

Gary King, Robert O. Keohane
y Sidney Verba

El diseño de la investigación social

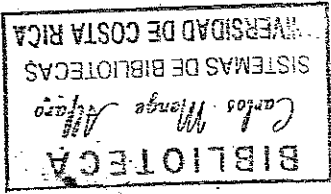
La inferencia científica
en los estudios cualitativos

Version de Jesús Cuéllar Menezo

Alianza Editorial

300.72
K52d

Reservados todos los derechos. El contenido de esta obra está protegido por la Ley, que establece penas de prisión y/o multas, además de las correspondientes indemnizaciones por daños y perjuicios, para quienes reprodujeren, plagiaren, distribuyeren o comunicaren públicamente, en todo o en parte, una obra literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier medio, sin la preceptiva autorización.



SISTEMA DE BIBLIOTECAS - UCR



579277

© de la traducción: Jesús Cuellar Menozo, 2000

© 1994 by Princeton University Press

© Ed. cast.: Alianza Editorial, S.A., Madrid, 2000, 2005, 2007

Calle Juan Ignacio Luca de Tena, 15; 28027 Madrid; telef. 91 393 88 88

www.alianzaeditorial.es

ISBN: 978-84-206-8697-4

Depósito legal: M. 35.519-2007

Fotocomposición: EFCO, S. A.

Poligono «Las Muelas»; 28850 Torrejón de Ardoz (Madrid)

Printed in Spain

SI QUIERE RECIBIR INFORMACIÓN PERIÓDICA SOBRE LAS NOVEDADES DE ALIANZA EDITORIAL, ENVÍE UN CORREO ELECTRÓNICO A LA DIRECCIÓN:

alianzaeditorial@anaya.es

Prefacio..... 9

1. La ciencia en las ciencias sociales..... 13

1.1 Introducción..... 13

1.1.1 Dos formas de investigar, una sola lógica inferencial..... 13

1.1.2 Definición de investigación científica en las ciencias sociales..... 17

1.3 Ciencia y complejidad..... 20

2. Principales componentes del diseño de investigación..... 23

2.1 Mejorar las preguntas de la investigación..... 24

2.2 Mejorar la teoría..... 29

2.3 Mejorar la calidad de los datos..... 33

2.4 Utilizar mejor los datos existentes..... 37

3. Temas de estudio de este volumen..... 39

3.1 La utilización de consecuencias observables para vincular la teoría y los datos..... 39

3.2 La maximización del control..... 39

3.3 Admitir la incertidumbre..... 42

3.4 Pensar como un científico social: escepticismo e hipótesis contrapuestas..... 42

2. La inferencia descriptiva..... 45

1. El conocimiento general y los hechos particulares..... 46

1.1 «Interpretación» e inferencia..... 48

1.2 «Singlularidad», complejidad y simplificación..... 53

1.3 Estudios de caso comparados..... 55

Signature line with 'SIBDI' text

1. La inferencia: el fin científico de la recogida de datos 57

2. La inferencia: el fin científico de la recogida de datos 60

3. Modelos formales de investigación cualitativa 62

4. Un modelo formal sobre la recogida de datos 64

5. El resumen de los portmenores históricos 66

6. La inferencia descriptiva 74

7. Criterios para juzgar las inferencias descriptivas 74

7.1 Inferencias no sesgadas 76

7.2 Eficiencia 76

3. Causalidad e inferencia causal 87

1. Definición de causalidad 88

1.1 Definición y un ejemplo cuantitativo 94

1.2 Un ejemplo cualitativo 96

2. Clantificación de las definiciones alternativas de causalidad 96

2.1 «Mecanismos causales» 98

2.2 «Causalidad múltiple» 100

2.3 Causalidad «simétrica» y «asimétrica» 102

3. Supuestos necesarios para la estimación de efectos causales 102

3.1 Homogeneidad de las unidades 105

3.2 Independencia condicional 106

4. Criterios para evaluar las inferencias causales 107

5. Reglas para elaborar teorías causales 108

5.1 Regla 1: elaborar teorías falsables 116

5.2 Regla 2: construir teorías que tengan coherencia interna 118

5.3 Regla 3: seleccionar cuidadosamente las variables dependientes 120

5.4 Regla 4: maximizar lo concreto 120

5.5 Regla 5: formular teorías de forma tan incluyente como sea posible 123

4. Precisar qué va a observarse 125

1. Diseños de investigación imprecisos 128

1.1 Más inferencias que observaciones 129

1.2 Multicolinealidad 131

2. Límites de la selección aleatoria 134

3. Sesgo de selección 138

3.1 Seleccionar en función de la variable dependiente 139

3.1.1 Ejemplos de sesgo de selección producido por el investigador 142

3.1.2 Ejemplos de sesgo de selección producido por el mundo 145

3.2 Seleccionar en función de una variable explicativa 147

3.3 Otros tipos de sesgo de selección 149

4. Selección intencionada de observaciones 150

4.1 Selección en función de la variable explicativa 150

4.2 Seleccionar un abanico de valores de la variable dependiente 151

4.3 Seleccionar observaciones utilizando tanto las variables explicativas como las dependientes 153

4.4 Seleccionar observaciones de manera que la variable causal clave sea constante 157

4.5 Seleccionar observaciones de manera que la variable dependiente sea constante 158

5. Comentarios Finales 159

5. Entender qué debe evitarse 161

1. Error de medida 162

1.1 Error de medida sistemático 167

1.2 Error de medida no sistemático 169

1.2.1 Error de medida no sistemático en la variable dependiente 170

1.2.2 Error de medida no sistemático en una variable explicativa 173

2. Exclusión de variables relevantes: el sesgo 178

2.1 Cómo calibrar el sesgo de las variables omitidas 178

2.2 Ejemplos de sesgo de la variable omitida 186

3. Inclusión de variables irrelevantes: la ineficiencia 193

4. Endogeneidad 199

4.1 Corregir inferencias sesgadas 200

4.2 Precisar cuáles son los componentes de la variable dependiente 201

4.3 Convertir la endogeneidad en un problema de variable omitida 202

4.4 Seleccionar observaciones para evitar la endogeneidad 205

4.5 Precisar cuáles son los componentes de la variable explicativa 207

5. Asignación de valores a la variable explicativa 211

6. Controlar la situación que se investiga 218

7. Observaciones finales 218

6. Aumentar el número de observaciones 221

1. Inferencias causales con diseños de una sola observación 222

1.1. Estudios de caso «cruciales» 222

1.2. Razonamiento analógico 225

2. ¿Cuántas observaciones son suficientes? 226

3. Aumentar el número de observaciones a partir de unas pocas 230

3.1 Las mismas medidas con otras unidades 232

3.2 Nuevas medidas con las mismas unidades 236

3.3 Nuevas medidas con nuevas unidades 237

4. Observaciones finales 241

Notas 243

Bibliografía 255

Índice analítico 265

En este libro se propone un mismo enfoque para la extracción de inferencias válidas, tanto causales como descriptivas, en la investigación cualitativa, donde las medidas numéricas resultan imposibles o no recomendables. Lo que queremos señalar es que la lógica de un buen proyecto de investigación no difiere fundamentalmente por que éste sea cuantitativo o cualitativo. Nuestro enfoque se aplica por igual a estas dos formas de estudio separado.

Con este libro pretendemos animar a los investigadores cualitativos a

que se tomen en serio la inferencia científica y la incorporen a su trabajo. Confiamos en que nuestra lógica inferencial única y el intento de demostrar que puede ser útil para los cualitativistas ayuden a mejorar los trabajos en nuestra disciplina y, quizá también, los de otros estudios sociales. Por lo tanto, esperamos que este trabajo lo lean y examinen de forma crítica política y científica. Utilizamos algunas fórmulas hasta académicas ya asentadas. Utilizamos algunas fórmulas matemáticas porque resultan especialmente útiles para clarificar ciertos conceptos de los métodos cualitativos; sin embargo, no presuponemos en el lector conocimientos matemáticos o estadísticos, y es posible saltarse la mayoría de las fórmulas sin perder el hilo argumental.

Las autoridades universitarias suelen decir que la docencia y la investigación son complementarias. De hecho, ambas actividades son prácticas-

mente coincidentes porque las dos entrañan la adquisición de nuevos conocimientos y su transmisión a otras personas, aunque de una manera ligera-mente diferente en uno y otro caso. Este libro atestigua el carácter sincrónico de ambas actividades. Desde 1989, hemos estado trabajando en este libro y dirigiendo juntos en el departamento de gobierno de la Universidad de Harvard un seminario de tercer ciclo titulado «Métodos cualitativos en las ciencias sociales». El seminario ha sido muy animado y con frecuencia ha continuado en los pasillos y en los largos mensajes que hemos intercambiado entre nosotros y con nuestros estudiantes. Las batallas intelectuales han sido siempre amistosas, pero las reglas del combate dictaban que nosotro no estabamos verdaderamente convencido de algo, creiamos que nuestra obligación era continuar el debate. Al final, los autores aprendimos bastante unos de otros y de los estudiantes acerca de los métodos de investigación cualitativos y cuantitativos, y cambiamos muchos de nuestros puntos de vista iniciales. Al margen de sus objetivos principales, este libro constituye una declaración del enfoque unánime de la inferencia científica en la investigación cualitativa al que tanto nos costó llegar.

La primera versión del libro se terminó en 1991, y en los años posteriores la hemos revisado considerablemente. Gary King fue el primero en sugerir que lo escribiríamos, esbozó las primeras versiones de la mayoría de los capítulos y llevó la voz cantante en el largo proceso de revisión. Sin embargo, la redacción del libro ha sido alterada en tan gran medida por Robert Keohane y Sidney Verba, así como por el propio Gary King, que nos resulta imposible determinar de manera fiable la autoría de muchos fragmentos y apartados.

Durante este largo proceso pasamos borradores a compañeros de todos los Estados Unidos, a los que estamos muy agradecidos por la extraordinaria generosidad de sus comentarios. Damos también las gracias a los estudiantes de doctorado, tanto de Harvard como de otras universidades, que han ido conociendo el texto y cuya reacción ha sido importante a la hora de revisarlo. Es muy aventurado pretender mencionar a todos aquellos que han colaborado en un proyecto como este (calculamos que existe una posibilidad de 0,92 de que sin darnos cuenta omitamos a alguien que nos haya hecho comentarios importantes). Sin embargo, quisieramos mencionar a las siguientes personas: Christopher H. Achen, John Aldrich, Hayward Alker, Robert H. Bates, James Battista, Nathaniel Beck, Nancy Burns, Michael Cobb, David Collier, Gary Cox, Michael C. Desch, David Dessler, Jorge Domínguez, George Downs, Mitchell Duneter, Matthew Evangelista, John Ferjold, Andrew Gelman, Alexander George, Joshua Goldstein, Andrew Green, David Green, Robin Hanna, Michael Hiscox, James E. Jones, padre, Miles Kahler, Elizabeth King, Alexander Kozhemakim, Stephen D. Krasner, Herbert Kritzer, James Kuklinski, Nathan Lane, Peter Lange, Tony Lavelle, Judy Layzer, Jack S. Levy, Daniel Little, Sean Lynn-Jones, Lisa L.

Martin, Helen Miller, Gerardo L. Munck, Timothy R. Nokken, Joseph S. Nye, Charles Ragin, Swarna Rajagopalan, Shamara Shantu Riley, David Rooke, David Rohde, Frances Rosenbuth, David Schweder, Collins G. Shackelford, padre, Kenneth Shepsle, Daniel Walsh, Carolyn Warner, Steve Aviv Yetiv, Mary Zerbino y Michael Zürn. Vaya nuestro agradecimiento a Steve Voss por confeccionar el índice analítico y al personal de Princeton University Press: Walter Lippincott, Malcolm DeBevoise, Peter Dougherty y Alessandra Bocco. También damos las gracias a la National Science Foundation de los Estados Unidos por la ayuda a la investigación SBR-9223637 que concedió a Gary King. Robert O. Keohane agradece a la John Simon Guggenheim Memorial Foundation el periodo de beca que le permitió concluir el trabajo en este libro.

Los tres autores (ordenados y combinados de diversas maneras) tuvimos la inmensa fortuna de poder presentar versiones anteriores de este libro en seminarios y grupos de trabajo de las reuniones de la Midwest Political Science Association (Chicago, 2-6 de abril de 1990); del Grupo sobre Metodología Política (Universidad de Duke, 18-20 de julio de 1990); de la American Political Science Association (Washington, D.C., del 29 de agosto al 1 de septiembre de 1991); del Seminario sobre metodología y filosofía de las ciencias sociales (Universidad de Harvard, Center for International Affairs, 25 de septiembre de 1992); de la serie de coloquios del Interdisciplinary Consortium for Statistical Applications (Universidad de Indiana, 4 de diciembre de 1991); del ciclo de seminarios del Institute for Global Cooperation and Change (Universidad de California, Berkeley, 15 de enero de 1993); así como en la Universidad de Illinois, Urbana-Champaign (18 de marzo de 1993).

