede atribuir en mayor medida a una transformación de la cultura política del país.

Obviamente, los periodos de campaña electoral son un momento de gran producción de documentos sobre la vida política de un país. Entre la multitud de estudios que han analizado las campañas electorales, los programas de los partidos o la propaganda política, podemos citar la extensa investigación realizada en los años ochenta por tres investigadores británicos sobre los programas electorales de los partidos de 19 democracias desde el final de la Segunda Guerra Mundial hasta 1980 (Budge, Robertson y Hearl, 1987). Utilizando la documentación oficial de los partidos o, cuando ésta no estaba disponible, artículos de periódicos, los autores reunieron los programas electorales de unos 90 partidos. Analizaron los contenidos de los programas, clasificando los temas que éstos mencionaban en 54 categorías (como descolonización, paz, justicia social, moralidad, planificación económica, etc.), y analizaron las diferencias entre los partidos y entre las naciones, así como su transformación en el curso del tiempo. Por ejemplo, llegaron a la conclusión de que, en casi todos los países, había una división clara entre la izquierda y la derecha que podía reconocerse en todos los programas. Otros temas que abordaron en su investigación fueron la convergencia/divergencia de los programas de los partidos en el tiempo, la clasificación de los partidos en familias ideológicas y los cambios en las semejanzas/diferencias entre los países con el paso del tiempo.

Entre los estudios sobre propaganda electoral y comunicación política podemos destacar la investigación realizada por Roderick Hart (2000) sobre las campañas electorales en Estados Unidos. El autor analizó las elecciones presidenciales durante un largo periodo de tiempo (1948-1996), basándose en documentos como la transcripción de los discursos y los debates, reportajes de televisión y retransmisiones de propaganda electoral, artículos de periódico e incluso cartas de los ciudadanos a los periódicos. El objetivo de la investigación era identificar los cambios en la retórica de la comunicación política que se produjeron en un periodo de cincuenta años.

11.3.5. Documentos empresariales y administrativos

En este apartado abordaremos los documentos generados por organizaciones relacionadas con el trabajo. A lo largo de su existencia, empresas, escuelas, hospitales, colegios profesionales, sociedades anónimas, etc., producen toda una serie de documentos que pueden dar una imagen completa del fragmento de sociedad que representan: balances, cartas, circulares, organigramas, listas de socios, inventarios de almacén, declaraciones tributarias, actas de reuniones, resúmenes anuales, informes financieros, etc.

Se trata de material muy heterogéneo. Ciertos tipos de documentos, como escriben Atkinson y Coffey [1997: 48], «constituyen, utilizando una analogía literaria, *géneros,* con sus estilos específicos y sus convenciones», en el sentido de que hay ciertos lenguajes especializados que se asocian con partes específicas de nuestra vida cotidiana. Podríamos decir que, así como existen los géneros literarios, existen también *géneros burocráticos*. Pensemos en el lenguaje que se utiliza para la presentación de una exposición de arte o en la forma de describir el *buqué* de un determinado vino, o incluso en un balance empresarial, o el prototipo de lenguaje críptico, accesible sólo para quienes lo conocen, el lenguaje médico.

Se han realizado numerosas investigaciones a partir de estos documentos. Ilustraremos brevemente algunas de ellas para que el lector pueda hacerse una idea del potencial de este material.

Empezaremos por el mundo empresarial en sentido estricto. En este campo destacan los estudios de Max Weber, que ante las profundas transformaciones que produjo la implantación de la industria en la sociedad, se preguntó qué consecuencias tendría «en el ámbito personal, el destino profesional, o el "estilo de vida" extraprofesional de los obreros», así como «qué cualidades físicas y psíquicas desarrollaría en ellos y cómo se manifestarían éstas en la conducta de los trabajadores en todos los aspectos de su vida» [Weber, 1924]. Y para profundizar en estos interrogantes realizó en 1908 una investigación sobre lo que él llama la «psicofísica del trabajo industrial», en una fábrica textil de Westfalia. Weber estudió, por ejemplo, las variaciones en la productividad obrera, relacionándola con las variaciones climáticas, con el tipo de retribución (por el trabajo hecho o por horas), con las características personales de los obreros (diferencias entre hombres y mujeres, entre jóvenes y adultos, entre casados y solteros, etc.), con los horarios de trabajo, con el tipo de máquinas empleadas (diferentes tipos de telares), etc. Lo que nos interesa aquí es el hecho de que él realizó su investigación a partir de documentos empresariales, pasando, como escribe su viuda Marianne Weber, «largas semanas estudiando los libros contables y los registros de la producción» [*ibídem*, 67].

Los balances de las administraciones públicas son una base documental importante para el estudio de las políticas públicas. A principios de los años ochenta se realizó en Italia una investigación que empleó este tipo de material documental y en particular los balances municipales [Felicori, 1984]. Las elecciones municipales de 1975-76 permitieron por primera vez a la izquierda hacerse con la administración de muchas ciudades importantes: Roma, Milán, Turín, Venecia, Florencia, etc. Entonces se planteó la necesidad de estudiar las innovaciones que introdujo la gestión

de la izquierda, en especial en las actividades relacionadas con la cultura.

En estos años se consolidó lo que se llamaría actividades culturales «efímeras» en los municipios: grandes eventos espectaculares con una amplia participación de masas, con carácter de fiesta colectiva, en los que la administración pública se convertía en industria del espectáculo, poniendo a disposición sus capitales y sus plazas. Entre los interrogantes que se planteó la investigación estaba el de comprobar si, en comparación con años anteriores, esta nueva política cultural estaba efectivamente presente en todos los municipios o si se sólo limitaba a los grandes municipios, que atraían una atención mayor, y si efectivamente se trataba de una política propia de la gestión de partidos de izquierda o más bien de una tendencia general que había afectado a todo el país.

La investigación consiguió llamar la atención sobre estos interrogantes a partir de un análisis de los balances municipales, que permitían diferenciar, dentro de los presupuestos destinados a la cultura, los que se usaban para financiar las «fiestas populares» conocidas como «cultura efímera», de los que se usaban para financiar actividades tradicionales (museos, bibliotecas, teatros, etc.).

El mundo de la escuela también ofrece una gran cantidad de material documental útil para la investigación social. Por ejemplo, las redacciones de los alumnos, que se conservan durante mucho tiempo en los archivos de los centros escolares, podrían emplearse para estudiar las transformaciones de la cultura juvenil con el paso del tiempo.

En los años setenta, un sociólogo inglés, Peter Woods, realizó una investigación en un instituto de una pequeña ciudad industrial del Reino Unido. Durante su estudio, realizado principalmente con observación participante, Woods también utilizó material documental archivado en la escuela. Por ejemplo, estudió el proceso de formación de las evaluaciones de los profesores sobre los alumnos y las categorías mentales empleadas para este fin, mediante las «fichas», que los profesores elaboraban y enviaban a los padres de forma periódica (al final de cada trimestre o cuatrimestre), y en las que se indicaba para cada estudiante y para cada asignatura, además de la nota, una breve evaluación (por ejemplo: «es muy trabajador», «le falta voluntad para mejorar», etc.). Woods utilizó estas evaluaciones para analizar la cultura de los profesores y su forma de evaluar. Según el autor, «en ningún otro documento queda reflejado el sistema de categorización de los profesores de forma tan clara y concisa como en las fichas escolares», [Woods, 1979, 171]. Woods realiza un análisis del contenido de estas evaluaciones, explica sus conceptos y sus «tipos ideales» de referencia, y asegura que este proceso de clasificación rápida de los alumnos con estereotipos más bien toscos es típico de las instituciones, que están obligadas a elaborar unos procedimientos estandarizados para «garantizar un tratamiento de masas lo más eficaz posible».

El sistema sanitario produce una gran cantidad de documentos que se pueden utilizar para la investigación social. Un ejemplo de ello sería la investigación realizada por Lindsay Prior sobre la forma de clasificar las enfermedades y las causas de muerte de la Organización Mundial de la Salud. El autor afirma que los sistemas de clasificación no son la representación exacta de un orden externo inherente a la realidad objetiva, sino una expresión de las reglas y los temas que predomi-

nan en un determinado contexto sociocultural. Corresponde al investigador cualitativo la tarea de descubrir hasta qué punto es la proyección de la cultura de la sociedad más que el reflejo de una realidad externa. Bajo este punto de vista, Prior analiza la clasificación de las causas de muerte y su evolución en el tiempo (está disponible a partir de 1903). El autor se pregunta cuándo desaparecieron como causa de muerte la vejez y la decadencia física e intelectual. El concepto de las enfermedades que tenemos actualmente es enormemente «somático», está centrado en el cuerpo como sistema biológico, en los síntomas, más que en las causas. Prior se pregunta también si en algún momento de la historia la «pobreza» se ha incluido entre las causas de muerte.