GARY KING
ROBERT O. KEOHANE
SIDNEY VERBA
Cambridge, Massachusetts

1. La ciencia en las ciencias sociales

1. Introducción

Este libro trata de la investigación en las ciencias sociales y tiene un objetivo práctico: diseñar investigaciones que generen inferencias válidas sobre la vida social y política. Nos centramos en la ciencia política, pero nuestro argumento es aplicable a disciplinas como la sociología, la antropología, la historia, la economía y la psicología, así como a áreas que no se consideran una disciplina, como las técnicas periciales, la investigación en temas de educación y el razonamiento clínico.

Nuestra obra no se ocupa de la filosofía de las ciencias sociales, ni tanto poco es un manual para tareas específicas de la investigación como el diseño de encuestas, la realización del trabajo de campo o el análisis de datos estadísticos. En realidad, trata del diseño de la investigación: de cómo plantear preguntas y moldear los estudios académicos para extraer inferencias desprengidas y causales válidas. Por lo tanto, se sitúa en un zona intermedia entre las abstractas polémicas filosóficas y las técnicas prácticas para centrarse en la lógica intrínseca que subyace en toda investigación social científica.

1.1 Dos formas de investigar, una sola lógica inferencial

Nuestro principal objetivo es relacionar las tradiciones de lo que se suele denominar investigación «cuantitativa» y «cualitativa» mediante una misma lógica inferencial. Ambas tendencias parecen bastante diferentes y, de he-

proceso de investigación

mente a esta idea. Esta diferencia de opiniones genera una viva polémica, pero, por desgracia, también divide la disciplina en dos ramas, una cuantitativa-sistemática-generalizadora y otra cualitativa-humanista-discursiva. A medida que la primera aumenta la complejidad de sus análisis estadísticos (y su trabajo se hace menos comprensible para los que no han estudiado las técnicas), la segunda se convence cada vez más de que tales análisis resultan irrelevantes para acontecimientos que, como aquellos que interesan a sus partidarios, no parecen poder reproducirse o generalizarse.

Uno de los principales propósitos de este libro es demostrar que las diferencias entre la tradición cuantitativa y la cualitativa son sólo de tipo estilístico y que su importancia en términos metodológicos y de contenido es mínima. Puede considerarse — con razón — que toda buena investigación procede de la misma lógica inferencial subyacente. Tanto los estudios cuantitativos como los cualitativos pueden ser sistemáticos y científicos. La investigación histórica puede ser analítica si lo que pretende es evaluar diversas explicaciones mediante un proceso inferencial causal válido. La historia o la sociología histórica, no es incompatible con las ciencias sociales

Dif. son de métodos est. sociales

Para romper estas barreras es preciso que comencemos por cuestionar el concepto mismo de investigación «cualitativa». Hemos utilizado el término en el título de este libro para hacer referencia a nuestro objeto de estudio, pero sin querer decir con ello que la investigación «cualitativa» se diferencia fundamentalmente de la «cuantitativa», excepto en el estilo.

La mayoría de las investigaciones no corresponden claramente a una u otra categoría, y las mejores suelen combinar características de los dos tipos. En el mismo proyecto de investigación, algunos de los datos recogidos pueden ser objeto de análisis estadístico, pero no así otros elementos igualmente relevantes. Las pautas y tendencias del comportamiento social, político o económico son más susceptibles de análisis cuantitativo que el flujo de ideas entre personas o la influencia determinante que tiene un líder extraordinario. Para entender un mundo social en cambio constante tenemos que barajar información que no es fácilmente cuantificable y también la que sí lo es. Además, todas las ciencias sociales necesitan compararse, lo cual supone que se evalúe qué fenómenos se parecen «más» o «menos» en cuestión de grado (o sea, diferencias cuantitativas) o de tipo (diferencias cualitativas).

combinar ambos

Dos magníficos estudios recientes son ejemplo de este punto de vista. En *Coercive Cooperation* (1992), Lisa L. Martin quiso explicar el grado de cooperación internacional en materia de sanciones económicas, analizando cuantitativamente 99 casos en los que se intentó aplicar tales medidas después de la Segunda Guerra Mundial. Aunque este análisis cuantitativo ofrecía mucha información valiosa, algunas de las inferencias causales que indicaban los datos eran ambigüas; de ahí que Martin decidiera estudiar en detalle seis casos de sanciones con la intención de recabar más pruebas de relevancia para su inferencia causal. Para escribir *Making Democracy Work*

cho, a veces se consideraran enfrentadas. Para nosotros, tales diferencias sólo son una cuestión de estilo y de técnicas específicas. Los dos tipos de investigación participan de una misma lógica subyacente, que suele explicarse y formalizarse claramente al analizar métodos cuantitativos. Sin embargo, esta lógica inferencial es igual a la que respalda las buenas investigaciones cualitativas, y tanto los cualitativos como los cuantitativos se beneficiarían si se le concediera una atención más explícita al diseñar investigaciones.

Los estilos de la investigación cuantitativa y cualitativa son muy diferentes. La primera se sirve de números y métodos estadísticos. Suele basarse en medidas numéricas de ciertos aspectos de los fenómenos; parte de casos concretos para llegar a una descripción general o para comprobar hipótesis causales y busca medidas y análisis que otros investigadores puedan reproducir fácilmente.

Por el contrario, la investigación cualitativa abarca una amplia gama de enfoques, pero, por definición, ninguno de ellos se basa en medidas numéricas. Este tipo de trabajo se centra generalmente en un caso o en un reducido número de ellos; se sirve de entrevistas en profundidad o de análisis de estudios de forma global o exhaustiva un acontecimiento o unidad. Los cualitativos, aunque sólo tengan un pequeño número de casos, suelen sacar a la luz en sus estudios una gran cantidad de información. A veces, en las ciencias sociales este tipo de trabajo tiene que ver con estudios de área o de caso en los que el centro de atención es un hecho determinado o una decisión crucial en una nación, unas elecciones, una decisión transcendental o una crisis mundial. ¿Por qué se destruyó tan subitamente el régimen germánico oriental en 1989? De forma más general, ¿por qué casi todos los regímenes comunistas de Europa del Este se vinieron abajo en 1989? A veces, aunque desde luego no siempre, se puede elegir un acontecimiento como ejemplo de un cierto tipo de hechos, como sería el caso de una revolución política o de una determinada comunidad que decide oponerse a la instalación de un depósito de residuos. Este tipo de trabajo suele estar relacionado con estudios de área en los que se aborda la historia y la cultura de una determinada parte del mundo. El lugar o acontecimiento en cuestión se analiza de cerca y con todo detalle.

Durante varias décadas, los politólogos han venido contrastando las ventajas de los estudios de caso con las de los estadísticos, las de los estudios de área con las de los trabajos comparativos, así como las de las investigaciones políticas «científicas», que usan métodos cuantitativos, con las de las «históricas», que se basan en una profunda comprensión de las fuentes escritas o los contextos. Algunos cuantitativistas creen que en las ciencias sociales el único camino que lleva a la verdad es el análisis estadístico sistemático. Los partidarios de la investigación cualitativa se oponen violentamente a esta postura.

pers. lógica

Además, nada en nuestro conjunto de reglas implica que tengamos que llevar a cabo el experimento perfecto (si es que existe tal cosa) o reunir todos los datos relevantes antes de poder extraer inferencias científicas válidas. Merece la pena estudiar un asunto importante aunque se disponga de poca información. Si aplicamos cualquier diseño de investigación a esta situación las conclusiones serán relativamente inciertas, pero si se admite honestamente tal incertidumbre esta clase de estudio será muy útil. Con frecuencia, la investigación social se caracteriza por disponer de una información limitada. Como el mundo social cambia rápidamente, los análisis que nos ayudan a comprender las transformaciones han de describirlos e intentar interpretarlas en su contexto, aunque la incertidumbre de nuestras conclusiones sea alta. La necesidad de resolver un problema puede ser tan grande que quizás haya datos —recogidos con los métodos científicos más útiles— que se queden obsoletos antes de organizarlos. Si una persona empujaba un cuestionario de cinco páginas sobre psicopatía no sea la mejor estrategia. Joseph Schumpeter citó una vez a Albert Einstein, que había dicho que «siempre que nuestras proposiciones sean ciertas no dirán nada acerca de la realidad, y siempre que no digan nada acerca de la realidad, no serán ciertas» (Schumpeter [1936], 1991, pp. 298-299). Sin embargo, aunque la certeza no pueda alcanzarse, la seguridad en nuestras conclusiones, así como su fiabilidad, validez y sinceridad, podrán incrementarse si prestamos atención a las reglas de la inferencia científica. Las ciencias sociales que propugnamos pretendan traer del mundo inferencias descriptivas y causales. Los que no compartan la premisa de que es posible un conocimiento parcial e imperfecto ni aspiran a una comprensión descriptiva y causal tendrán que buscar en otra parte inspiración o batallas sobre paradigmas en las que participar.

En resumen, aquí no se dan recetas para la investigación científica empírica. Ofrecemos ciertos preceptos y reglas cuya intención es disciplinar el pensamiento, no sofocararlo. Tanto en la investigación cuantitativa como en la cualitativa aplicamos de manera imperfecta reglas inferenciales teóricas a diseños de investigación y a datos empíricos inherentemente imperfectos. Toda reglamentación que tenga sentido admitirá excepciones, pero hay que pedir que estas se justifiquen de manera explícita, que se evalúen sus consecuencias para la fiabilidad de la investigación y que se plantee la incertidumbre de las conclusiones. No buscamos un dogma, sino un pensamiento disciplinado.

1.2 Definición de investigación científica en las ciencias sociales

Nuestra definición de «investigación científica» es un ideal al que toda investigación cuantitativa y cualitativa, incluso la más cuidadosa, sólo puede aproximarse. Sin embargo, necesitamos definir lo que es una buena investi-

(1993), Robert D. Putnam y sus colegas entrevistaron a 112 parlamentarios regionales italianos en 1970, a 194 en 1976 y a 234 entre 1981 y 1982. También enviaron por correo un cuestionario a más de 500 líderes comunitarios de todo el país en 1983. Además, ex profeso para este trabajo, se llevaron a cabo cuatro encuestas de alcance nacional. No obstante, entre 1976 y 1989 los autores del libro realizaron estudios de caso pormenorizados sobre la vida política de seis regiones. Los investigadores, con la intención de someterse a la «prueba del trauma intercultural», «llegaron a conocer con detalle las maniobras y personalidades políticas que habían animado la política regional en las dos últimas décadas» (Putnam, 1993, p. 190).

Las lecciones de estos análisis debieran ser claras: cualquiera que sea el tema de estudio, ni la investigación cuantitativa será mejor que la cualitativa, ni al contrario. Dado que muchos asuntos de interés para las ciencias sociales, si han de tener sentido, no pueden formularse de manera que sea posible la comprobación estadística de hipótesis mediante datos cuantitativos, no queremos animar a nadie a que utilice únicamente técnicas cuantitativas. No pretendemos sacar de la biblioteca a todos los científicos sociales para llevarlos a centros informáticos ni sustituir conversaciones irrepetibles por entrevistas estructuradas. En realidad, lo que señalamos es que los estudios no estadísticos tendrían resultados más fiables si los investigadores prestaran más atención a las reglas de la inferencia científica, que a veces se formulan mejor con el estilo de la investigación cuantitativa. Los precisos métodos estadísticos que sustentan esta tendencia conforman modelos formales abstractos que son aplicables a todo tipo de investigación, incluso a aquellas en las que las variables no pueden medirse cuantitativamente. El carácter extremadamente abstracto, e incluso poco realista, de los modelos estadísticos es lo que hace que las reglas inferenciales resalten con tanta claridad.

Las reglas de la inferencia que analizamos no son relevantes para todos los problemas que los científicos sociales consideran importantes. Muchos asuntos cruciales de la vida política —relacionados con conceptos como el de acción (*agency*), obligación, legitimidad, ciudadanía o soberanía y con la correcta relación entre las sociedades nacionales y la política internacional— tienen un carácter filosófico más que empírico. Sin embargo, estas reglas sí que tienen importancia para toda investigación en la que el objetivo sea conocer lo que ocurre en el mundo real. De hecho, lo que distingue a las ciencias sociales de la observación ocasional es que la primera pretende hacer inferencias válidas mediante el uso sistemático de procedimientos de investigación contrastados. El hecho de que aquí nos centremos en la indagación empírica supone dejar de lado muchos problemas de la filosofía de las ciencias sociales, así como las polémicas que se ocupan del papel del postmodernismo, del carácter y existencia de la verdad, del relativismo y de asuntos similares. Partimos de la base de que es posible lograr cierto conocimiento del mundo exterior, pero tal conocimiento siempre será incierto.

El diseño de la investigación científica tiene como objetivo la extracción de inferencias descriptivas o explicativas a partir de la información empírica que se tenga del mundo. Para hacer estudios científicos suele ser indispensable describir con cuidado ciertos fenómenos, pero la acumulación de hechos no es, en sí misma, suficiente. Estos pueden ser remiñidos (por investigadores cualitativos o cuantitativos) de forma más o menos sistemática, y, evidentemente, la primera es mejor que la segunda; sin embargo, nuestra definición de ciencia conlleva un paso adicional, que es el de utilizar los datos inmediatos para hacer inferencias que conduzcan a algo más amplio que no se observa directamente. Ese algo puede entusar una inferencia descriptiva — utilizar observaciones del mundo para revelar otros hechos que no se han observado — o una inferencia causal — conocer efectos causales a partir de los datos observados. El alcance de la inferencia puede delimitarse en el espacio y el tiempo — el comportamiento electoral en las elecciones estadounidenses desde 1960, los movimientos sociales en Europa del Este desde 1989 — o puede ser más general — el comportamiento humano desde la aparición de la agricultura. En cualquiera de los casos, lo que distingue la investigación científica es el objetivo de hacer inferencias que vayan más allá de las observaciones específicas que se han recogido.