11.3.6. Huellas físicas

Hasta ahora hemos hablado sólo de documentos escritos. Pero el hombre deja también huellas físicas, *documentos materiales* que permiten deducir las actividades que los han producido. Ésta es la tarea del detective, que a partir del hallazgo de un cadáver en el bosque y observando el entorno circundante, la posición de la víctima, la disposición de la ropa, las huellas sobre el terreno, fragmentos de tejido, etc., puede reconstruir con todo detalle el asesinato. Igual que, de un análisis de los restos de una fiesta celebrada la noche anterior, se puede deducir su desarrollo, el número y la edad de los participantes, las actividades realizadas, su duración, las bebidas consumidas, e incluso hechos particulares como una pelea, etc.

En los años sesenta, Webb y otros autores [1996; 2.ª ed. 1981], en una conocida obra que abordó esta cuestión por primera vez, distinguieron dos categorías de huellas físicas: las *huellas de erosión* y las *huellas de incremento*. Se producen huellas de erosión siempre que una actividad humana provoca el desgaste de un determinado soporte físico. Y la medida de este desgaste puede servir de indicador de la actividad que lo ha producido. Los autores señalan, por ejemplo, la posibilidad de medir la cantidad de público que frecuenta un determinado lugar observando el desgaste del suelo, y sugieren que se emplee esta técnica para estudiar, por ejemplo, la popularidad de las distintas salas de un museo. Los autores citan el caso del Museo de la Ciencia de Chicago, en el que había una sección donde se exponían pollitos vivos y donde el revestimiento del suelo debía renovarse cada seis meses, mientras que en otras partes del museo el suelo se conservaba en buen estado durante años. Del mismo modo, la popularidad de los libros de una biblioteca puede medirse por el deterioro de las páginas (bordes desgastados, marcados, arrugados y amarilleados por el uso). Un ejemplo de ello es la *International Encyclopedia of the Social Sciences* de una biblioteca universitaria, que presentaba partes distinguibles a simple vista por el color distinto de las páginas, a causa de su mayor uso [*ibídem*, 7-9].

Las huellas de incremento aparecen cuando la actividad realizada provoca una acumulación de material. Podemos citar los estudios sobre la basura de Rathje, que organizó equipos que retiraban la basura doméstica de una muestra de población y la clasificaban, y llegó a conclusiones interesantes sobre los estilos de vida

de las familias [Rathje y Murphy, 1992]. También podemos recordar en este contexto los estudios sobre productividad y sobre la forma de trabajar de un departamento de mantenimiento de máquinas a partir del polvo acumulado sobre las piezas que se revisaban, o la operacionalización de la situación de aislamiento del mundo exterior (o de la interacción mantenida con el mismo) de los internos de un hospital psiquiátrico, a partir de la cantidad de correo que recibían [Webb *et al.*, 1966].

Se han realizado varios estudios sobre las pintadas en las paredes, que, aunque se realicen con un fin de mera diversión, pueden revelar impulsos más profundos. Por ejemplo, Kinsey, en su notable estudio sobre la sexualidad de los estadounidenses, encargó recoger varios centenares de inscripciones de las paredes de los aseos públicos, y encontró diferencias significativas en las inscripciones eróticas entre los aseos para hombres y mujeres (lo que confirmó su tesis de que en la sociedad contemporánea existe una cultura sexual diferente entre hombres y mujeres) [Kinsey *et al.*, 1953]. Se han realizado otros estudios similares a éste en algunos aspectos sobre los *graffiti* y las pintadas en las paredes que han acompañado a los movimientos de protesta social, como los movimientos estudiantiles de los años sesenta y setenta.

Entre otros signos físicos que no se pueden encuadrar dentro de las categorías de la erosión y el incremento podemos citar dos ejemplos: la presencia de señales de circulación vial y avisos al público en idiomas distintos del oficial, como signo de reconocimiento y aceptación de la presencia de etnias distintas o la disposición de las plazas para ciudadanos blancos y negros en una reunión pública como indicador de integración racial [Bouchard, 1976, 283].

Concluimos este epígrafe con un apunte sobre los documentos visuales, cuyo tratamiento requeriría en realidad un epígrafe específico dedicado a la *sociología visual,* ya que estos documentos son difíciles de encuadrar bajo la categoría de los documentos materiales. La fotografía, el cine y la televisión, además de objeto de estudio como producto cultural o formas de comunicación de masas, se pueden utilizar también como instrumentos de investigación, como documentos sociales de un valor cognitivo relevante. La historia de las imágenes de lo social comienza con la fotografía a partir del año 1830 aproximadamente.

Durante casi un siglo la realidad social se representó mediante fotografías: retratos, fotos de guerra, documentos sobre las grandes metrópolis de la civilización industrial, reportajes de viajes a los países exóticos, fotos de recuerdo de la vida cotidiana y familiar. A continuación viene el cine, después la televisión y más recientemente las cámaras de vídeo, que, al permitir al propio investigador grabar los fenómenos sociales en directo, añaden una nueva herramienta de investigación a su instrumental tradicional. Efectivamente, una reflexión sobre esta cuestión debería considerar tanto la aportación histórico-documental de los documentos visuales (fotografías y reportajes sobre acontecimientos ya sucedidos), como las nuevas posibilidades que ofrece el hecho de poder grabar movimientos y secuencias «sobre el terreno», proporcionando al investigador nuevas formas para observar y analizar la sociedad.

11.4. LEER LOS DOCUMENTOS HUMANOS

11.4.1. Leer los documentos personales

Tal y como hemos apuntado al hablar de otras técnicas cualitativas, en última instancia los científicos pueden adoptar dos enfoques a la hora de abordar los documentos personales.

El primer enfoque es fundamentalmente «reductivo». No puede negarse el atractivo literario y el poder evocador de las imágenes de los documentos personales, pero, al tratarse de relatos con una finalidad única e invariable, los documentos personales quedan relegados a un ámbito precientífico, y como mucho se les reconoce una función integradora (o a lo sumo «ilustrativa») que sirve de apoyo a los avances teóricos obtenidos por otra vía.

El propio Herbert Blumer, uno de los teóricos más importantes del enfoque cualitativo en investigación social, no oculta la utilidad de los documentos personales: «ayudan al investigador a conocer mejor el tipo de realidad que está estudiando [...], ofreciéndole la oportunidad de ampliar su experiencia y mejorar su imaginación sociológica [...]» [Blumer, 1939, 80-81]. Pero, según Blumer, los documentos personales no son suficientes para validar una teoría científica: «el contenido interpretativo depende en gran medida de la preparación y del esquema teórico con los que se lee el documento. Una persona, en función de su experiencia y sus intereses, puede descubrir en un documento cosas que otra podría no percibir [...]; los documentos humanos se prestan fácilmente a diversas interpretaciones. Esto puede comprobarse por la facilidad con la que pueden analizarse utilizando diferentes teorías. Las teorías parecen capaces de ordenar los datos» [*ibídem*, 76-77].

En definitiva, este enfoque implica un pesimismo de fondo que afirma la subjetividad de las interpretaciones basadas en los documentos personales, declarando explícitamente su «insuficiencia para validar una teoría científica».

La otra postura es la de los defensores más convencidos del criterio interpretativo. Según Schwartz y Jacobs, por ejemplo, «por un lado existe un *mundo subjetivo* y, por otro, la experiencia individual de este mundo» [1979; 61]. Ambos mundos merecen el interés del investigador social. Pero si éste decide estudiar el segundo, también el estudio de un solo caso se convierte en fuente relevante de información empírica:

> El método idiográfico [3] (empleado en el enfoque [...] «cualitativo») subraya la legitimidad de la investigación que estudia la vida de un solo individuo [...]. Los defensores de este método afirman que es válido desde el punto de vista científico y correcto desde el punto de vista metodológico estudiar el comportamiento de un ser humano y considerarlo no sólo representativo de un grupo, sino también una totalidad independiente, de la que pueden extraerse generalizaciones.

[3] Método idiográfico y método nomotético: ver capítulo 1, apartado 5.

> Por otra parte, afirman que el auténtico objetivo de las ciencias sociales es *comprender* el comportamiento, no cuantificar, clasificar o producir unos modelos de comportamiento. Están convencidos de que este resultado puede lograrse empleando sobre todo el método del análisis de los casos individuales, dado que éste permite estudiar la totalidad de una única personalidad.
>
> [*ibidem*, 69].

Ello quiere decir que si nuestro objetivo no es explicar los hechos sino comprenderlos, entonces también el estudio de casos únicos puede arrojar luz sobre la realidad que estamos estudiando y favorecer nuestra comprensión. En otras palabras, podemos decir que aunque los documentos personales resultan ciertamente inadecuados para la *descripción* y probablemente también para la *explicación* de los fenómenos sociales, pueden ofrecer aportaciones fundamentales para su *interpretación*.

Por último, no olvidemos que los documentos personales se llaman así porque se refieren a la «vida personal» de los individuos estudiados. La reconstrucción de las vicisitudes subjetivas de las personas estudiadas también permite comprender cómo esas circunstancias biográficas únicas han podido producir determinados comportamientos. En este sentido los documentos personales, y en particular los historiales de vida, permiten obtener una visión diacrónica de los sujetos estudiados que no se puede obtener con otros instrumentos. El análisis de estos historiales hace posible reconstruir itinerarios de vida que en algunos casos podrían ser el único instrumento que permite comprender la forma en que ciertas personas han llegado a ser lo que son y se han comportado de una forma concreta, lo que aporta una contribución significativa y única a la interpretación de los fenómenos sociales.

11.4.2. Leer los documentos institucionales

Toda actividad humana reconocible en nuestra sociedad produce documentos. La sociedad moderna se autodocumenta continuamente: podemos decir que no hay acto institucional o actividad organizada que no deje huellas documentales, lo que quiere decir que no hay fenómeno de nuestra vida social que no pueda estudiarse mediante el análisis de documentos.

Las ventajas más evidentes de la utilización de documentos para los fines de la investigación social pueden enunciarse bajo tres títulos: *a*) no reactividad; *b*) posibilidad de análisis diacrónico; y, *c*) costes reducidos. Ya hemos tratado los dos primeros puntos al principio de este capítulo. En cuanto al tercero, se trata de una peculiaridad sobre todo de los documentos institucionales. Por definición, el uso de documentos no implica costes de producción de la información (como, por ejemplo, en el caso de una entrevista), puesto que se trata de emplear material ya producido. Hay que añadir que las instituciones producen gran cantidad de documentos, que cuando son accesibles suelen agruparse en lugares bien definidos; por tanto, el coste añadido de registro de cada documento es bajo, lo que permite obtener muestras muy numerosas con un coste discreto.

También son evidentes los inconvenientes ligados al uso de los documentos ins-

titucionales, que podemos resumir en dos puntos: *a*) el carácter incompleto de la información; y, *b*) el carácter oficial de la representación. Sobre el primer punto remitimos de nuevo a lo dicho en el primer apartado de este capítulo (al producirse para fines distintos de los objetivos de la investigación, la información contenida en estos documentos suele ser escasa e insuficiente, y tampoco puede ser integrada, ya que el documento es «encontrado» y no producido por el investigador). En cuanto al segundo punto, hay que decir que los documentos no suelen ser representaciones objetivas de la realidad institucional a la que se refieren, sino una representación «oficial» de la misma. Así, los balances de los partidos o los de una empresa se elaboran para responder a la normativa vigente, y suelen ofrecer una imagen distorsionada del cuadro empresarial real de la organización. Del mismo modo, no será posible estudiar la distribución real del poder en una empresa basándose únicamente en el organigrama empresarial, y las redacciones escolares escritas por los niños sobre lo que piensan de la patria o de la escuela evocarán muchas veces las convenciones sociales más que la opinión genuina de los niños.

En otras palabras, los documentos institucionales reflejan exactamente lo que su definición sugiere, es decir, la dimensión institucional de los fenómenos estudiados, que ciertamente es distinta de la dimensión personal. Recordemos de nuevo que existe una diferencia clara en la forma de considerar la realidad social entre los documentos personales y los documentos institucionales, aunque obviamente, esta diferencia no significa inferioridad de ninguna de las dos perspectivas.

11.5. SÍNTESIS DEL CAPÍTULO 11

1. Las tres acciones básicas que fundamentan las técnicas de investigación cualitativa son *observar, preguntar* y *leer*. En este capítulo abordamos la tercera: leer los documentos producidos por los individuos y las instituciones. Un *documento* se define como cualquier material que proporciona información sobre un determinado fenómeno social y que existe al margen de la acción del investigador.
2. Los *documentos personales* tienen *naturaleza privada,* dado que son producidos por el individuo para su uso personal. También se denominan documentos «expresivos», puesto que expresan los sentimientos, los asuntos y, más en general, la personalidad de quien los produce. Estos se clasifican en *autobiografías* (relatos sobre toda la vida de la persona, escritos en un periodo limitado de tiempo, y por lo tanto con una visión retrospectiva), *diarios* (escritos para uso estrictamente personal y paralelamente a los hechos), *cartas, testimonios orales* (relatos de experiencias personales solicitadas por el investigador). Los testimonios orales pueden dividirse en *historiales de vida* (narración de la vida de una persona que hace ésta ante el entrevistador) e *historia oral* (relatos de acontecimientos específicos de los que el sujeto ha sido testigo).
3. Los *documentos institucionales* son producidos por las instituciones o por individuos en el contexto de sus funciones institucionales. A diferencia de los do-

cumentos personales, tienen *naturaleza pública*. Según su procedencia se clasifican en: medios de comunicación de masas, narrativa, textos pedagógicos, cuentos populares, material judicial, documentos políticos, documentos empresariales y administrativos, y huellas físicas.

11.6. LECTURAS COMPLEMENTARIAS

Para una introducción bastante completa sobre el uso de los documentos en general, como diarios, cartas, autobiografías, historia oral, fotografías, películas, etc., remitimos a la obra de K. Plummer *Documents of Life* (Allen y Unwin, 1983, 175 pp.).

Con respecto a los relatos personales, encontramos una reseña sintética sobre su uso en investigación social en T. L. Orbuch «People's AcCounts count: The Sociology of Accounts» *(Annual Review of Sociology, 1997)*. Sobre el uso de la historia biográfica y familiar en la investigación social, desde una perspectiva amplia que tiene en cuenta los distintos planteamientos metodológicos realistas, neopositivistas y narrativos, véase la obra de R. L. Miller, *Researching Life Stories and Family Histories* (Sage, 2000, 192 pp.).

En relación con los documentos institucionales, una introducción sobre el uso de las fuentes documentales en la investigación social se encuentra en la obra de J. Scott *A Matter of Record: Documentary Sources in Social Research* (Polity Press 1990, 233 pp.). Por lo que se refiere a la cultura material, huellas físicas, documentos audiovisuales, etc., remitimos al clásico de E. J. Weebb, D. T. Campbell, R. D. Schwartz y L. Sechrest *Unobtrusive Measures* (edición revisada, Corwin Press, 1999, 240 pp.). Una obra más reciente sobre el mismo tema es la de A. Kellehear *The Unobtrusive Researcher* (Allen y Unwin 1993, 177 pp.). Por último, señalamos una reseña sobre la «sociología del lugar» (forma material de la ciudad, contexto natural y físico, tecnología y transportes, paisaje urbano) en T. F. Gieryn «A Space for Place in Sociology» (*Annual Review of Sociology, 2000,* 26:463-96).

BIBLIOGRAFÍA

Abbagnano, N. (1971): *Dizionario di filosofia,* Torino, Utet.

Adorno, T. W, Frenkel-Brunswik, E., Levinson, D. J. y Nevitt, R. (1950): *The Authoritarian Personality,* New York, J. Wiley; citas de la edición de 1964.

Alberoni, F. et al. (1967): *L'attivista di partito,* Bologna, Il Mulino.

Alvira Martín, F. (2004): *La encuesta: una perspectiva general metodológica,* Madrid, CIS.

Anderson, N. (1923): *The Hobo,* Chicago, The University of Chicago Press.

Arendell, T. (1986): *Mothers and Divorce,* Berkeley, University of California Press.

Aronson, E., Brewer, M. y Carlsmith, J. M. (1985) «Experimentation in Social Psychology», en G. Lindzey y E. Aronson (eds.), *Handbook of Social Psychology,* Nueva York, Random House.

Atkinson, P. (1995): *Medical Talk and Medical Work,* London, Sage.

Atkinson, P. y Coffey, A. (1997): *Analyzing Documentary Realities,* en D. Silverman, *Qualitative Methodology and Sociology,* Aldershot, glover.

Babbie, E. R. (1979): *The Practice of Social Research,* Belmont, Wadsworth.

Bales, R. F. (1951): *Interaction Process Analysis,* Reading, Addison-Wesley.

Bandura, A., Ross, D. y Ross, S. A. (1963) «Imitation of Film Mediated Aggressive Models», *Journal of Abnormal and Social Psychology,* 66: 3-11.

Banfield, E. C. (1958): *The Moral Basis of a Backward Society,* Glencoe, The Free Press.

Barbagli, M. (1995): *L'occasione e l'uomo ladro,* Bologna, Il Mulino.

Bastin, G. (1961): *Les techniques sociométriques,* París, Puf.

Bayley, K. D. (1994): *Methods of Social Research,* New York: Free Press.

Bechhofer, F. (1974): «Current Approaches to Empirical Research: Some Central Ideas», en J. Rex (ed.), *Approaches to Sociology: an Introduction to Major Trends in British Sociology,* London, Routledge y Kegan, Paul.

Becker, H. (1963): *Outsiders. Studies in the Sociology of Deviance,* Glencoe, Free Press.

Bernardi, F. (2006), *Análisis de la historia de acontecimientos,* Madrid, CIS.