1. El objetivo es la inferencia

Para generar y analizar datos la investigación científica utiliza métodos explícitos, codificados y públicos que, por lo tanto, pueden evaluarse. Gran parte de la investigación social cualitativa sigue procedimientos de búsqueda o inferenciales menos precisos. Como afirmó Robert K. Merton ([1949], 1968, pp. 71-72): «El análisis sociológico de datos cualitativos suele depender de un mundo privado de ideas penetrantes pero insondables y de interpretaciones inefables [...] [Sin embargo,] la ciencia [...] es pública, no privada». La afirmación de Merton no es aplicable a todos los cualitativistas (y, por desgracia, si a algunos cuantitativistas), pero muchos se comportan como si carecieran de método y, a veces, como si la utilización de medios explícitos menoscabara su creatividad. No obstante, no pueden dejar de utilizar alguno. De alguna forma observan los fenómenos, plantean preguntas, extraen información sobre el mundo a partir de esas ob-

2. Los procedimientos son públicos

3. Las conclusiones son inciertas

La inferencia es, por definición, un proceso imperfecto. Su objetivo es utilizar datos cuantitativos y cualitativos para conocer el mundo que los ha producido. Evidentemente, es imposible llegar a conclusiones perfectamente ciertas utilizando datos inciertos. De hecho, la incertidumbre es un aspecto crucial de toda investigación o conocimiento del mundo. Si este factor no se calcula de manera razonable, no se podrá interpretar ninguna descripción del mundo real ni ninguna inferencia causal que le afecte. Un investigador que no contemple abiertamente la incertidumbre estará afirmando que lo sabe todo a la perfección o que no tiene ni idea de la certeza o incertidumbre de sus resultados. En cualquier caso, las inferencias que carecen de un cálculo de incertidumbre no son ciencia tal como aquí se define.

4. El contenido es el método

Para terminar, la investigación científica propugna un conjunto de normas inferenciales de las que depende su validez. Explicar las más importantes es una de las tareas primordiales de este libro. El principal contenido de la «ciencia» son sus métodos y reglas, no su objeto de estudio, ya que podemos utilizar tales métodos para estudiar prácticamente todo. Esta idea fue aceptada hace unos cien años cuando Karl Pearson (1892, p. 16) explicó que «el campo de la ciencia es ilimitado; su materia es infinita; cada conjunto de fenómenos naturales, cada fase de la vida social, cada estadio del desarrollo pasado o presente es materia científica. La unidad de todas las ciencias se basa únicamente en su método, no en su materia».

A su vez, estas cuatro características tienen otra consecuencia: cuando la ciencia es buena, es una empresa social. El trabajo de todo investigador o equipo de investigadores está sometido a limitaciones de conocimiento e ideas y los errores son inevitables, pero es probable que otros los señalen.

la ciencia es pública
 la ciencia es pública
 la ciencia es pública

Inferencia

Comprender el carácter social de la ciencia puede resultar liberador, ya que implica que nuestro trabajo no tiene que estar por encima de la crítica para hacer una aportación importante —ya sea en cuanto a la descripción de un problema o a su conceptualización, a la teoría o a la evaluación de una en concreto. Una aportación será posible siempre que nuestro trabajo contenga explícitamente (o pretenda reinterpretar) las preocupaciones de la comunidad académica y utilice medios públicos para hacer inferencias que respeten las reglas científicas y la información de que disponemos. Incluso la aportación de un artículo menor será siempre más grande que la de un «trabajo importante» que permanezca para siempre en un cajón o continúa en un ordenador.

1.3 Ciencia y complejidad

Las ciencias sociales pretenden explicar situaciones del mundo social que consideramos más o menos complejas. Sin embargo, hay que reconocer que lo que percibimos como complejidad no es algo inherente a los fenómenos: el mundo no está dividido de forma natural en conjuntos de acontecimientos simples o complejos. Por el contrario, la complejidad que se percibe en una situación depende en parte de en qué medida podemos simplificar la realidad adecuadamente, y nuestra capacidad de simplificar depende de si podemos extraer con coherencia resultados y variables explicativas. Tener más observaciones puede ayudarnos en este proceso, pero suele resultar insuficiente. Por lo tanto, «la complejidad» depende, en parte, del estado de nuestra teoría.

Los métodos científicos pueden ser tan valiosos para acontecimientos intrínsecamente complejos como para otros más simples. Es probable que la complejidad haga más inciertas nuestras inferencias, pero no tiene por qué hacerlas menos científicas. La incertidumbre y unos datos limitados no tienen que llevarnos a abandonar una investigación científica. Por el contrario: la compensación más grande a la hora de utilizar las reglas de la inferencia científica se logra precisamente cuando los datos son limitados, los instrumentos de observación tienen defectos, las medidas no están claras y las relaciones son inciertas. Si hay relaciones claras y datos no ambiguos, quizá el método no sea tan importante, ya que incluso reglas o inferencias defectuosas podrían producir respuestas medianamente aceptables. Consideremos algunos acontecimientos complejos y, en cierto sentido, únicos que tuvieron enormes ramificaciones. La caída del Imperio romano, la Revolución francesa, la Guerra Civil de los Estados Unidos, la Primera Guerra Mundial, el holocausto y la reunificación alemana de 1990 son ejemplos de este tipo. El que estos acontecimientos tuvieran lugar parece ser el resultado de una compleja interacción de muchas fuerzas cuya conjunción se antoja crucial para que así fuera. Esto quiere decir que series de

pop. y reglas complejas

300.72 K52d 579277 1. La ciencia en las ciencias sociales

acontecimientos y de fuerzas, cuyas causas son independientes entre sí, convergieron en un lugar y momento determinados, de manera que su interacción parece que ocasionó los acontecimientos que se observan (Hirschman, 1970). Además, suele ser difícil creer que tales acontecimientos fueran consecuencias inevitables de fuerzas históricas a gran escala: algunos parecen haber dependido, en parte, de la idiosincrasia de ciertas personalidades, instituciones o movimientos sociales. De hecho, desde el punto de vista de nuestras teorías, con frecuencia el azar parece haber tenido su papel: factores que se encuentran fuera del alcance de la teoría vincularon de forma determinante la sucesión de los acontecimientos.

Una de las formas de comprender estos hechos es mediante la búsqueda de generalizaciones, que conceptualizan cada caso dentro de un tipo de acontecimientos del que se puede decir que tiene una serie de características generales. Este método suele funcionar bien para guerras o revoluciones ordinarias, pero algunos de estos procesos, al ser mucho más extremos que otros, constituyen «valores atípicos» en la distribución estadística. Además, hay guerras o revoluciones precursoras de gran importancia que pueden tener una influencia tan considerable en fenómenos posteriores del mismo tipo —pensamos de nuevo en la Revolución francesa— que hay que tener cuidado al comparatlas con sus secuelas, ya que, en cierto sentido, éstas pueden ser fruto de la imitación. Expandir el tipo de acontecimiento puede ser útil, pero no siempre resulta apropiado.

También se puede proceder de manera científica con acontecimientos peculiares de gran magnitud utilizando un análisis contrastado, que es «la construcción mental de una serie de acontecimientos que se altera mediante la modificación de una o más "condiciones"» (Weber [1905], 1949, p. 173). La aplicación de esta idea de forma sistemática y científica se pone de manifiesto en un ejemplo especialmente extremo de un peculiar acontecimiento histórico. Stephen J. Gould ha indicado que una de las formas de distinguir entre características evolutivas sistemáticas y acontecimientos estocásticos, aleatorios, puede ser imaginar como sería el mundo si se fijaran todas las condiciones, hasta un determinado punto, y se volviera a dar cuenta al resto de la historia. Señala que si fuera posible «volver a poner en marcha la cinta de la vida» y dejar que la evolución tuviera de nuevo lugar desde el principio, los organismos del mundo actual serían completamente diferentes (Gould, 1989a).

Un acontecimiento único en el que se han centrado últimamente los que estudian la evolución es la súbita desaparición de los dinosaurios hace 65 millones de años. Gould (1989a, p. 318) afirma que «debemos aceptar que nuestra conciencia no se habría desarrollado en el planeta si una catástrofe cósmica no se hubiera cobrado la vida de los dinosaurios». Si tal afirmación fuera cierta, la extinción de estos animales habría sido para los seres humanos tan importante como cualquier otro hecho histórico. Sin embargo,

su desaparición no se ajusta con propiedad al tipo de acontecimientos que se podrían estudiar directamente de forma sistemática y comparativa mediante la aplicación de leyes generales.

No obstante, la extinción de los dinosaurios puede abordarse científicamente, ya que se pueden desarrollar hipótesis alternativas y contrastarlas con sus consecuencias observables. Una de las hipótesis que explican este fenómeno, que desarrollaron Luis Alvarez y sus colaboradores de Berkeley a finales de los años setenta (W. Alvarez y Asaro, 1990), postula que hubo una colisión cósmica: un meteorito chocó con la Tierra a unos 72.000 kiló-

metros por hora, lo cual generó una explosión mayor que la de una guerra nuclear total. Si esta hipótesis fuera correcta, su consecuencia observable sería que el tridido (un elemento habitual en los meteoritos, pero escaso en la Tierra) se encontraría en una determinada capa de la corteza terrestre que corresponde a los sedimentos depositados hace 65 millones de años, de hecho, el descubrimiento de este metal que confirma parcialmente esta teoría. Aunque se ha considerado una prueba que confirma esta teoría. Aunque este acontecimiento es sin duda único, hay otras muchas consecuencias observables. Para dar sólo un ejemplo, tendría que ser posible encontrar el carácter del meteorito en algún lugar de la Tierra (y ya se han encontrado algunos candidatos).

El problema que plantea la causa (o causas) de la extinción de los dinosaurios sigue sin resolverse, aunque la polémica ha generado investigaciones muy valiosas. Por lo que a nosotros respecta, la importancia de este ejemplo radica en que las generalizaciones científicas son útiles incluso cuando se estudian acontecimientos muy poco habituales que no pertenecen a un tipo importante. La hipótesis de Alvarez no puede contrastarse con un conjunto de acontecimientos normales, pero sí tiene consecuencias observables en otros fenómenos evaluables. Sin embargo, hay que señalar que una hipótesis no se considera una explicación razonablemente cierta hasta que no se evalúa empíricamente y se somete a ciertas pruebas estrictas. Como mínimo, sus consecuencias deben ser compatibles con nuestro conocimiento del mundo exterior, y cuando la hipótesis es realmente buena debe predecir lo que Imre Lakatos (1970) denomina «hechos nuevos», o sea, aquellos que no se habían observado hasta entonces.

La cuestión es que incluso acontecimientos aparentemente únicos, como la extinción de los dinosaurios, pueden estudiarse de forma científica si nos ocupamos de mejorar las teorías, los datos y la forma de utilizarlos. Matemática nuestra teoría mediante la clarificación conceptual y la especificación de variables puede generar más consecuencias observables e incluso poner a prueba teorías causales referidas a acontecimientos únicos como la extinción de los dinosaurios. Perfeccionar nuestros datos nos permitirá contemplar un mayor número de consecuencias observables, mientras que utilizar los mejor hasta posible que extrajamos mas consecuencias de ellos. El hecho de estudiar acontecimientos muy complejos no hace irrelevante la

2 Principales componentes del diseño de investigación

Cuando la investigación social es de calidad, constituye un proceso creativo en el que la intuición y el descubrimiento surgen dentro de una consolidada estructura de estudio científico. Para el científico social de primera categoría el diseño de una investigación no es un programa que haya que seguir mecánicamente para recoger y evaluar datos. Por el contrario, el investigador debe ser lo suficientemente flexible como para rechazar las viejas formas de mirar el mundo y plantear nuevas preguntas, revisar los diseños de investigación apropiadamente y recoger después datos diferentes de los que había pensado en un principio. Sin embargo, para que sus conclusiones sean válidas y aceptables para los expertos del área, todas estas revisiones y nuevas consideraciones han de producirse siguiendo procedimientos explícitos que respeten las normas de la inferencia. Un proceso de investigación dinámico tiene lugar dentro de una normativa estable.

Con frecuencia, los científicos sociales comienzan sus investigaciones con un diseño ponderado, recogen algunos datos y extraen conclusiones. Sin embargo, el proceso no suele carecer de obstáculos, y este orden no siempre es el mejor, ya que las conclusiones pocas veces se desprenden fácilmente del diseño de investigación y de los datos que se recogen de acuerdo con él. Una vez que el investigador ha reunido esos datos, tal como indicado con su proyecto, a menudo se encuentra con que el engranaje entre las principales preguntas de la investigación, la teoría y los datos disponibles es imperfecto. En este momento, suele acudir el desánimo y se cree, equivocadamente, que otros científicos sociales hallan un vínculo estrecho e inmediato entre datos e investigación. Esta percepción se debe a que con frecuencia los investigadores retratan los andamiajes después de levantar sus edificios intelectuales, sin dejar apenas rastro de las penalidades e incertidumbres de la construcción. De ahí que el proceso investigador parezca más mecánico y previsible de lo que realmente es.