Biemer, P. P., Groves, R. M., Lyberg, L. E., Mathiowetz, N. A. y Sudman, S. (eds.) (1991): *Measurement Errors in Surveys,* Nueva York, John Wiley.

Bishop, G. F., Tuchfarber, A. J. y Oldendick, R. W. (1986): «Opinions in Fictious Issues: The Pressure to Answer Survey Questions», en *Public Opinion Quarterly,* 50: 240-250.

Black, T. R. (1999): *Doing Quantitative Research in The Social Sciences: An Integrated Approach to Research Design, Measurement and Statistics,* Londres, Sage.

Blaikie, N. (1993): *Approaches to Social Inquiry*, Cambridge, Polity Press.

Blalock, H. M. (1961): *Causal Inferences in Nonexperimental Research,* Chapel Hill, The University of North Carolina Press.

Blalock, H. M. (1970): *An Introduction to Social Research,* Englewood Cliffs, Prentice Hall: las citas son de la edición en lengua original. Hay traducción al español: *Introducción a la investigación social* (1970), Buenos Aires, Amorrortu.

Blau, P. y Duncan, O. D. (1967): *The American Occupational Structure,* Nueva York, Wiley.

Blumer, H. (1939): *An Appraisal of Thomas and Znaniecki's «The Polish Peasant in Europe and America»,* Nueva York, Social Science Research Council, Bulletin 44.

Blumer, H. (1969): *Symbolic Interactionism: Perspective and Method,* Englewood Cliffs, Prentice Hall: las citas son de la edición en lengua original. Hay traducción al español. *El interaccionismo simbólico: perspectiva y método* (1982). Barcelona hora cop.

Bogardus, E. S. (1925): «Measuring Social Distance», en *Journal of Applied Sociology,* 9: 299-308.

Boltanski, L. (1990): *L'Amour et la Justice comme compétences,* París, Métailié.

Borman, K. M., LeCompte, M. D. y Goetz, J. P. (1986): «Ethnographic and Qualitative Research Design and Why It Doesn't Work», en *American Behavioral Scientist,* 30: 42-57.

Bottomore T. B. y Rubel, M. (1956): *Karl Marx: Selected Writings in Sociology and Social Philosophy,* Nueva York, McGraw Hill.

Bouchard, T. J. (1976) «Unobtrusive Measures: An Inventory of Uses», *en Sociological Methods and Research,* 4: 267-300.

Boudon, R. (1984): *La place du désordre,* París, Presses Universitaires de France; citas de: *Theories of Social Change,* Cambridge, Polity Press, 1986.

Bourget, D., Gagne, P. y Moamai, J. (2000): «Spousal Homicide and Suicide in Quebec», en *Journal of the American Academy of Psychiatry and the Law,* 28(2): 179-182.

Bryman, A. (1988): *Quantity and Quality in Social Research,* Londres, Routledge.

Bryman, A. y Burgess, R. G. (1994): *Analyzing Qualitative Data,* London, Routledge.

Budge, I., Robertson, D. y Hearl, D. (1987): *Ideology, Strategy and Party Change,* Cambridge, Cambridge University Press.

Bunge, M. (1959): *Causality,* Cambridge, Harvard University Press.

Burawoy, M. (1979): *Manifacturing Consent: Changes in the Labor Process*

under Monopoly Capitalism, Chicago, University of Chicago Press.
Callejo, J. (2001): *El grupo de discusión: introducción a una práctica de investigación,* Barcelona, Ariel Practicum.
Campbell, A., Gurin, G. y Miller, W. E. (1954): *The Voter Decides,* Evanston, Row and Peterson.
Campbell, D. T. y Stanley, J. C. (1963): *Experimental and Quasi-experimental Designs for Research,* Chicago, Rand McNally.
Cantor, M. G. y Pingree, S. (1983): *The Soap Opera,* Beverly Hills, Sage.
Cantril, H. y Free, L. A. (1962): «Hopes and Fears for Self and Country: the Self-Anchoring Scale in Cross-Cultural Research», en *American Behavioral Scientist,* 6, suppl. October: 1-30.
Capecchi, V. y Livolsi, M. (1973): *Televisione e bambini,* Roma, Rai - Servizio opinioni.
Caplow, T., Hicks, L. y Wattenberg, B. J. (2001): *The First Measured Century: An Illustrated Guide to Trends in America, 1900-2000,* Washington D. C., AEI Press.
Caplow, T., Bahr, H. M., Modell, J. y Chadwick, B. A. (1991): *Recent Social Trends in the United States, 1960-1990,* Montreal y Kingston, McGill-Queen's University Press.
Carley, M. (1981): *Social Measurement and Social Indicators,* London, George Allen y Unwin, 1981.
Carmines, E. G. y Zeller, R. A. (1983): *Reliability and Validity Assessment,* Newbury Park, Sage.
Cartocci, R. (1993): «Rilevare la secolarizzazione: indicatori a geometria variabile», en *Rivista italiana di scienza politica,* 23: 119-152.
Cavan, R. S. (1928): *Suicide,* Nueva York, Russell y Russell.
Chirban, J. T. (1996): *Interviewing in Depth: The Interactive-Relational Approach,* Newbury Park, Sage.
Cicourel, A. V. (1964): *Method and Measurement in Sociology,* Nueva York, Free Press.
Clayman, S. E. (1991): «The Diversity of Ethnomethodology», en *Annual Review of Sociology, 1991,* 17: 385-418.
Clayman, S. E. (1993): «Booing: the Anatomy of a Disaffiliative Response», en *American Sociological Review,* 58: 110-130.
Cochran, W. G. (1977): *Sampling Techniques,* 3.ª ed., New York, Wiley. Hay traducción al español: *Técnicas de muestreo* (1987), Compañía Editorial Continental.
Cole, R. (1981): *Work, Mobility and Participation,* Berkeley, University of California Press.
Collett, P. y Marsh, P. (1974): «Patterns of Public Behaviour: Collision Avoidance on a Pedestrian Crossing», *Semiotica,* 12: 281-299.
Collins, M., Sykes, W., Wilson, P. y Blackshaw, N. (1988): *Nonresponse: the UK Experience,* en Groves *et al. Telephone Survey Methodology,* Nueva York, Wiley.
Collins, R. (1988): *Theoretical Sociology,* Orlando, Harcourt Brace Jovanovich.

Converse, P. E. (1970): «Attitudes and Non-attitudes: Continuation of a Dialog», en E. R. Tufte (ed.): *The Quantitative Analysis of Social Problems,* Reading, Mass, Addison-Wesley.

Cook, T. D. y Campbell D. T. (1979): *Quasi-experimentation: Design and Analysis Issues for Field Settings,* Boston, Houghton Mifflin.

Converse, J. M. y Presser, S. (1986): *Survey Questions: Handcrafting the Standardized Questionnaire,* Thousands Oaks, Sage, 1986.

Corbetta, P. y Parisi, A. M. L. (1987): «Il calo della partecipazione elettorale: disaffezione delle istituzioni o crisi dei riferimenti partitici?», en *Polis,* 1: 29-65.

Cordero, M. (1998): «Banco de datos», en *Cuadernos metodológicos del CIS,* n.º 24.

Corsaro, W. A. y Rizzo, T. A. (1990): *Disputes in the Peer Culture of American and Italian Nursery School Children,* en Grimshaw (1990).

Coser, L. (ed.) (1963): *Sociology Through Literature,* Englewood Cliffs, Prentice Hall.

Costrich, N. *et al.* (1975): «When Stereotypes Hurt: Three Studies of Penalties for Sex-role Reversals», en *Journal of Experimental Psychology,* vol. 11: 520-530.

Coulon, A. (1995): *Ethnomethodology,* Thousands Oaks, Sage.

Creswell, J. W. (1994): *Qualitative and Quantitative Approaches,* Thousands Oaks, Sage.

Dahl, R. A. (1961): *Who Governs?,* New Haven, Yale University Press.

Davis, F. (1973): «The Martian and the Covert: Ontological Polarities in Social Research», en *Urban Life,* 3: 333-343.

Del Piero, A. (1994), *L'isola della quarantina,* Florencia, Giunti.

Denzin, N. K. (1989): *The Research Act,* Englewood Cliffs, Prentice-Hall.

Denzin, N. K. y Lincoln, Y. S. (eds.) (1994): *Handbook of Qualitative Research,* Newbury Park, Sage; 2ª ed. (2000).

Denzin, N. K. (1997): *Interpretive Ethnography: Ethnographic Practices for the Twenty-first Century,* Thousand Oaks, Sage.

Desrosières, A. (1998): *The Politics of Large Numbers: A History of Statistical Reasoning,* Cambridge, Harvard University Press.

DeVellis, R. (1991): *Scale Development: Theory and Applications,* Newbury Park, Sage.

Díaz de Rada, V. (2005) *Manual de trabajo de campo en la encuesta,* Madrid, CIS.

Dillman, D. (1978): *Mail and Telephone Survey: The Total Design method,* New York, Wiley.

Dillman, D.A. (1991): «The Design and Administration of Mail Surveys», en *Annual Review of Sociology 1991,* 17: 225-49.

Dilthey, Wilhelm: *Introducción a las ciencias del espíritu,* traducción de J. Marías, Madrid, 1966.

Duncan, O. D. (1984): *Notes on Social Measurement: Historical and Critical,*

New Yoork, Russell Sage Foundation, 1984.
Durkheim, E. (1895): *Les règles de la méthode sociologique,* París, Alcan; las citas son de la traducción al español: *Las reglas del método sociológico y otros escritos sobre filosofía de las ciencias sociales* (1988), Madrid, Alianza Editorial.
Durkheim, E. (1897): *Le suicide: Étude de sociologie,* París, Alcan; las citas son de la traducción al español de *El suicidio* (1982), Akal Universitaria.
Edwards, A. L. (1957): *Techniques of Attitude Scale Construction,* Nueva York, Appleton.
Evans, G. E. (1975): *The Days that We Have Seen,* London, Faber and Faber.
Felicori, M. (1984): «Feste d'estate: indagine sulla politica culturale dei comuni italiani», en A. M. L. Parisi (ed.), *Luoghi e misure della politica,* Bolonia, Il Mulino.
Firebaugh, G. (1997): *Analyzing Repeated Surveys,* Thousand Oaks, Sage.
Fisher, R. A. (1935): *The Design of Experiments,* Edimburgo, Oliver y Boyd.
Foucault, M. (1975): *Surveiller et punir. Naissance de la prison,* París, Gallimard.
Fowler, F. J. Jr. (1995): *Improving Survey Questions: Design and Evaluation,* Thousand Oaks, Sage.
Friedrichs, R. W. (1970): *A Sociology of Sociology,* Nueva York, Free Press.
Gallino, L. (1978): *Dizionario di sociologia,* Turín, Utet.
Galtung, J. (1967): *Theory and Methods of Social Research,* Oslo, Universitetsforlaget.
Gambetta, D. (1987): *Where They Pushed or Did They Jump?,* Cambridge, Cambridge University Press.
Gambetta, D. (1992): *La mafia siciliana,* Turín, Einaudi.
Gans, H. (1962): *The Urban Villagers,* Nueva York, Free Press.
García Fernando, M., Ibáñez, J. y Alvica, F. (1986): *El análisis de la realidad social,* Madrid, Alianza Universidad Textos (1.ª ed.).
Geertz, C. (1973): *The Interpretation of Cultures,* New York, Basic Books; las citas son de la edición en lengua original. Hay traducción al español: *La interpretación de las culturas* (primera edición, 1981), Barcelona, Gedisa.
Genders, E. y Player, E. (1995): *Grendon: a Study of a Therapeutic Prison,* Oxford, Clarendon Press.
Gieryn, T. F. (2000): «A Space for Place in Sociology», en *Annual Review of Sociology, 2000,* 26: 463-96.
Glaser, B. G. y Strauss, A. L. (1976): *The Discovery of Grounded Theory: Strategies for Qualitative Research,* Nueva York, Aldine.
Glass, G. V. (1976): «Primary, Secondary and Meta-analysis of Research», en *Educational Researcher,* 5, n. 10: 3-8.
Glass, G. V., McGaw, B. y Smith, M. L. (1987): *Meta-analysis in Social Research,* Beverly Hills, Sage.
Glock, C. Y. (1965): «There is a Religious Revival in the United States?», en J. Zahn (ed.), *Religion and the Face of America,* Berkeley, University Extension, University of California.

Glueck, S. y Glueck, E. (1950): *Unraveling Juvenile Delinquency,* Nueva York, The Commonwealth Fund.
Goffman, E. (1959): *The Presentation of Self in Everyday Life,* Nueva York, Doubleday.
Goffman, E. (1967): *Interaction Ritual,* Nueva York, Doubleday.
Goldfarb, W. (1945): «Psychological Privation in Infancy and Subsequent Adjustment», en *American Journal of Orthopsychiatry,* 15: 247-253.
Goldthorpe, J. H. (2000): *On Sociology: Numbers, Narratives, and the Integration of Research and Theory,* Oxford, Oxford University Press.
Goode, W. J y Hatt, P. K. (1952): *Methods in Social Research,* Nueva York, McGraw-Hill.
Granovetter, M. (1995): *Getting a Job,* Chicago, The University of Chicago Press.
Grimshaw, A. D. (ed.) (1990): *Conflict Talk,* Cambridge, The University Press.
Groves, R. M. (1989): *Survey Errors and Survey Costs,* Nueva York, Wiley.
Groves, R. M. (1990): «Theories and Methods of Telephone Surveys», en *Annual Review of Sociology 1990,* 16: 221-40.
Groves, R. M. y Lyberg, L. E. (1988): *An Overview of Nonresponse Issues in Telephone Surveys,* en Groves *et al.* (1988).
Groves, R. M. et al. (1988): *Telephone Survey Methodology,* Nueva York, Wiley.
Guba, E. G. y Lincoln, Y. S. (1994): *Competing Paradigms in Qualitative Research,* en Denzin y Lincoln (1994).
Gubrium, J. F. y Holstein J. (1997): *The New Language of Qualitative Method,* New York, Oxford University Press.
Gubrium, J. F. y Holstein, J. A. (eds.) (2002): *Handbook of Interview Research: Context and Method,* Thousand Oaks, Sage.
Gurr, T. R. (1972): *Politimetrics,* Englewood Cliffs, Prentice Hall.
Guttman, L. A. (1944): «A Basis for Scaling Qualitative Data», en *American Sociological Review,* 9: 139-50.
Guttman, L. A. (1950) «The Basis for Scalogram Analysis», en S. Stouffer (ed.), *Measurement and Prediction,* Princeton, Princeton University Press.
Hall, G. (1976): «Workshop as a Ballerina: an Exercise in Professional Socialization», *Urban Life and Culture,* 6: 193-220.
Halsey, A. H. y J. Webb (eds.) (2000): *Twentieth-Century British Social Trends,* Houndmills, Palgrave Macmillan.
Hammersley, M. y Atkinson, P. (1983): *Ethnography: Principles in Practice,* Londres, Tavistock.
Hammersley, M. y Atkinson, P. (1994): *Ethnography and Participant Observation,* en Denzin y Lincoln.
Hanson, (1958): *Patterns of Discovery,* Cambridge, Cambridge University Press.
Hart, R. P. (2000): *Campaign Talk: Why Elections are Good for Us,* Princeton, Princeton University Press.
Henry, G. T. (1990): *Practical Sampling,* Newbury Park, Sage.

Heritage, J. (1984): *Garfinkel and Ethnomethodology,* Cambridge, Polity Press.
Holland, P. W. (1986): «Statistics and Causal Inference», en *Journal of the American Statistical Association,* 81: 945-960.
Horn, R. V. (1993): *Statistical Indicators for the Economic y Social Sciences,* Cambridge, Cambridge University Press.
Houtkoop-Steenstra, H. (2000): *Interaction and the Standardized Survey Interview: The Living Questionnaire,* Cambridge, Cambridge University Press, 2000.
Hughes, J. A. (1980): *The Philosophy of Social Research,* New York, Longman.
Hyman, H. (ed.) (1954): *Interviewing in Social Research,* Chicago, The University of Chicago Press.
Iceberg, S. (1967): *Pimp: the Story of My Life,* Los Angeles, Holloway.
Inglehart, R. (1977): *The Silent Revolution,* Princeton, Princeton University Press.
Irwin, J. (1970): *The Felon,* Englewood Cliffs, Prentice Hall.
Irwin, J. (1980): *Prisons in Turmoil,* Boston, Little, Brown y Co.
ISR - Institute for Social Research, (1976), *Interviewer's Manual,* Ann Arbor, University of Michigan.
Janowitz, M. (1958): «Inferences about Propaganda Impact from Textual and Documentary Analysis», en W. E. Daugherty y M. Janowitz (eds.), *A Psychological Warfare Casebook,* Baltimore, Johns Hopkins Press.
Jenkins, R. (1983): *Lads, Citizens and Ordinary Kids,* Londres, Routledge y Kegan.
Jorgensen, D. L. (1989): *Participant Observation,* Newbury Park, Sage.
Kahn, R. L. y Cannel, C. F. (1967): *The Dynamics of Interviewing,* New York, Wiley; nueva edición, *Dynamics of Interviewing: Theory, Technique and Cases,* Malabar, Krieger, 1983.
Kalton, G. (1983): *Introduction to Survey Sampling,* Beverly Hills, Sage.
Kaplan, A. (1964): *The Conduct of Inquiry,* New York, Harper and Row.
Kasprzyk, D., Duncan, G., Kalton, G. y Singh, M. P. (eds.) (1989): *Panel Surveys,* Nueva York, John Wiley.
Kellehear, A. (1993): *The Unobtrusive Researcher,* St. Leonards, Allen y Unwin.
Kendon, A. (1990): *Conducting Interaction,* Cambridge, Cambridge University Press.
Kertzer, D. (1980): *Comrades and Christian: Religion and Political Struggle in Communist Italy,* Cambridge, Cambridge University Press.
Kertzer, D. (1988): *Ritual, Politics and Power,* New Haven, Yale University Press.
Kiecolt, K. J. y Nathan, L. E. (1985): *Secondary Analysis of Survey Data,* Beverly Hills, Sage.
King, G., Keohane, R. O. y Verba, S. (1994): *Designing Social Inquiry,* Princeton, Princeton University Press. Las citas son de la versión española: *El diseño de la investigación social: la inferencia científica en los estudios cualitativos* (2000), Madrid, Alianza Editorial.
Kinsey, A. C., Pomeroy, W. B., Martin, C. E. y Grebhard, P. H. (1953): *Sexual Behavior in the Human Female,* Philadelphia, Saunders.