Algunos de nuestros consejos van dirigidos a los investigadores que intentan relacionar la teoría y los datos. A veces pueden concebir procedimientos de recogida de datos más apropiados para la mejor evaluación de una teoría, en otras ocasiones pueden utilizar los datos de que disponen y reformular una pregunta teórica (o incluso plantear otra completamente diferente que no se contempló en un principio) con el fin de generar un proyecto de investigación más importante. Si el estudio sigue las normas inferenciales, seguirá siendo científico y producirá inferencias fiables acerca del mundo.

Sea flexible

23

to a una política, y también existen manuales para llevar a cabo observación participante en una oficina. Sin embargo, no hay reglas que indiquen cómo elegir el proyecto de investigación que se va a realizar y, si decidimos llevar a cabo trabajo de campo, tampoco las hay para indicarnos dónde tenemos que hacerlo.

Podemos proponer maneras de hacer una muestra de comunidades con el fin de estudiar el impacto de diferentes políticas educativas o formas de conceptualizar los conflictos étnicos que nos llevan a la formulación y comprobación de hipótesis relativas a su incidencia. Sin embargo, no existen normas que nos indiquen si hay que estudiar la política educativa o el conflicto étnico. Desde el punto de vista de la metodología de las ciencias sociales, hay formas mejores y peores de estudiar la caída del régimen de Alemania del Este en 1989, al igual que las hay para abordar la relación entre las opiniones de un candidato en materia fiscal y sus probabilidads de éxito electoral. Sin embargo, no hay ninguna manera de determinar si es mejor estudiar la caída del régimen germano oriental o el papel de los impuestos en las elecciones estadounidenses.

El tema concreto que estudia un científico social puede tener un origen personal e idiosincrásico. No es casual que, habitualmente, sean los miembros de un grupo los que comienzan a estudiarlo: las mujeres han solidificado la batuta en la historia de su propio género, los negros en la de su etnia y los inmigrantes en la de la emigración. Los temas también pueden verse influidos por inclinaciones o valores personales. Es probable que quien estudia el Tercer Mundo tenga más ganas de viajar y tolere mejor condiciones de vida difíciles que el que se ocupa de la elaboración de políticas en el Congreso estadounidense, del mismo modo que quien analice la cooperación internacional puede tener una especial aversión a los conflictos violentos.

Estas experiencias y valores personales suelen proporcionar la motivación para convertirse en un científico social y, posteriormente, para elegir una determinada pregunta de investigación. En consecuencia, pueden ser las «auténticas» razones para embarcarse en un determinado proyecto, y es justo que sea así. Sin embargo, independientemente de lo personales o psicológicas que sean las razones para elegir un tema, los métodos científicos y normas inferenciales que se analizan en este libro ayudarán a los académicos a hacer proyectos de investigación más sólidos. Desde el punto de vista de la posible aportación a las ciencias sociales, las razones personales no son ni necesarias ni suficientes para justificar la elección de un tema. En la mayoría de los casos, no tendrían que aparecer en los escritos académicos. Para decirlo de manera más directa pero menos delicada, a nadie le importa nuestra opinión: a la comunidad académica sólo le interesa lo que podemos demostrar.

Siempre que sea posible, los investigadores, deben también perfeccionar el diseño de su investigación antes de llevar a cabo cualquier trabajo de campo. Sin embargo, los datos disciplinan el pensamiento a su manera. Es muy habitual que un diseño excelente haga aguas cuando se recogen las primeras observaciones: no es que la teoría esté equivocada, sino que los datos no son apropiados para responder a las preguntas planteadas en un principio. Entender desde el comienzo lo que se puede y no se puede hacer en las etapas finales puede ayudar al investigador a prever, al menos, algunos problemas cuando haga el primer diseño de investigación.

Para un mejor análisis, dividimos todos estos diseños en cuatro partes: la *pregunta de la investigación*, la *teoría*, los *datos* y la *utilización de los datos*. Estos componentes no suelen desarrollarse por separado y los estudiosos no se ocupan de ellos con un orden preestablecido. De hecho, para los cualitativistas que comienzan el trabajo de campo antes de elegir una pregunta precisa para su investigación, primero van los datos y después lo demás. Sin embargo, esta peculiar ruptura, que explicamos entre los apartados 2.1 y 2.4 de este capítulo, es especialmente útil para comprender el carácter de los diseños de investigación. Con el fin de determinar con precisión lo que *podría* hacerse si se reorientaran los recursos, nuestros consejos en el resto de este apartado presuponen que el investigador dispone de tiempo y medios limitados. Evidentemente, en cualquier investigación real siempre hay que hacer concesiones. Creemos que si el investigador entiende de los consejos que damos para los cuatro componentes, esto le ayudará a hacer concesiones que mejoren al máximo el proyecto de su investigación, aunque ésta sea objeto de condicionantes externos.

2.1 Mejorar las preguntas de la investigación

A lo largo de este libro se plantea una vez que identificamos el objeto de estudio. Cuando hay una pregunta para la investigación, ¿cómo podemos llevar ésta a cabo para dar explicaciones válidas a los fenómenos sociales y políticos? Nuestro análisis comienza con una pregunta para la investigación y continúa después con las fases de diseño y realización. Pero, ¿dónde se originan las preguntas de un estudio? ¿Cómo elige un investigador el tema de análisis? No hay respuesta sencilla a estas preguntas. Al igual que otros autores, Karl Popper (1968, p. 32) ha señalado que «no existe algo que pueda llamarse método lógico para tener nuevas ideas [...] El descubrimiento contiene un "elemento irracional" o una "inición creativa». En los primeros estadios del proceso de investigación, las normas de elección están menos formalizadas que las de otras actividades de ese mismo proceso. En el área de la elección social existen textos que tratan sobre cómo diseñar experimentos de laboratorio; en estadística hay criterios para seleccionar la muestra de una encuesta que trate sobre las actitudes respec-

Lo ideal sería que todos los proyectos de investigación de las ciencias sociales cumplieran dos condiciones. En primer lugar, un proyecto de investigación tiene que plantear una pregunta «importante» para el mundo real. El tema ha de ser relevante para la vida política, social o económica, para entender algo que afecte de manera significativa a la vida de muchas personas o para comprender o predecir acontecimientos que pudieran resultar dañinos o benéficos (véase Shively, 1990, p. 15). En segundo lugar, un proyecto de investigación tiene que hacer una aportación concreta a lo escrito en un área académica identificable, aumentando la capacidad colectiva de dar explicaciones científicas verificables a algún aspecto del mundo. Esta segunda condición no supone que todas las investigaciones que aporten algo a nuestra reserva de explicaciones para las ciencias sociales pretendan en realidad hacer inferencias causales. En ocasiones, la situación del conocimiento en un determinado campo hace que se precisen muchos datos y descripciones antes de afrontar el reto de dar una explicación. A veces, la aportación de un proyecto sólo es la inferencia descriptiva, mientras que en otros el objetivo ni siquiera es hacer un resumen de hechos históricos. Sin embargo, en este caso se cumple nuestra segunda condición, porque tales acontecimientos son un requisito imprescindible para la explicación.

La primera condición dirige nuestro interés al mundo real de los fenómenos políticos y sociales, y también a cómo se registran los acontecimientos y problemas actuales e históricos que configuran la vida de las personas. Determinar si una pregunta de investigación cumple dicho criterio es algo primordialmente social. La segunda condición nos sitúa en la bibliografía académica de las ciencias sociales, es decir, en las perpallidades intelectuales que no se han planteado todavía, en las que están por resolver, así como en las teorías y métodos científicos de que se dispone para resolverlas.

A los politólogos no les resulta difícil encontrar un objeto de estudio que cumpla la primera condición. A lo largo de los últimos cuarente años diez guerras importantes se han cobrado la vida de casi treinta millones de personas (Levy, 1985, p. 372); algunas «guerras limitadas», como las que libraron los Estados Unidos y Vietnam del Norte, o Irán e Irak, han producido cada una un número de víctimas cercano al millón; del mismo modo, si hubiera una guerra nuclear, ésta podría eliminar a miles de millones de seres humanos. La mala gestión de los políticos, tanto en el ámbito nacional como en el internacional, ha generado —como en los años treinta— privaciones económicas de carácter global y crisis regionales y locales, tal como lo demuestran las trágicas experiencias de gran parte de África y América Latina durante los años ochenta. En general, las variaciones entre países en lo tocante a instituciones políticas se relacionan con las diversas condiciones de la vida ordinaria, las cuales se reflejan en la diferente esperanza de vida o mortalidad infantil que tienen países con un grado de desa-

trollo económico similar (Russett, 1978, pp. 913-28). Dentro de los Estados Unidos, la eficacia de los programas que combaten la pobreza o la desarticulación social parece haber variado de forma considerable. No cabe duda de que será importante toda investigación que contribuya, aunque sea de forma marginal, al conocimiento de estos problemas.

Aunque los científicos sociales tienen muchas cuestiones significativas que investigar, las herramientas para entenderlas son escasas y poco elaboradas. Hay muchos escritos sobre la guerra o el sufrimiento en las sociedades que poco contribuyen a la comprensión de tales problemas porque no los describen de forma sistemática o no hacen inferencias causales o descripciones válidas. Las ideas brillantes pueden mejorar la comprensión al aportar nuevas e interesantes hipótesis, pero la brillantez no es un método de investigación empírica. Todas las hipótesis han de ser contrastadas empíricamente antes de que hagan una aportación al conocimiento. Este libro no da consejos sobre cómo ser brillante. Sin embargo, lo que sí puede hacer es recalcar la importancia de que la investigación se lleve a cabo de manera que constituya una aportación al conocimiento.

Nuestro segundo criterio para elegir una pregunta de investigación, «hacer una aportación», significa explícitamente que hay que situar el diseño del estudio dentro del marco de la bibliografía científica existente. Esto garantiza que el investigador comprenda el «estado de la cuestión» y reduzca al mínimo la posibilidad de repetir lo que ya se ha hecho. También garantiza que el trabajo sea importante para otros, aumentando así el éxito del conjunto de la comunidad académica. Se puede hacer una aportación explícita a lo escrito sobre el tema de varias maneras. A continuación se enumeran algunas de ellas:

1. Eligiendo una hipótesis que los estudiosos consideren importante en la bibliografía pero de la que no se haya realizado un estudio sistemático. Si encontramos pruebas a favor o en contra de la hipótesis, estaremos haciendo una aportación.
2. Eligiendo una hipótesis aceptada en la bibliografía que sospechemos es falsa (o creamos que no se ha demostrado adecuadamente) e investigando si realmente lo es o si otra teoría es correcta.
3. Intentando resolver o presentando más pruebas a favor de uno de los bandos en una polémica que esté presente en lo escrito hasta el momento; demostrando quizá que toda esa polémica carece de fundamento desde el principio.
4. Diseñando investigaciones que arrojen luz o evalúen premisas no cuestionadas por la bibliografía.
5. Señalando que en la bibliografía no se ha concedido atención a un asunto importante y proceder a aportar a ese campo un estudio sistemático.
6. Señalando que las teorías o pruebas relativas a cierto objetivo en un campo podrían aplicarse a otro para solucionar un problema existente, pero aparentemente se alejado.

Si nos preocupamos demasiado por hacer una aportación a la bibliografía académica sin prestar cierta atención a asuntos importantes para el mundo real, corremos el riesgo de formular preguntas insignificantes desde el punto de vista político. Por el contrario, centrarse en el contexto político actual sin preocuparse de hasta qué punto un problema social es susceptible de investigarse sistemáticamente dentro del marco de un cuerpo de conocimiento conduce a trabajos descuidados que poco añaden a una comprensión profunda.

Nuestros dos criterios para elegir preguntas de investigación no son necesariamente opuestos. A largo plazo, la comprensión de los fenómenos del mundo real aumenta con la elaboración y evaluación de hipótesis explicativas que propugna el método científico. Sin embargo, de forma más inmediata, puede haber contradicción entre lo que es útil desde el punto de vista práctico y lo que finalmente tiene valor. Por ejemplo, Mankiw (1990) señala que la teoría macroeconómica y la macroeconomía aplicada se apartaron considerablemente la una de la otra en los años setenta y ochenta: modelos cuya incoherencia teórica se había demostrado siguieron utilizándose para hacer pronósticos sobre la economía estadounidense, mientras que los nuevos modelos teóricos, concebidos para corregir esos fallos, continuaron teniendo un carácter especulativo y no se desarrollaron lo suficiente como para realizar predicciones ajustadas con ellos.