Kish, L. (1965): *Survey Sampling,* Nueva York, Wiley. Hay traducción al español: *Muestreo de encuestas* (1979).

Klecka, W. R. y Tuchfarber, A. J. (1978): «Random Digit Dialing: A Comparison to Personal Surveys», en *Public Opinion Quarterly,* 42: 105-14.

Knoke, D. y Kulinski, J. H. (1982): *Network Analysis,* Newbury Park, Sage.

Krebs D. y Schmidt P. (eds.) (1993): *New Directions in Attitude Measurement,* Nueva York and Berlín, Walter de Gruyter.

Kruskal, J. B. y Wish, M. (1978): *Multidimensional Scaling,* Beverly Hills, Sage.

Kuhn, T. S. (1962): *The Structure of Scientific Revolutions,* Chicago, The University of Chicago Press; las citas son de la edición en lengua original del año 1970. Hay traducción al español: *La estructura de las revoluciones científicas* (primera edición, 1971), Breviarios del Fondo de Cultura Económica.

Kutchinsky, B. (1973): «The Effect of Easy Availability of Pornography on the Incidence of Sex Crimes», en *Journal of Social Issues,* 29: 163-181.

Lasswell, H. D. y Leites, N. (eds.) (1949): *Language of Politics*: *Studies in Quantitative Semantics,* Nueva York, Stewart.

Lavrakas, P. J. (1993)*: Telephone Survey: Sampling, Selection and Supervision,* 2ª ed., Thousand Oaks, Sage.

Layder D. (1988): *Sociological Practice: Linking Theory and Social Research,* London, Sage.

Lazarsfeld, P. F. (1948): «The Use of Panels in Social Research», en *Proceedings of the American Philosophical Society,* 42: 405-410.

Lazarsfeld, P. F. (1959): «Problems in Methodology», en R. K. Merton, L. Broom, L. S. Cottrell (eds.), *Sociology Today,* Nueva York, Basic Books.

Lazarsfeld, P. F. y Oberschall, R. (1965): «Max Weber and Empirical Research», en *Americam Sociological Review,* 30: 185-199.

Lazarsfeld, P. F. y Rosenberg, M. (eds.) (1955): *The Language of Social Research,* Nueva York, The Free Press.

Likert, R. (1932): «A Technique for the Measurement of Attitudes», *Archives of Psychology,* n.º 140.

Livolsi, M., Schizzerotto, A., Porro, R. y Chiari, G. (1974): *La macchina del vuoto,* Bolonia, Il Mulino.

Lofland, J. (1971): *Analizing Social Settings,* Belmont, Wadsworth.

Lofland, J. y Lofland, L. H. (1995): *Analizing Social Settings,* Belmont, Wadsworth.

Long, Scott (1997): *Regression Models for Categorical and Limited Dependent Variables,* Advanced Quantitative Techniques in the Social Sciences, 7. Beverly Hills: Sage Publications.

Lord, F. M. y Novick, M. R. (1968): *Statistical Theories of Mental Scores,* Reading, Addison-Wesley.

Lynd, R. S. y Lynd, H. M. (1929): *Middletown,* Nueva York, Harcourt, Brace y World; citas en lengua original de la edición de 1957.

Lynd, R. S. y Lynd, H. M. (1937): *Middletown in Transition,* Nueva York, Harcourt, Brace y World.

Maas, S. y Kuypers, J. A. (1974): *From Thirty to Seventy: a 40 Year Longitudinal Study of Adult Life Styles and Personality,* San Franciso, Jossey Bass.
Madge, J. (1962): *The Origins of Scientific Sociology,* New York, The Free Press.
Malinowski, B. (1922): *Argonauts of the Western Pacific,* Londres, Routledge y Kegan Paul.
Mangione, T. W. (1995): *Mail Surveys: Improving the Quality,* Thousand Oaks, Sage.
Manheim, J. R. y Rich, R. C. (1988): *Análisis político empírico. Métodos de investigación en ciencia politica,* Madrid, Alianza Editorial.
Mann, L. (1969): «Queue Culture: the Waiting Line as a Social System», en *American Journal of Sociology,* 75: 340-354.
Mann, L. (1973): «Learning to Live with Lines», en J. Helmer y N. A. Eddington (eds.), *Urban Man: The Psychology of Urban Survival,* Nueva York, Free Press.
Maranell G.M. (ed.) (1974): *Scaling: A Sourcebook for Behavioral Scientists,* Chicago, Aldine.
Marradi, A. (1980): *Concetti e metodi per la ricerca sociale,* Florencia, Giuntina.
Marradi, A. (1981): «Misurazione e scale: qualche riflessione ed una proposta», en *Quaderni di sociologia,* 29: 595-639.
Marradi, A. (1996): «Una lunga ricerca sui valori ed alcuni suoi strumenti», en A. Marradi y G. P. Prandstraller (eds.), *L'etica dei ceti emergenti,* Milán, Angeli.
Marsh, C. (1982): *The Survey Method: The Contribution of Surveys to Sociological Explanation,* Londres, George Allen y Unwin.
Maynard, D. W. y Clayman, S. E. (1991): «The Diversity of Ethnomethodology», en *Annual Review of Sociology 1991,* 17: 385-418.
McCracken, G. (1988): *The Long Interview,* Newbury Park, Sage.
McIver, J. P.y Carmines, E. G. (1981): *Unidimensional Scaling,* Newbury Park, Sage.
Menard, S. (1991): *Longitudinal Research,* Newbury Park, Sage.
Merton, R. K. (1965): *On the Shoulders of Giants,* Nueva York, The Free Press.
Merton, R. K. (1968): *Social Theory and Social Structure,* 2ª ed., Nueva York, The Free Press.
Merton, R. K., Fiske, M. O. y Kendall, P. L. (1956): *The Focused Interview,* Nueva York, The Free Press.
Milgram, S. (1974): *Obedience to Authority: An Experimental View,* Nueva York, Harper y Row.
Mill, J. S. (1843): *A System of Logic Ratiocinative and Inductive,* Londres, Longmans; citas en lengua original de la edición de 1973, Londres, Routledge y Kegan, Paul.
Miller, R. L. (2000): *Researching Life Stories and Family Histories,* Thousand Oaks, Sage.
Mishler, E. G. (1986): *Research Interviewing: Context and Narrative,* Cambridge, Harvard University Press.

Moreno, J. (1970): *Barrios in Arms,* Pittsburgh, University of Pittsburgh Press.
Moreno, J. L. (1934): *Who Shall Survive?,* Washington D. C., Nervous and Mental Desease Publishing.
Moser, C. A. y Kalton, G. (1971): *Survey Methods in Social Investigation,* Londres, Heinemann.
Mowrer, E. R. (1927): *Family Disorganization,* Chicago, University of Chicago Press.
Newman I. y Benz C. R. (1988): *Qualitative-Quantitative Research Methodology: Exploring the Interactive Continuum,* Carbondale, Southern Illinois University Press.
Newson, J y Newson, E. (1978): *Seven Years Old in the Home Environment,* Harmondsworth, Penguin Books.
Nunnally, J. C. (1978): *Psychometric Theory,* Nueva York, McGraw Hill.
Orbuch T. L. (1997): «People's Accounts Count: The Sociology of Accounts», en *Annual Review of Sociology, 1997,* 23: 455-78.
Osgood, C. E. (1952): «The Nature and Measurement of Meaning», en *Psychological Bulletin,* 49: 197-237.
Osgood, C. E., Suci, G. J. y Tannenbaum, P. H. (1957): *The Measurement of Meaning,* Urbana, University of Illinois Press.
Palmer, V. M. (1928): *Field Studies in Sociology: a Student's Manual,* Chicago, University of Chicago Press.
Patton, M. Q. (1990): *Qualitative Evaluation and Research Methods,* Newbury Park, Sage, 1990.
Pawson, R. y Tilley, N. (1997): *Realistic Evaluation,* London, Sage.
Pedhazur, E. y Pedhazur Schmelkin L. (1991): *Measurement Design and Analysis: An Integrated Approach,* Hillsdale, Lawrence Erlbaum Associates.
Pheterson, G. I., Kiesler, S. B. y Goldberg, P. A. (1971): «Evaluation of the Performance of Women as a Function of their Sex, Achievement, and Personal History», en *Journal of Personality and Social Psychology,* 19: 114-118.
Plummer, K. (1983): *Documents of Life,* Boston, Allen y Unwin.
Polsby, N. W. (1963): *Community Power and Political Theory,* New Haven, Yale University Press.
Popper, K. R. (1934): *Logik der Forschung,* Viena; citas de la traducción al inglés *The Logic of Scientific Discovery,* London, Routledge, 1992. Existe traducción al castellano: *La lógica de la investigación científica* (primera edición, 1962), Madrid, Tecnos.
Prior, L. (1997): *Following in Foucault's Footsteps: Text and Context in Qualitative Research,* en Silverman.
Propp, V. (1928): *Morfologjia skazki,* Leningrad, Academia; *Morphology of the Folktale,* Austin, University of Texas Press, 1968. Hay traducción al español: *Morfología del cuento* (2000), Madrid, Editorial Fundamentos.
Rathje, W. y Murphy, C. (1992): *Rubbish! The Archaeology of Garbage,* Nueva York, Harper y Collins.

Reher, D. y Valero, A. (1995): «Fuentes de información demográfica en España», en *Cuadernos Metodológicos del CIS,* n.º 13, Madrid, CIS.
Reichenbach, H. (1951): *The Rise of Scientific Philosophy,* Berkeley, University of California Press.
Ricci, A. y Salierno, G. (1971): *Il carcere in Italia,* Turín, Einaudi.
Riessman, C. K. (1990): *Divorce Talk,* New Brunswick, Rutgers University Press.
Ritzer G. y Smart B. (2001): *Handbook of Social Theory,* Londres, Sage.
Robertshaw, P. (1985): «Semantic and Linguistic Aspect of Sex Discrimination Decision. Dichotomised Woman», en D. Carzo y B. S. Jackson (eds.), *Semiotics, Law and Social Science,* Roma, Gangemi.
Robinson, J. P., Shaver, P. R. y Wrightsman, L. S. (1991), *Measures of Personality and Social Psychologial Attitudes,* San Diego, Academic Press.
Robinson, J. P., Shaver, P. R. y Wrightsman, L. S. (1999): *Measures of Political Attitudes,* San Diego, Academic Press.
Rodríguez Osuna, J. (1991): «Métodos de muestreo», en *Cuadernos Metodológicos del CIS,* n.º 1, Madrid, CIS.
Rodríguez Osuna, J. (1993): «Métodos de muestreo», en *Cuadernos Metodológicos del CIS,* n.º 6, Madrid, CIS.
Romero, M. (1992): *Maid in the USA,* Nueva York, Routledge.
Rosenberg, M. (1968): *The Logic of Survey Analysis,* Nueva York, Basic Books.
Rosenthal, R. (1966): *Experimenter Effects in Behavioral Research,* Nueva York, Appleton Century Crofts.
Rosenthal, R. y Jacobson, L. (1968): «Teacher Expectations for the Disadvantaged», en *Scientific American:* 19-23.
Rossi, P. H., Wright, G. D. y Anderson, A. B. (1983): *Handbook of Survey Research,* San Diego, Academic Press.
Rotschild-Witt, J. (1979): «The Collectivistic Organisation: an Alternative to Rational-Bureaucratic Models», en *American Sociological Review,* 44: 509-527.
Rotton, J. y Kelly, I. W. (1985): «Much Ado About the Full Moon: A Meta-Analysis of Lunar-Lunacy Research», en *Psychological Bulletin,* 97: 286-306.
Roy, D. (1952) «Quota Restriction and Goldbricking in a Machine Shop», *The American Journal of Sociology,* 57: 427-442.
Rubin, D. B. (1974): «Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies», en *Journal of Educational Psychology,* 66: 688-701.
Rugg, D. (1941): «Experiments in Wording Questions: II», en *Public Opinion Quarterly,* 5: 91-92
Runciman, W. G. (1966): *Relative Deprivation and Social Justice,* Londres, Routledge y Kegan.
Salzinger, K. (1958): «A Method of Analysis of the Process of Verbal Communication between a Group of Emotionally Disturbed Adolescents, and their Friends and Relatives», en *The Journal of Social Psychology,* 47: 39-53.
Sampson, R. J y Laub, J. H. (1993): *Crime in the Making,* Cambridge, Harvard

University Press.
Sánchez-Jankowski, M. (1991): *Islands in the Street,* Berkeley, University of California Press.
Schuman, H. y Presser, S. (1981): *Questions and Answers in Attitude Surveys,* Nueva York, Academic Press.
Schutz, A. (1932): *Der Dinnhafte Aufbander sozialen Welt; eine Einleitung in die Verstehende Soziologie,* Wien, Springer; citas de la versión inglesa *The Phenomenology of Social World,* Evanston, Northwestern University Press, 1967.
Schwartz, H. y Jacobs, J. (1979): *Qualitative Sociology,* Nueva York, The Free Press.
Sciolla, L. y Ricolfi, L. (1989): *Vent'anni dopo,* Bolonia, Il Mulino.
Scott, J. (1990): *A Matter of Record: Documentary Sources in Social Research,* Cambridge, Polity Press.
Scott, M. (1968): *The Racing Game,* Chicago, Aldine.
Selltiz C., Wrightsman, L. S. y Cook, S. W. (1976): *Research Methods in Social Relations,* Nueva York, Holt, Rinehart and Winston.
Shaw, B. (1941): *The Doctor's Dilemma,* Nueva York, Dodd, Mead y Co., 1941 (1.ª ed 1911).
Sheatsely, P. B. (1983): *Questionnaire Construction and Item Writing,* en Rossi *et al.* (1983).
Sherif, M. (1967): *Group Conflict and Co-operation: Their Social Psychology,* Londres, Routledge y Kegan.
Silverman, D. (ed.) (1997): *Qualitative Research,* London, Sage.
Smith, T. W. (1984): «Nonattitudes: A Review and Evaluation», en C. F. Turner y E. Martin (eds.), *Surveying Subjective Phenomena,* Nueva York, Russel Sage Foundation.
Sniderman, P. M. y Grob, D. B. (1996): «Innovations in Experimental Design in Attitude Surveys», en *Annual Review of Sociology 1996,* 22: 377-99.
Spradley, J. P. (1979): *The Ethnographic Interview,* Nueva York, Holt, Rinheart y Winston.
Spradley, J. P. (1980): *Participant Observation,* Nueva York, Holt, Rinheart y Winston.
Steffensmeier D. J. (1986): *The Fence,* Totowa, Rowman y Littlefield.
Stevens, S. S. (1946): «On the Theory of Scales of Measurement», en *Science,* 103: 677-80.
Stewart, D. W. y Shamdasani, P. N. (1990): *Focus Groups: Theory and Practice,* Newbury Park, Sage.
Stinchcombe, A. L. (1968): *Constructing Social Theories,* Nueva York, Harcourt, Brace y World.
Strauss, A. L. y Corbin, J. (1990): *Basics for Qualitative Research: Grounded Theory Procedures and Techniques,* Newbury Park, Sage.
Sudman, S. y Bradburn, N. R. (1982): *Asking Questions,* San Francisco, Jossey-Bass.
Sudman, S., Bradburn, N. M. y Schwarz, N. (1996): *Thinking About Answers:*