Cuando un investigador elige un tema, puede parecer que los criterios de aplicabilidad al mundo real y aportación al progreso científico son opuestos. Algunos investigadores comenzarían con un problema del mundo real que sea muy significativo socialmente, como la amenaza de guerra nuclear, la disparidad de ingresos entre hombres y mujeres o la transición a la democracia en Europa del Este. Otros quizá partan de un problema intelectual planteado en la bibliografía de las ciencias sociales: la contradicción entre diversos estudios experimentales de tomas de decisiones en condiciones de incertidumbre o la existente entre diversas teorías de voto referidas a las elecciones para el Congreso estadounidense y los últimos resultados en las urnas. Evidentemente, no hay una distinción estricta entre los criterios. Algunas preguntas de investigación responden a los dos desde el principio, pero, a la hora de diseñar el estudio, los investigadores suelen comenzar es-

Cuando el punto de partida es un problema significativo del mundo real y no un área bibliográfica ya establecida, resulta esencial que elaboremos un plan de estudio factible. Si se propone un asunto que no puede desarrollarse hasta convertirse en un proyecto de investigación concreto, qué permite la extracción de inferencias descriptivas y causales válidas, hay que modificarlo o abandonarlo. Del mismo modo, también hay que cambiar toda propuesta que no haga alguna aportación a los escritos académicos. Al elegir provisionalmente un tema, estamos dialogando con su bibliografía. ¿Qué preguntas de interés para nosotros se han respondido ya? ¿Cómo se puede plantear y pulir nuestra pregunta para que parezca posible contestarla con las herramientas disponibles? Podemos partir de un asunto candente, pero tendremos que hacer frente tanto a lo escrito en las ciencias sociales como a problemas inferenciales.

2.2 Mejorar la teoría

En las ciencias sociales, una teoría es una especulación razonada y precisa sobre la respuesta que cabe dar a la pregunta de una investigación, e incluye una declaración de por qué tal respuesta es correcta. Las teorías suelen conllevar hipótesis descriptivas o causales más específicas. Una teoría debe estar en consonancia con los datos disponibles anteriormente sobre una pregunta de investigación. «Una teoría que no tiene en cuenta los datos existentes es una incoherencia. Si dispusiéramos de algo equivalente a la legislación sobre "veracidad" en la publicidad», esta incoherencia no se llamaría teoría» (Liebertson, 1992, p. 4; véase también Woods y Walton, 1982).

En primer lugar, hay que elegir teorías que puedan estar equivocadas. En realidad, se aprende mucho más de las que están equivocadas que de aquellas que se enuncian de forma tan general que puede que ni siquiera lo estén en principio. Es necesario poder responder directamente a la pregunta: ¿qué datos nos convencen de que estamos equivocados? Si no hay respuesta para esta pregunta, tampoco hay teoría.

En segundo lugar, para asegurarse de que una teoría es falsable, hay que elegir una que pueda generar tantas consecuencias observables como sea posible. Tal elección posibilitará que la teoría se someta a más pruebas, utilizando más datos y más variados, que esté más veces en peligro de ser falsada y que se puedan recoger más datos para darle una mayor consistencia.

En tercer lugar, al diseñar teorías es preciso ser lo más concreto posible. Las teorías e hipótesis que se formulan con vaguedad no sirven más que para ofuscar. Se puede demostrar más fácilmente que una teoría está equivocada cuando se ha enunciado con precisión y hace predicciones específicas; por consiguiente, esta teoría será mejor.

Algunos investigadores recomiendan que se siga el principio de «conciencia». Por desgracia, este término se ha utilizado tantas veces en conversaciones informales y en escritos académicos que el principio ha quedado eclipsado (véase Sobor [1988] para un análisis completo). Jeffreys (1961, p. 47) dio la definición más clara de conciencia: «Las teorías más simples tienen en principio más probabilidades». Por lo tanto, la conciencia es un juicio, e incluso un supuesto, sobre la naturaleza del mundo: se presupone que es simple. Elegir teorías cuya premisa es la existencia de un mundo simple es una regla que se aplica con claridad en situaciones en las que se está muy seguro de que el mundo es así realmente. Los físicos parecen cómodos con la conciencia, pero los biólogos suelen creer que es absurda. En las ciencias sociales, los hay que la defienden con vehemencia para sus debates (por ejemplo, Zelnor, 1984), pero nosotros creemos que sólo es útil en ciertas ocasiones. Dada la precisa definición de conciencia como supuesta acerca del mundo, nunca deberíamos insistir en que fuera un principio general a la hora de diseñar teorías, aunque sea útil en aquellas situaciones en las que haya indicios de que el mundo que estudiamos es simple. Lo que queremos decir es que no recomendamos al investigador que busque la conciencia como un bien en sí mismo, porque, a menos que ya se-

podamos mucho de un asunto, no parece que haya muchas razones para practicarla. Ni siquiera necesitamos este principio para evitar teorías excesivamente complicadas, porque está implícito en la máxima de que la teoría debe ser tan complicada como indican los datos. Las situaciones en las que no hay información suficiente sobre la complejidad de la teoría que se investiga pueden conducir a lo que se denomina «diseños de investigación imprecisos» (véase el apartado 1 del capítulo 4), pero éstos son problemas del diseño de la investigación y no supuestos acerca del mundo.

Todos los consejos que hemos dado hasta pueden aplicarse si aún no hemos reunido nuestros datos ni hemos comenzado a analizarlos. Sin embargo, si están ya recogidos, podemos utilizar esas reglas para modificar nuestra teoría y conseguir más datos, generando de este modo más consecuencias observables sobre la nueva teoría. Por supuesto, este proceso es caro, cuesta mucho tiempo y quizá vayamos a desperdiciar los datos que ya se han recogido. ¿Qué ocurre entonces cuando está claro que hay que mejorar nuestra teoría pero no podemos permitirnos recoger más datos? Esta situación —en la que con frecuencia se encuentra el investigador— ha de tratarse con cautela y dominio de uno mismo. Cualquiera investigador inteligente puede elaborar una teoría «plausible» con el conjunto de datos que surge de un hecho, sin embargo, esto no diría nada sobre la veracidad de la teoría, que puede ajustarse muy bien a los datos y ser completamente falsa —y, de hecho, su falsedad ser demostrable con otros muchos datos. A los seres humanos se les da bien reconocer pautas, pero no tan bien detectar la falta de las mismas (! la mayoría vemos pautas en manchas de tinta salteadas!). A una teoría hay que racionarle los ajustes realizados sobre la marcha que no sean coherentes con los datos disponibles, y hay que hacerlos con bastante disciplina.

Todavía nos queda el problema de qué hacer cuando hayamos terminado tanto la recogida de datos como nuestro análisis y queremos mejorar nuestra teoría. En esta situación, recomendamos que se respeten dos normas: en primer lugar, si nuestra predicción se halla condicionada por diversas variables y queremos librarnos de una de las condiciones, podemos hacerlo. Por ejemplo, si al principio nuestra hipótesis era que los países democráticos con sistemas avanzados de protección social no luchan entre sí, sería aceptable extender tal hipótesis a todas las democracias contemporáneas y de este modo contrastar nuestra teoría en un mayor número de casos y aumentar las posibilidades de que sea falsada. El propósito general es que, después de ver los datos, podamos modificar la teoría para que se pueda aplicar a una gama más amplia de fenómenos. Esta alteración de nuestras tesis, al exponerlas aún más a la falsación, no debería llevarnos a explicaciones para salir del paso que sólo pretendan «salvar» una teoría inadecuadamente ajustándola a fenómenos con los que ya se ha comprobado que concuerda.

Sin embargo, la práctica opuesta no suele ser apropiada. Después de observar los datos, no deberíamos únicamente añadir una condición restrictiva

para determinar si nuestras especulaciones son correctas. Este ejercicio puede ser bastante útil, especialmente a la hora de señalar áreas de las que los futuros investigadores podrían ocuparse.

Como hemos señalado anteriormente, ya se sabe que las ciencias sociales no siguen normas estrictas: ¡la necesidad de creatividad a veces exige que se prescindiera del manual! Y los datos pueden disciplinar el pensamiento. De ahí que a los investigadores, una vez que se enfrentan a los datos, se les ocurra cómo deberían haber construido la teoría en un principio. Aunque sea de forma restrictiva, tal modificación quizá valga la pena si podemos convencernos a nosotros mismos y a los demás de que reformar la teoría tal como lo proponemos es algo que habríamos hecho antes de recoger los datos si se nos hubiera ocurrido. Sin embargo, hasta que la teoría no se contraste con nuevos datos, su posición seguirá siendo incierta, y así habrá que considerarla.

Una de las consecuencias importantes de estas normas es que los proyectos piloto suelen ser muy útiles, especialmente en investigaciones en las que los datos han de recogerse mediante entrevistas u otros medios participativamente costosos. La recogida de datos preliminar puede llevarnos a atender las preguntas de la investigación o a modificar la teoría. Posteriormente, se pueden recabar otros datos para contrastar la nueva teoría, evitándose así el problema de utilizar la misma información para elaborar y comprobar una teoría.

2.3 Mejorar la calidad de los datos

Los «datos» son informaciones sobre el mundo recogidas de forma sistemática y pueden ser de tipo cuantitativo o cualitativo. A veces se reúnen datos para evaluar una teoría muy concreta, pero no es inusual que los académicos los recojan antes de saber exactamente qué es lo que les interesa encontrar. Además, aunque los datos se reúnan para contrastar una hipótesis determinada, puede que al final a los investigadores les interesen preguntas que no se les habían ocurrido anteriormente.

En cualquier caso, ya sea que los datos se reúnan para un determinado propósito o que se utilicen para otro que no está claro cuando se realiza la recogida, existen ciertas normas que mejoran la calidad de tales datos. En principio, podemos considerar esas normas al margen de las que se han mencionado en el apartado 2.2 para mejorar la teoría. En la práctica, para toda recogida de datos, se necesita cierto grado de teoría, del mismo modo que para formular cualquier teoría se precisan algunos datos (véase Combs, 1964).

Nuestra primera y más importante directriz para mejorar la calidad de los datos es: *registrar y detallar el proceso con el que se generan los datos*. Sin esta información no podemos determinar si la utilización de procedi-

y continuar como si esa matización hubiera demostrado que nuestra teoría es correcta. Si la teoría original fuera que las democracias contemporáneas no luchan entre sí porque tienen sistemas constitucionales, sería menos aceptable que, al encontrar excepciones a nuestra «regla», limitáramos la proposición a las democracias que disponen de un sistema avanzado de protección social una vez que se ha constatado mediante la inspección de los datos que esta matización parece convertir nuestra proposición en correcta. O suponíamos que nuestra teoría inicial fuera que las revoluciones sólo tienen lugar en situaciones de grave crisis económica, pero nos damos cuenta de que no es verdad en uno de nuestros estudios de caso. En esta situación no sería razonable añadir únicamente condiciones generales como: las revoluciones nunca tienen lugar durante períodos de prosperidad excepto a allí donde el ejército es débil, los líderes políticos son represivos, la economía es correcta, excepto en el país X». Ya que hemos descubierto que nuestra teoría no se aplica en dicho país, no es de mucha ayuda convertir esta falsación en una generalización espumosa. Si no nos esforzamos por recabar nuevos datos, no responderemos de pruebas aceptables que sostengan la nueva versión de la teoría.

Por lo tanto, nuestra regla básica en cuanto a la alteración de una teoría una vez que se han observado los datos es: *podemos hacer que la teoría sea menos restrictiva (de forma que contemple una gama más amplia de fenómenos y se vea expuesta a más oportunidades de falsación), pero esto no ha de hacerse sin haber recogido más datos, con los que contrastar la nueva versión de la teoría*. Si no podemos recabar más datos, estaremos en punto muerto y no podemos proponer ningún método mágico para dejar de estar hecho, algunos resultados negativos pueden ser bastante útiles para un área académica. ¿Quién no habría de preferir una sólida conclusión negativa an-

Además, si estamos equivocados, no tenemos por qué dejar de escribir una vez que hayamos admitido nuestra derrota. Podemos añadir un apartado a nuestro artículo o un capítulo en el libro para ocuparnos de futuras investigaciones empíricas y de la presente especulación teórica. En este contexto, disponemos de bastante libertad y podemos señalar nuevas condiciones que sería plausible introducir en nuestra teoría — si creemos que pueden solucionar el problema —; plantear la modificación de otra teoría existente o apuntar hacia una gama de teorías completamente diferente. En tal situación, no podemos llegar a ninguna conclusión con mucha certeza (excepto, quizá, al afirmar que la teoría que planteamos al principio está equivocada), pero sí podemos permitirnos el lujo de inventar nuevos datos de investigación o proyectos de recogida de datos que podrían utilizarse para salir del paso?

mientos normalizados de análisis, producirá inferencias sesgadas. Sólo si conocemos el proceso de obtención de los datos podremos hacer inferencias descriptivas y causales válidas. En una encuesta de opinión de tipo cuantitativo, para registrar el proceso de obtención de datos es preciso conocer exactamente de qué manera se ha realizado la muestra y qué preguntas se han hecho. En un estudio de caso comparado de tipo cualitativo es crucial enumerar las normas precisas que hemos utilizado para elegir el reducido número de casos que se van a analizar. En el capítulo 6 damos más directrices para la selección de casos en investigación cualitativa, pero aún más importante que elegir un buen método es registrar y detallar con cuidado cualquier que se utilice, así como la información necesaria para que otros lo apliquen.