The Application of Cognitive Processes to Survey Methodology, San Francisco, Jossey-Bass.
Sussman, L. (1963): *Dear F. D. R.,* Nueva York, Bedminster.
Sutherland, E. H. (1937): *The Professional Thief,* Chicago, University of Chicago Press.
Tashakkori A. y Teddlie C. (1998): *Mixed methodology: Combining Qualitative and Quantitative Approaches,* Thousand Oaks, Sage.
Thomas, W. I. y Znaniecki, F. (1918-20): *The Polish Peasant in Europe and America,* Chicago, The University of Chicago Press.
Thomas, W. I. y Znaniecki, F. (1919): *The Polish Peasant in Europe and America, vol. III, Life record of an Immigrant,* Boston, The Gorham Press, 1919
Thompson, P. (1972): «Problems of Method in Oral History», en *Oral History,* 4, n. 2: 1-47.
Thompson, P. (1975): *The Edwardians. The Remaking of British Society,* London, Weidenfeld.
Thurstone, L. L. (1927): «A Law of Comparative Judgement», en *Psychological Review,* 34: 273-86.
Thurstone, L. L. (1928): «Attitudes Can Be Measured», en *American Journal of Sociology,* 33: 529-54.
Thurstone, L. L. (1931): «Rank Order as a Psychophysical Method», en *Journal of Experimental Psychology,* 6: 187-201.
Thurstone, L. L y Chave, E. J. (1929): *The Measurement of Attitude,* Chicago, The University of Chicago Press.
Tocqueville, A. de (1856): *L'Ancien régime et la révolution;* las citas son de la versión española: *El Antiguo Régimen y la Revolución,* Madrid, Alianza Editorial, 1982.
Tulis, J. (1987): *The Rethorical Presidency,* Princeton, Princeton University Press.
Van Maanen, J. (1988): *Tales of the Field,* Chicago, The University of Chicago Press.
Van Maanen, J., Dabbs, J. M. y Faulkner, R. R. (1982): *Varieties of Qualitative Research,* Beverly Hills, Sage.
Vuchinich, S. (1990): *The Sequential Organization of Closing in Verbal Family Conflict,* Grimshaw (1990).
Wagenaar, A. C. (1981): «Effect of the Raised Legal Drinking Age on Motor Vehicle Accidents in Michigan», en *HSRI Research Review,* 11: 1-8.
Wallerstein, J. S. y Kelly, J. B. (1980): *Surviving the Breakup,* Nueva York, Basic Books.
Walum, L. R. (1974): «The Changing Door Ceremony», en *Urban Life and Culture,* 2: 506-515.
Wasserman, S. y Faust, K. (1994): *Social Network Analysis,* Cambridge, Cambridge University Press.
Webb, E., Campbell, D. T., Schwartz, D. y Sechrest, L. (1966): *Unobtrusive Measures,* Chicago, Rand Mcnally; 2ª ed. *Nonreactive measures in the Social Sciences,* Boston, Houghton y Mifflin, 1981.

Weber, M. (1904): *Über die «Objectivität' sozial wissenshaftlicher und sozialpolitsher Erkenntins* reeditado en Weber (1922b).
Weber, M. (1922a): *Wirtschaft und Gesellshaft,* Tübingen, Mohr.
Weber, M. (1922b): *Gesammelte Aufsätze zur Wissenschaftslehre,* Tübingen, Mohr.
Weber, M. (1924): *Gesammelte Aufsätze zur Soziologie un Sozialpolitik,* Tübingen, Mohr.
Weisberg, H. F y Rusk, J. G. (1970) «Dimensions of Candidate Evaluation», en *American Political Science Review,* 64: 1167-1185.
Weiss, J. y Gruber, G. (1986): «The Managed Irrelevance of Federal Education Statistics» en W. Alonso y P. Starr (eds.), *The Politics of Numbers,* Nueva York, Russell Sage Foundation.
Whyte, W. F. (1943): *Street Corner Society. The Social Structure of an Italian Slum,* Chicago, The University of Chicago Press.
Whyte, W. F. (1984): *Learning from the Field,* Beverly Hills, Sage.
Winship, C. y Morgan, S. (1999): «The Estimation of Causal Effects from Observational Data», en *Annual Review of Sociology 1999,* 25: 659-707.
Wiseman, J. (1979): «Close Encounters on the Quasi-Primary Kind: Sociability in Urban Second-Half Clothing Stores», en *Urban Life,* 8: 23-51.
Wolfinger, R. (1973): *The Politics of Progress,* Englewood Cliffs, Prentice Hall.
Woods, P. (1979): *The Divided School,* Londres, Routledge y Kegan.
Zimmerman, D. H. y Wieder, D. L. (1977): «The Diary-interview Method», en *Urban Life,* 5: 479-97.

ÍNDICE ANALÍTICO

Confianza
  grado de, 294
  intervalo de, 294
Control
  de la comunidad científica, 72
  de la teoría, 77
  de laboratorio, 132
  de las variables externas, 132
  estadístico (depuración), 123
Constante, 88
Convariación, 119-120
Cuestión
  de la equivalencia, 129
  de la invariación, 128
Cuestionario, 160-161, 368
  autocumplimentado, 198-201
  postal, 199

Datos, 74
  agregados, 271
Deducción, 73
Definición operativa, 83
Depuración (control estadístico), 123
Descripción, 350, 351
  y explicación, 317, 427
Deseabilidad social, 166, 181
Descubrimiento (contexto del), 72
Desviación típica, 294
Diarios, 199, 403, 405
Diferencial semántico, 248
Diseño
  de la investigación, 74
  factorial, 145, 307-308
Documentos, 400, 401
  institucionales, 413-414

personales, 401-406

Efecto
contaminación, 187
del experimentador (Hawthorne), 80, 81, 152
Elemento, 224, 225
análisis de los, 234
Emparejamiento, 138, 151
Empirismo lógico, 16
Encuesta
social, 158-162
longitudinal/transversal, 209-211, 212
retrospectiva, 210
transversal repetida, 211, 212
Entrevista, 369-374
cara a cara, 192-195
clínica, 382
de grupo, 383, 384, 385
estructurada, 160, 374-376
libre, 160
no dirigida, 379
no estructurada, 377-378
semiestructurada, 376
telefónica, 195-198
telemática, 201-202
Entrevistador, 193, 194, 195, 388, 389
Enunciados de posibilidad, 23
Epistemología, 8
Equivalencia, 106
Error
accidental y sistemático, 103-104, 107
de cobertura, 104, 105, 292, 309-311, 316
de muestreo, 105, 292-296, 299-300
de no respuesta,105, 292, 311-312, 316
de observación, 104-105
de selección, 104-105
de tratamiento de datos, 104-105
global, 106
reducción proporcional del, 589, 590
Escalas
aditivas, 233
por intervalos, 95
de autoposicionamiento, 229
de Guttman (escalograma), 242-248
de Likert, 233-242
de Rasch, 248
Estado (de una propiedad), 83, 91
Estandarización
de las preguntas y las respuestas, 160 171
Estilo de escritura, 354
Estimación, 294
Estudio
descriptivo/explicativo, 317
diacrónico, 209
de comunidad, 331
explorativo, 175, 203
longitudinal/transversal, 88, 209
Experimento, 124, 125-132
cuasiexperimentado, 141, 146-151
de laboratorio, 132-126
en el campo, 136-138
ex postfacto, 151
verdaderos experimentos, 141-146
Explicación e interpretación, 20
Etnografía, 328, 334
Etnometodología, 362, 395

Fiabilidad, 107-108
Focus group, 384
Fuentes estadísticas, 264-287
Formas equivalentes, 131

Grupo experimental/de control, 131

Hecho social, 12
Hipótesis, 76
Historia
de vida, 409-411
oral, 412
Holística (perspectiva), 54
Huellas físicas, 424-426

Indeterminación, principio de, 98
Indicación, relación de, 98
Indicador, 98-102
Índice, 100
Inducción, 13, 75
Inferencial causal, 127

Panel, 167, 210, 217
Paradigma, 4-7
Población, 316
Ponderación, 303, 313, 316
Pos-estratificación, 314
Positivismo, 9-13
  lógico, 15
  neopositivismo, 13-18
  postpositivismo, 13-15
Preguntas
  abiertas y cerradas, 159, 167, 171-175
  batería de, 189-192
  codificadas en el campo, 173
  de hechos y motivaciones, 170
  de prestigio, 166
  de respuesta múltiple, 484
   embudo, 187
   filtro, 179
   formulación de las, 175-189
   primarias y secundarias, 386
Pretest, postests, 142, 203-204
Propiedad, 82
  discreta y continua, 92, 95
  operacionalizada, 87, 88
  sociodemográfica, 170

Reduccionismo, 25
Registro dela observación, 346-349
Representatividad de la muestra, 152, 153
Reproducibilidad, 106, 243
Respuesta normativa (deseabilidad social de las respuestas), 166
Revolución científica, 4, 5
Seudoopiniones, 166
Sociología
  comprensiva, 20-22
Sociometría, 253-258
Sondeo (entrevista), 159-160
Subcultura, 332

Temas culturales, 353, 354
Teoría, 75
  y generalización empírica, 77
  e hipótesis, 76
  e investigación, 72-75
Testimonios orales, 406-412

Test-retest, 107
Tipo ideal, 22, 353
Tipología, 352, 392

Unidad
  de análisis, 84-87
  de medida, 95
  de recuento, 96
  de recogida, 86
Universo, 292

Valores, 83
  asignación de, 93
  diferencias entre, 94
  observados y verdaderos, 103
Variables
  agregadas, 89, 270
  constantes, 88
  contextuales, 89
  cuasi-cardinales, 96, 97, 226, 230
  dicotomómicas, 93
  de intervalos y ratios, 95
  dependientes e independientes, 53, 119, 120
  discretas y continuas, 91, 95
  fundamentales y derivadas, 96
  globales, 89
  individuales y colectivas, 89
  latentes y observadas, 89
  lenguaje de las, 16, 89
  manipulables, 89, 120
  nominales, ordinales y cardinales, 91-97, 226-230

Verstehen, 21, 22

Z (estándar), *ver* estandarización