En el apartado 2.2 defendimos teorías de las que pudieran extraerse muchas consecuencias observables. Nuestra segunda directriz para mejorar la calidad de los datos es: *para evaluar mejor una teoría hay que recoger datos acerca de la mayor cantidad posible de consecuencias observables.* Esto supone recoger tantos datos como sea posible en los contextos más diversos. Cada consecuencia adicional de nuestra teoría que observemos aportará un nuevo contexto en el que poner a prueba su veracidad. Cuantas más consecuencias observables encontremos que sean coherentes con la teoría, más penetrante será la explicación y más ciertos sus resultados.

Al añadir datos sobre nuevas consecuencias observables de una teoría, prevé la teoría, tendremos más confianza en ella. Pensemos, por ejemplo, en la teoría de la disuasión racional: los potenciales iniciadores de una guerra calculan los costes y beneficios de atacar a otros estados, y en estos cálculos pueden influir amenazas de represalia crebles. La comprobación más directa de esta teoría sería evaluar si, dadas ciertas amenazas de guerra, la decisión de atacar se asocia con factores como el equilibrio de fuerzas militares entre los posibles contendientes o con los intereses que están en juego para el país atacado (Huth, 1988). Sin embargo, aunque utilizar únicamente casos en los que se lanzan amenazas conforme un conjunto de consecuencias observables de la teoría, éstas son sólo una parte de las que se podrían recoger (y si sólo se contemplan éstas se podría producir un sesgo de selección), ya que aquellas situaciones en las que se disuade incluso de amenazar serían excluidas de los datos. Por lo tanto, también podría merecer la pena recoger información sobre otra variable dependiente (o sea, sobre otro conjunto de consecuencias observables) que calibrara si los estados amenazan cuando tienen incentivos para hacerlo.

Siempre que se carezca de datos suficientes sobre la disuasión en política internacional, también puede ser útil poner a prueba otra teoría que parta de presupuestos similares respecto a la motivación y que se aplique a otra variable dependiente en diferentes condiciones, sin dejar de ser una consecuencia observable de la misma teoría. Por ejemplo, podríamos concebir un experimento de laboratorio para ver si, en circunstancias simuladas, la capacidad militar y unas negociaciones firmes, en vez de acentuar las «amenazas», hacen que no se produzcan. También podríamos examinar si, en situaciones análogas, actores como los oligopolios que compiten por su cuota de mercado o las familias del crimen organizado que luchan por el territorio utilizan estrategias de disuasión y hasta qué punto tienen éxito en diversas circunstancias. De hecho, los economistas que trabajan en el área de la organización industrial han utilizado la teoría de juegos no cooperativa, en la que también se basa la de la disuasión, para estudiar problemas como el de la entrada en ciertos mercados o las estrategias de precios (Fudenberg y Tirole, 1989). Como las teorías se parecen tanto, los datos empíricos en los que se basan las predicciones de la teoría de juegos acerca del comportamiento de las empresas aumentarían la plausibilidad de otras hipótesis relacionadas con la acción del Estado en política internacional. Seguiría habiendo incertidumbre al aplicar las conclusiones de un campo al otro, pero el problema es lo suficientemente importante como para justificar que se intente sacar ideas y pruebas de donde sea posible.

Es evidente que una recogida de datos que siempre prescinda del análisis impedirá la finalización de una investigación útil en vez de facilitarla. En la práctica, las limitaciones de tiempo y de recursos siempre condicionarán los esfuerzos por recoger datos. Aunque más información, más casos, más entrevistas, una nueva variable y otras formas relevantes de recopilación de datos siempre vayan a dar más certeza a nuestras inferencias en algún sentido, puede que haya investigadores prometedores que se vean perjudicados en igual medida por el exceso de información que por su escasez. Insistir en leer un libro más o en hacerse con otra matriz de datos brutos sin escribir una línea es recetar imprudencia.

La tercera directriz es: *máximizad la validez de nuestras mediciones.* La validez tiene que ver con calibrar lo que creemos que estamos calibrando. El índice de paro puede ser un buen indicador de la situación económica, pero uno y otra no son sinónimos. En general, lo más fácil es maximizar la validez ajustándose a los datos sin permitir que se nos interpongan conceptos no observados o no mensurables. Si un entrevistado responde a nuestra pregunta aduciendo ignorancia, lo que sabemos es que *dijo* que no lo sabía, y para eso sí tenemos una medida válida. Sin embargo, lo que *quiso decir* en realidad constituye un concepto completamente diferente que no puede medirse con un alto grado de confianza. Por ejemplo, en países con gobiernos represivos, aducir ignorancia puede ser, para algunas personas, una forma de criticar el sistema político, para otros es como decir «No sé».

Nuestra cuarta directriz es: asegurarse de que los métodos de recogida de datos son fiables. Fiablez significa que si se aplica el mismo procedimiento de la misma manera siempre se obtendrá la misma medida. Cuando se utiliza un procedimiento fiable en diferentes ocasiones, sin que nada haya ocurrido entre tanto que transforme la «verdadera» situación del objeto que estamos calibrando, se observará el mismo resultado¹⁰. Las medidas fiables también ofrecen los mismos resultados cuando las realizan diferentes investigadores, y ello depende, evidentemente, de que se puedan seguir procedimientos explícitos¹¹.

Nuestra última directriz es que todos los datos y análisis deben ser en la medida de lo posible, reproducibles. La capacidad de reproducción no sólo se aplica a los datos, con el fin de que podamos comprobar si nuestras medidas son fiables, sino que también tiene que ver con el proceso de razonamiento que se utiliza para alcanzar conclusiones. Partiendo del informe de nuestra investigación, otro investigador debería poder copiar los datos y rastrear la lógica de nuestras conclusiones. La capacidad de reproducción es importante aunque nadie repita nuestro estudio. Los procedimientos y métodos utilizados sólo se podrán evaluar si el estudio se describe con el suficiente detalle como para que pueda reproducirse.

En algunos tipos de investigación reproducir los datos puede resultar difícil o imposible: quizá los entrevistados mueran o desaparezcan y no sea posible reproducir las observaciones directas de acontecimientos del mundo real que han hecho testigos o participantes en ellos. La capacidad de reproducción también significa cosas diferentes en cada tradición investigadora. En los estudios cuantitativos se centra en repetir el análisis a partir de los mismos datos. Como bien sabe cualquiera que haya intentado alguna vez reproducir los resultados cuantitativos de obras publicadas, la tarea suele ser bastante más difícil de lo que debería — aunque las obras sean de renombrado — y siempre resulta más útil de lo que parece al principio (acerca de la reproducción en investigación cuantitativa, véase Dewald y otros, 1986).

En la investigación cualitativa tradicional la analogía la proporcionan las notas a pie de página y las revisiones bibliográficas. Utilizando estas herramientas, los sucesivos investigadores tendrían que poder localizar las fuentes utilizadas en una obra publicada y hacer sus propias evaluaciones sobre las inferencias que han surgido de esa información. Si los estudios se basan en la observación directa, la reproducción es más difícil. Un investigador puede tomar prestadas las notas de campo de otro o la grabación de sus entrevistas para comprobar si este material fundamenta las conclusiones extraídas en primera instancia. Este nuevo análisis de resultados, utilizando los mismos datos, no suele hacerse, ya que gran parte de la información del trabajo de campo se basa en conversaciones, impresiones y otros actos de participación no registrados. Sin embargo, se podría avanzar bastante si hubiera más investigadores que intentaran hacer este tipo de reproducción, y

así se animaría también a otros a tomar notas de campo más detalladas. Alguna vez se ha reproducido un proyecto de investigación completo, incluyendo la recogida de datos. La reproducción no puede ser perfecta porque es imposible retrotraerse en el tiempo; no obstante, el intento puede resultar bastante fructífero. Quizá la reproducción más exhaustiva que se haya llevado a cabo de un proyecto cualitativo sea la del estudio sociológico de la ciudad estadounidense de Middletown, en Indiana, que comenzaron Robert y Helen Lynd. Su primer estudio de «Middletown» se publicó en 1929 y se reprodujo en un libro que apareció en 1937. Unos cincuenta años después del primer proyecto se está publicando una larga serie de libros y artículos que reproduce los primeros estudios (véase Caplow y otros, 1983a, 1983b y las citas que incluyen). La investigación cualitativa no tiene por qué ser tan exhaustiva, pero este gran proyecto de investigación debería servir como modelo de lo que se puede hacer.

Toda investigación tendría que aspirar a ser lo más reproducible posible: los investigadores habrían de registrar siempre los métodos, normas y procedimientos exactos que han utilizado en la recogida de información y extracción de inferencias para que otros pudieran hacer lo mismo y llegar (esperamos) a la misma conclusión. La capacidad de reproducción también supone que los académicos que utilizan fuentes inéditas o privadas tendrían que comprometerse a que, en el futuro, otros pudieran acceder a este material en condiciones similares; aprovecharse de un acceso privilegiado sin buscar que otros lo tengan imposible la reproducción y pone en cuestión la calidad científica del trabajo. Lo normal es que las investigaciones no sean reproducidas, pero nuestra responsabilidad es comportarnos como si alguien pudiera querer hacerlo. Aunque la obra no sea reproducida, aportar los materiales para que si lo sea hará posible que los lectores comprendan y evalúen nuestro trabajo.

2.4 Utilizar mejor los datos existentes

Solucionar los problemas que plantean los datos mediante la recopilación de otros nuevos y mejores es casi siempre preferible a utilizar de otro modo los datos defectuosos de que disponemos; sin embargo, no siempre es posible recoger otros. Los científicos sociales suelen enfrentarse a datos problemáticos y a pocas posibilidades de adquirir otros mejores, de manera que tienen que sacar el máximo partido a los que tienen.

Como utilizar mejor los datos ya recogidos es el tema principal de las clases de métodos estadísticos y, de hecho, es la mayor aportación de la estadística inferencial a las ciencias sociales. Los preceptos de este asunto, que tan claros están en el estudio de la estadística inferencial, también se aplican a la investigación cualitativa. En el resto del libro nos ocupamos de ellos con más detalle, y aquí sólo vamos esbozar brevemente las

directrices que hay que seguir para utilizar mejor datos recogidos con an- tenidad

En primer lugar, siempre que sea posible, para hacer inferencias hay que utilizar datos que no estén «sesgados», es decir, que como promedio sean correctos. Para entender esta idea tan concreta de la investigación estadística, imaginemos que se aplica la misma metodología (ya sea en la investigación cuantitativa o en la cualitativa) para analizar y extraer conclusiones de datos procedentes de varias matrices de datos brutos. Es probable que, a causa de pequeños errores en los datos o en la aplicación del procedimiento, utilizar sólo una vez esta metodología nunca dará resultados del todo correctos. Un procedimiento «no sesgado» será correcto si se considera como una media de muchas aplicaciones, aunque, por separado, éstas no sean correctas. El procedimiento no inclinará sistemáticamente el resultado en una u otra dirección.

Es evidente que la extracción de inferencias no sesgadas depende tanto de la primera recogida de datos como de su uso posterior y que, como hemos señalado anteriormente, siempre es mejor prevenir los problemas antes de comenzar a reunir los datos. Sin embargo, aquí mencionamos brevemente algunas dificultades porque al utilizar los datos es preciso poner un especial cuidado en el análisis de posibles fuentes de sesgo que hayan pasado desapercibidas cuando se recogían dichos datos. Una de esas fuentes, que puede falsear las inferencias, es la del sesgo de selección, que consiste en elegir observaciones de tal manera que se distorsiona sistemáticamente la población de la que proceden. Aunque un ejemplo obvio es el de escoger a propósito sólo aquellos casos que apoyan nuestra teoría, el sesgo de selección puede ser mucho más sutil. Hay otro sesgo que puede surgir de las variables omitidas y que consiste en excluir alguna variable de control que pudiera influir en una aparente conexión causal entre las variables explicativas y la que queremos aclarar. Entre los capítulos 2 y 6 analizaremos estas y otras posibles dificultades que se presentan en la extracción de inferencias no sesgadas.

La segunda directriz se basa en el concepto estadístico de «eficiencia»: para que los datos se utilicen de manera eficiente hay que maximizar la información que se ha tenido en cuenta para hacer las inferencias descriptivas o causales. Para maximizar la eficiencia no sólo es preciso utilizar todos nuestros datos, sino usar también toda la información relevante que hay en ellos para mejorar las inferencias. Por ejemplo, si los datos están desagregados en pequeñas unidades geográficas, tenemos que utilizarlos de esa forma, no sólo como un agregado nacional. Los agregados más pequeños tendrán un mayor grado de incertidumbre, pero si son, al menos en parte, consecuencias observables de la teoría, contendrán información que puede traerse a colación en el problema inferencial.

3. Temas de estudio de este volumen

Finalizamos este capítulo general señalando los cuatro puntos importantes en la realización de un diseño de investigación que hemos planteado aquí y que analizaremos en detalle a lo largo del libro.

3.1 La utilización de consecuencias observables para vincular la teoría y los datos

En este capítulo hemos subrayado que para que una teoría merezca la pena debemos encontrar si es correcta. Estas consecuencias observables de la teoría han de guiar nuestra recogida de datos y ayudarnos a distinguir entre hechos relevantes e irrelevantes. En el apartado 6 del capítulo 2 analizamos cómo afecta la teoría a la recogida de datos y de qué manera éstos disciplinan la imaginación teórica. Aquí lo que queremos recalcar es que la teoría y la investigación empírica deben estar estrechamente relacionadas. Cualquier teoría que de verdad trabaje para nosotros tendrá consecuencias en la investigación empírica, y ésta no podrá llegar a buen puerto sin una teoría que le guíe en la elección de sus preguntas. Tanto la teoría como la recogida de datos son aspectos esenciales del proceso que nos lleva a decidir si una teoría puede considerarse de forma provisional verdadera o falsa, aunque este sujeto en ambos casos a la incertidumbre que caracteriza a todas las inferencias.

A toda teoría hay que preguntarle cuáles son sus consecuencias observables y en cualquier investigación empírica hay que saber si las observaciones son relevantes para las consecuencias de la teoría y, si es así, que nos permitan inferir sobre la corrección de la misma. En cualquier estudio social científico las consecuencias de la teoría y de la observación de los hechos han de ser elementos imbricados: las conclusiones de las ciencias sociales no pueden considerarse fiables si no se basan en teorías y datos estrechamente relacionados, que se hayan forjado mediante la formulación y análisis de las consecuencias observables de una teoría.

3.2 La maximización del control

El investigador que busca consecuencias adicionales para una hipótesis persiga uno de los objetivos más importantes de las ciencias sociales: *explorar tanto como sea posible de la forma más escueta posible*. La ciencia social de calidad pretende aumentar la trascendencia de lo que se explica mediante la información utilizada en la explicación. Si podemos exponer con precisión, a través de una o varias variables causales, lo que en princ-

El diseño de la investigación social

Hay varias maneras de aumentar el control sobre el problema de nuestra investigación. La principal consiste en incrementar el número de consecuencias observables de nuestra hipótesis y buscar el modo de confirmarlas. Como se ha descrito anteriormente, esta tarea puede suponer (1) una mejora de la teoría para que tenga más consecuencias observables, (2) una mejora de los datos para que realmente se detecten más consecuencias y se utilicen para evaluar la teoría y (3) una utilización más provechosa de esos datos, con el fin de extraer más consecuencias de aquellos que estén disponibles. Ninguna de estas posibilidades, ni tampoco el concepto general de maximización del control, son iguales a la idea de concepción, la cual, como explicamos en el apartado 2.2, es un presupuesto relacionado con la naturaleza del mundo más que una norma para diseñar investigaciones.

La maximización del control es tan importante y tan general que recordamos *ferventemente* que los investigadores hagan listas constantes de aquellas posibles consecuencias observables de sus hipótesis que pudieran detectarse en sus datos o en otros. Quizá se puedan comprobar algunas de esas nuevas consecuencias en la matriz de datos brutos original, siempre que la consecuencia no «surja de» los datos y sea una hipótesis que, de manera independiente, indiquen la teoría u otra matriz de datos. Sin embargo, utilizar otros datos es aún mejor. En este sentido, tendríamos que considerar también consecuencias que pudieran aparecer en nuevos datos —relativos a otras unidades o a otros aspectos de los que estudiamos, datos sobre diferentes grados de agregación y de otros períodos, como las predicciones sobre el futuro próximo— y evaluar las hipótesis en esos ámbitos. Cuantas más pruebas encontremos en diversos contextos, más convincente será nuestra explicación y más confianza podremos tener nosotros y los demás en las conclusiones.

En principio, algunos investigadores podrían oponerse a la idea de recoger consecuencias observables en cualquier sitio o en niveles de agregación diferentes de aquel para el que se ha diseñado la teoría. Por ejemplo, Lieberman (1985) aplica a la investigación cualitativa la idea estadística de «lacia ecológica» —utilizar de forma incorrecta datos agregados para hacer

inferencias sobre individuos— con el fin de prevenir contra las inferencias que utilizan diferentes niveles de agregación.¹² Estamos de acuerdo en que utilizar datos agregados puede llevarnos a hacer inferencias incorrectas sobre los individuos: cuando son éstos los que nos interesan, suele ser mejor estudiarlos directamente si podemos hacerlo con los datos adecuados. Sin embargo, si lo que queremos es que nuestra inferencia plantee una hipótesis con bastante validez, quizá sea mejor que nuestra teoría tenga consecuencias en muchos niveles de análisis diferentes y, a menudo, podremos utilizar datos de todos para dar alguna información sobre ella. De este modo, aunque lo que más nos interese esté en un nivel de análisis agregado, con frecuencia podremos lograr un mayor control sobre la veracidad de nuestra teoría si utilizamos datos de los otros.

Por ejemplo, si elaboramos una teoría para explicar las revoluciones, tendremos que buscar consecuencias observables de ella no sólo en los resultados más generales sino en fenómenos como las reacciones de los revolucionarios a entrevististas en profundidad, en las reacciones de los que residen en pequeñas comunidades de zonas apartadas del país o en declaraciones oficiales de los cargos de los partidos. Tendríamos que estar dispuestos a incorporar cualquier información que pudiéramos lograr, siempre que nos ayudara a conocer la veracidad de nuestra teoría. Si podemos contrastarla examinando los resultados de las revoluciones, perfecto. Pero, en la mayoría de los casos, existe poca información a ese nivel, quizá una o dos observaciones cuyo valor suele ser ambiguo o estar mal medido. La existencia de una revolución es compatible con muchas teorías diferentes. Solo al profundizar en el caso que nos ocupa o aportando información relevante sobre otros será posible diferenciar las teorías que ya existían con anterioridad.

El único problema, a la hora de utilizar información procedente de otros niveles y fuentes para estudiar una teoría que está diseñada con un determinado grado de agregación, reside en dilucidar si tales observaciones contienen alguna información relevante para evaluar las consecuencias de nuestra teoría. Si esas nuevas observaciones ayudan a comprobar la teoría, habrá que utilizarlas aunque no sean consecuencias de gran interés. Por ejemplo, quizá no nos importen en absoluto las opiniones de los revolucionarios, pero si sus respuestas a nuestras preguntas son compatibles con nuestra teoría de la revolución, será más posible que ésta sea correcta y la recogida de información adicional no habrá sido inútil. De hecho, una observación en el nivel de análisis más agregado —que tenga lugar una revolución predicha, por ejemplo— no es más que una consecuencia observada de la teoría y, dada la poca información que contiene, no habrá que concederle más importancia que a otras consecuencias observables. Tenemos que recabar información sobre tantas consecuencias observables de nuestra teoría como sea posible.

3.3 Admitir la incertidumbre

Tanto en la investigación cuantitativa como en la cualitativa el conocimiento y la inferencia son siempre inciertos. La medición en ambos tipos de estudio es propensa al error, pero sus causas pueden ser diferentes. El entrevistador cualitativo, que hace una larga entrevista en profundidad a una persona cuyo contexto ha estudiado, tiene menos posibilidades de calibrar mal la auténtica ideología política de ese sujeto que otro que entreviste de forma estructurada a una persona elegida de forma aleatoria y de la que no sepa nada (aunque lo contrario también es posible si, por ejemplo, el entrevistador se fia excesivamente de un entrevistado que no es fiable). Sin embargo, el encuestador tiende menos que el que investiga en profundidad a extrapolar de forma inapropiada a una población amplia, lo que sólo es aplicable a casos particulares. Ninguno de ellos es inmune a las incertidumbres de la medición o al inherente carácter probabilístico del mundo social.

Todos los buenos científicos sociales —ya sean cuantitativos o cualitativos— contemplan un margen aproximado de incertidumbre en sus inferencias. En la ciencia política, quizá el problema más importante de los estudios cualitativos sea la constante falta de estimaciones de incertidumbre razonables que se observa en las inferencias de los investigadores (véase King, 1990). Siguiendo las normas de este libro se puede hacer una inferencia válida casi en cualquier situación, independientemente de lo limitado que sean los datos, pero tendríamos que evitar sacar conclusiones muy ambiciosas a partir de datos poco consistentes. El problema no es que sea imposible extraer inferencias fiables en la investigación cualitativa, sino que siempre habrá que calcular de forma razonable el grado de incertidumbre que existe en cada una de ellas. Neustad y May (1986, p. 274), al ocuparse de campos en los que es difícil hacer estimaciones cuantitativas precisas, proponen un método útil para animar a los decisores políticos (que a menudo se enfrentan a la necesidad de llegar a conclusiones sobre qué política seguir, utilizando datos inadecuados) a juzgar la incertidumbre de sus conclusiones. Preguntan: «¿Cuánto dinero propio pondrás en ello?».

Esta pregunta tiene sentido siempre que también nos planteemos: «¿Con qué posibilidades?».

3.4 Pensar como un científico social: escepticismo e hipótesis

La incertidumbre de las inferencias causales tiene como consecuencia que los buenos investigadores sociales no las acepten fácilmente. Cuando se dice que A produce B, alguien que «piense como un científico social» se preguntará si esa relación es auténticamente causal. Es fácil hacer esas preguntas respecto a las investigaciones ajenas, pero resulta más importante

3.3 Admitir la incertidumbre

plantearlas en relación a los propios proyectos. Existen muchas razones para tomarse una explicación causal con escepticismo, por plausible que pueda parecer a primera vista. Leemos en el periódico que los japoneses comen menos carnes rojas y que sufren menos ataques cardíacos que los estadounidenses. Esta observación es interesante por sí misma y, además, la explicación resulta creíble: comer demasiados filetes conduce a la alta tasa de enfermedades cardíacas de los Estados Unidos. El científico social esceptico pregunta por la precisión de los datos: ¿cómo conocemos los hábitos alimentarios? ¿qué muestra se ha utilizado? ¿se clasifican los ataques al corazón de forma parecida en Japón y en los Estados Unidos, de manera que podamos comparar fenómenos similares? Suponiendo que los datos sean precisos, ¿qué otros factores podrían explicar los efectos? ¿hay más variables (otras diferencias dietéticas, rasgos genéticos o de forma de vida) que puedan explicar este resultado? ¿Acaso hemos intercambiado la causa y el efecto sin darnos cuenta? Es difícil imaginarse que el hecho de no sufrir un ataque al corazón pueda hacer que comamos menos carne roja, pero es posible. Quizá la gente pierda, con los años, las ganas de comer hamburguesas y filetes. Si así fuera, los que (por la razón que sea) no han sufrido un ataque al corazón vivirían más y comerían menos carne. Este hecho produciría la misma relación que llevó a los expertos a la conclusión de que la carne es la culpable de los ataques al corazón.

No pretendemos cuestionar esos estudios médicos; simplemente queremos poner un ejemplo de cómo los científicos sociales se enfrentan al problema de la inferencia causal: con escepticismo y teniendo en cuenta las explicaciones alternativas que podrían haberse pasado por alto. De este modo, la inferencia causal se convierte en un proceso en el que cada conclusión es una oportunidad de investigar aún más, con el fin de profundizar en esa inferencia y ponerla a prueba. Intentaremos, mediante aproximaciones sucesivas, acercarnos cada vez más a una inferencia causal precisa.

2. La inferencia descriptiva

En las ciencias sociales, la investigación, ya sea cuantitativa o cualitativa, tiene dos objetivos: describir y explicar. Unos investigadores se proponen describir el mundo y otros explicarlo, y ambos propósitos son esenciales. No podemos dar explicaciones causales con sentido sin una buena descripción, y esta, a su vez, pierde gran parte de su interés si no está vinculada a alguna relación causal. La descripción suele venir primero porque es difícil proponer explicaciones antes de saber algo acerca del mundo y qué ha de explicarse en función de qué características. Sin embargo, la relación entre descripción y explicación es interactiva. A veces nuestras explicaciones nos inducen a buscar descripciones de diferentes partes del mundo y, al contrario, éstas pueden llevarnos a nuevas explicaciones causales.

Tanto la descripción como la explicación dependen de las reglas de la inferencia científica. En este capítulo nos centramos en la descripción y en la inferencia descriptiva. Describir no es en absoluto algo mecánico o carente de problemas, ya que para hacerlo hay que elegir entre un número infinito de hechos que podrían registrarse. La descripción científica tiene varios aspectos fundamentales. Uno de ellos es que conlleva un proceso inferencial: describir consiste, en parte, en inferir información sobre hechos no observados a partir de aquellos que sí se han contemplado. Otro aspecto es su capacidad para distinguir lo que tienen de sistemático y de no sistemático los hechos observados.

Debe quedar claro que no estamos de acuerdo con los que denigran la «simple» descripción. Aunque explicar—relacionar causas y efectos—sea

tipos de hechos generales y de principios de comportamiento humano, y no dejarán de ser importantes aunque nuestro único objetivo sea comprender por qué ha dimitido el último ministro de Asuntos Exteriores brasileño. Por ejemplo, estudiando a otros ministros podríamos saber que todo el gobierno brasileño ha dimitido para protestar por las acciones del presidente, algo de lo que no nos habríamos dado cuenta si sólo hubiéramos examinado lo que hacía el de Exteriores.

En las ciencias sociales algunas investigaciones pretenden informar sobre un tipo de acontecimientos específico sin decir nada en particular sobre un hecho o unidad concreta. Los estudios de comportamiento electoral que utilizan encuestas masivas explican las decisiones del conjunto de los votantes, no el sufragio de un individuo concreto. Los estudios de las finanzas del Congreso estadounidense explican la influencia del dinero en los resultados electorales de todas las circunscripciones. La mayoría no mencionará la séptima circunscripción de Pensilvania ni ninguna otra a no ser de pasada o como una excepción a la regla general. Estos estudios siguen la pauta de Przeworski y Teune (1982): eliminar los nombres propios. Sin embargo, aunque no pretendan comprender el funcionamiento de ningún distrito en particular, no tendrían que pasar por alto — como, por desgracia, se hace de vez en cuando en esta tradición — que los hechos relativos a los diversos distritos que se incluyen en el análisis tienen que ser precisos.

Hay otras investigaciones que intentan decirnos algo acerca de un ejemplo concreto. Se centran en la Revolución francesa o en otro acontecimiento «importante» con la intención de explicar cómo o por qué sucedió. Los estudios de esta tradición serían pensables — y realmente carecerían de interés para la mayoría de sus lectores habituales — sin nombres propios. Un politólogo puede tratar con acierto las pautas de relación en el conjunto de las campañas para el Congreso estadounidense sin ocuparse de circunscripciones o candidatos concretos, pero imaginémoslo en el análisis que hizo Robert Caro (1983) de las elecciones tejanas de 1948 al Senado sin Lyndon Johnson y Coke Stevenson. Acontecimientos concretos como la Revolución francesa o las primarias demócratas de 1948 para el Senado en Texas pueden tener un interés intrínseco: suscitan nuestra curiosidad y, si fueran condiciones previas para que tuvieran lugar acontecimientos posteriores (como las guerras napoleónicas para la presidencia de Johnson), quizá fuera preciso conocerlas para comprender dichos acontecimientos. Además, tener un conocimiento general sobre las revoluciones, rebeliones o guerras civiles nos proporcionará información de valor incalculable para centrar mejor nuestro estudio sobre las causas concretas de la Revolución francesa.

Vamos a abordar estos asuntos analizando la «interpretación», que se presenta como alternativa a la inferencia científica (apartado 1.1), los conceptos de singularidad y complejidad del objeto de estudio (apartado 1.2) y el área general de los estudios de caso comparados (apartado 1.3).

el objetivo último, describir resulta determinante en toda explicación y es una actividad fundamental en sí misma. Lo que distingue los estudios científicos de otros tipos de investigación no es el enfrentamiento entre descripción y explicación, sino el hecho de que se hagan inferencias sistemáticas según procedimientos válidos. La inferencia, ya sea descriptiva o causal, cuantitativa o cualitativa, es el objetivo último de toda ciencia social de causalidad. La recogida sistemática de hechos es una empresa muy importante sin la que la ciencia no sería posible, pero no es ciencia en sí misma. El trabajo de archivo o el resumen de hechos históricos pueden conformar, si son correctos, una buena historia descriptiva, pero ninguna de estas actividades es suficiente para constituir ciencias sociales.

En este capítulo, distinguimos entre descripción — recogida de datos — e inferencia descriptiva. En el apartado 1 analizamos la relación entre los objetivos aparentemente contradictorios de la investigación académica: describir acontecimientos generales y conocer hechos particulares. De este modo, en el apartado 2, podemos explicar con más detalle el concepto de inferencia. En el resto del libro nuestra intención es presentar ideas a través de la expresión verbal y también mediante modelos de investigación algebraicos muy simples. En el apartado 3 entramos a considerar la naturaleza de esos modelos. Posteriormente, se analizan los relativos a la recogida de datos, al resumen de datos históricos y a la extracción de inferencias descriptivas en los apartados 4, 5 y 6, respectivamente. Para terminar, en el apartado 7, proporcionamos algunos criterios concretos para evaluar inferencias descriptivas.

1. El conocimiento general y los hechos particulares

El mundo que estudian los científicos sociales se compone de elementos particulares: votantes concretos o entidades gubernamentales específicas, así como ciudades, tribus, grupos, estados, provincias y naciones determinadas. La ciencia social de calidad pretende ir más allá de estos portmoneos para alcanzar un conocimiento más general. Sin embargo, la generalización no quita importancia al detalle. De hecho, el único objetivo que persigue ir de lo particular a lo general es aumentar nuestro conocimiento de ambas cosas. Los integrantes específicos del mundo social — o, más concretamente, los hechos concretos de esos integrantes — son la base en la que deben descansar las generalizaciones. Además, casi siempre aprendemos más de un caso específico cuando estudiamos conclusiones más generales. Si quisiéramos saber por qué ha dimitido el ministro de Asuntos Exteriores de Brasil, sería útil saber por qué lo han hecho otros ministros brasileños, por qué lo hacen los mismos cargos en otros países o por qué las personas, en general, dejan sus puestos en los gobiernos o incluso en otros ámbitos laborales. Cada una de estas razones nos ayudará a comprender diferentes

En las ciencias humanas hay historiadores y antropólogos que afirman que *lo único* que buscan son conocimientos concretos a través de lo que denominan «interpretación». Los interpretacionistas tratan de encontrar síntesis precisas de los acontecimientos y situar los acontecimientos que describen en un contexto inteligible dentro del que pueda explicarse el significado de las acciones. Como ha escrito Ferejohn (en Goldstein y Keohane, 1993, p. 228): «Queremos que las teorías de las ciencias sociales proporcionen explicaciones causales de los acontecimientos [...] [Y] que expliquen las razones de la acción social o su significado. No sólo queremos saber lo que llevó al agente a realizar un acto determinado, sino sus razones para hacerlo». Geertz (1973, p. 17) también escribe que «... no nos interesa quiénes son las propiedades que nos atraen».

Los investigadores que hacen hincapié en la «interpretación» pretenden arrojar luz sobre los aspectos del comportamiento humano uti-

lizando el concepto de *Verstehen* («compañía» que comprende el significado de las acciones e interacciones a partir del punto de vista de los que participan en ellas» [Eckstein, 1975, p. 81]). Los interpretacionistas quieren explicar las razones de las acciones intencionadas relacionadas con todos los conceptos y prácticas en los que están inmersas. También emplean normas de evaluación: «Las normas más evidentes son la coherencia y el campo de aplicación: una explicación interpretativa tiene que dar la máxima coherencia o inteligibilidad a un conjunto de prácticas sociales y, si lo que explica es un conjunto de prácticas determinado, tiene que ser compatible con otras prácticas o tradiciones de la sociedad» (Moon, 1975, p. 173).

Para los interpretacionistas, quizá la recomendación operativa más importante sea que los investigadores aprendan mucho sobre una cultura antes de elaborar preguntas de investigación, ya que sólo mediante una profunda inmersión cultural y un intenso conocimiento de la materia se pueden plantear preguntas correctas y formular hipótesis útiles. Dunner (1993), por ejemplo, estudió el comportamiento colectivo de los hombres de clase trabajadora negra en un bar antiojoso no segregado de Chicago. Al sumergirse en esta cultura local durante cuatro años, le sorprendieron ciertas cosas que antes no se le habían ocurrido. Por ejemplo, observó que, aunque estos hombres estaban muy en contra del Partido Republicano, su actitud ante muchos problemas era socialmente conservadora.

Algunos investigadores llevan aún más lejos el papel de la interpretación, llegando a indicar que, para las ciencias sociales, constituye un paradigma de investigación completamente diferente, «... no una ciencia experimental, llegando a indicar que, para las ciencias sociales, constituye un paradigma de investigación completamente diferente, «... no una ciencia experimental en busca de leyes sino otra de tipo interpretativo en busca de significado» (Geertz, 1973, p. 5). Sin embargo, para nosotros, la ciencia (tal como la hemos definido en el apartado 1.2 del capítulo 1) y la interpre-

tación no son empresas fundamentalmente diferentes que persigan objetivos distintos. Ambas dependen de la elaboración de cuidadosas descripciones, de una comprensión profunda del mundo, de que se planteen buenas preguntas, de la formulación de hipótesis falsables a partir de teorías más generales y de la recogida de las pruebas necesarias para evaluar tales hipótesis. La aportación primordial de la ciencia es un conjunto de procedimientos con los que se pueden dar *respuestas* a preguntas descriptivas y causales elaboradas apropiadamente.

Al insistir en la metodología inferencial no pretendemos menos cabar la importancia del proceso de formulación de preguntas fructíferas. Por el contrario, estamos de acuerdo con los interpretacionistas en que éste resulta crucial para entender en profundidad una cultura, antes de formular hipótesis o de diseñar un proyecto de investigación sistemático que encuentre una respuesta. Sólo queremos añadir que no se puede evaluar la veracidad de afirmaciones basadas en métodos como el de la observación participante si no se utiliza la lógica de la inferencia científica que nosotros describimos. Encontrar respuestas adecuadas para preguntas erróneas es un ejercicio pueril. La interpretación que parte del concepto de *Verstehen* suele ser fuente de penetrantes hipótesis. Por ejemplo, las detalladas observaciones que hizo Richard Fennel del Congreso estadounidense (Fennel, 1978) a través de lo que el denomina «empaparse y hurgar» han hecho grandes aportaciones al estudio de esta institución, principalmente porque han ayudado a plantear mejores preguntas de investigación. En su estudio de las regiones italianas, Putnam (1993, p. 12) afirma que para practicar el método de «empaparse y hurgar» «... es necesario que el investigador se sumerja en los pequeños detalles de una institución, con el fin de participar, como los que viven en ella a diario, en sus costumbres y prácticas, en sus éxitos y fracasos. Esta inmersión agudiza nuestras intuiciones y ofrece innumerables pistas sobre cómo se articula la institución y de qué manera se adapta a su medio». Cualquiera definición de ciencia que no deje lugar a las ideas relativas a la elaboración de hipótesis será tan trivial como una interpretación a la que no le preocupe descubrir la verdad.

Sin embargo, una vez que las hipótesis se han formulado, para demostrar que con correctas (dejando un margen a la incertidumbre) se precisan inferencias científicas válidas. Además, los procedimientos inferenciales de los científicos sociales interpretacionistas deben respetar las mismas reglas que los de otros investigadores cualitativos y cuantitativos. Esto significa que, aunque estamos de acuerdo en que las ciencias sociales de calidad necesitan interpretaciones penetrantes u otros métodos para generar buenas hipótesis, también hay que subrayar que la ciencia es esencial para que la interpretación sea precisa. Si pudiéramos entender el comportamiento humano únicamente a través del concepto de *Verstehen*, nunca podríamos falsar nuestras hipótesis descriptivas o aportar pruebas de ellas que no formaran parte de nuestra propia experiencia. En consecuencia, las conclusiones nun-

ca dejarían de ser hipótesis no comprobadas y las interpretaciones tendrían un carácter personal y no científico.

Uno de los mejores y más famosos ejemplos de la tradición interpretativa es el análisis que hizo Clifford Geertz del estudio de Gilbert Ryle sobre las diferencias entre un tic nervioso y un guiño. Geertz (1973, p. 6) escribe:

Pensemos en [...] dos chicos que contraen rápidamente los párpados de sus respectivos ojos derechos. En uno de ellos es un tic involuntario; en el otro, una señal de tipo conspícuo para un amigo. Los dos movimientos son, como tales movimientos, idénticos. Si se realizara una observación «fenomenológica» de estos movimientos aislados, como la haría una cámara fotográfica, no se podría distinguir entre el tic y el guiño o si, en realidad, ambos han sido una cosa u otra. Sin embargo, la diferencia que existe entre un tic nervioso y un guiño es enorme, aunque no sea fotografiable, como sabe cualquiera que haya tenido la desgracia de que confundieran la primera con la segunda. El que guiña se está comunicando de una forma precisa y determinada: (1) deliberadamente, (2) con alguien en concreto, (3) para transmitir un mensaje concreto, (4) según un código social establecido y (5) sin que lo note el resto de los presentes. Como señala Ryle, el que guiña hace dos cosas: contraer el párpado y guñar, mientras que el que tiene un tic sólo hace una: contraer el párpado. Cuando existe un código público por el que contraer los párpados se considera una señal de tipo conspícuo, hacer este movimiento es guñar.

Geertz está señalando un importante problema conceptual. Sin el concepto de «guiño», al que da sentido una teoría de la comunicación, el estudio más preciso de «las contracciones de los párpados en los seres humanos» carecería de sentido para los que estudian las relaciones sociales. En este ejemplo, la teoría, que surgió de meses de «empapar» y de minuciosos estudios culturales, resulta esencial para la pregunta básica de si las contracciones de los párpados pueden ser siquiera «tics» o «guiños». Este ejemplo indica claramente la gran importancia que tiene la interpretación, ya que proporciona nuevas formas de mirar el mundo (nuevos conceptos que considerar e hipótesis que evaluar). Sin una profunda inmersión en una determinada circunstancia quizá ni siquiera pudiéramos pensar en la que teorías hay que sopesar. En el ejemplo anterior, si no pensáramos en la diferencia entre tics y guiños, todo estaría perdido. Si la interpretación —o cualquier otra cosa— nos ayuda a abordar nuevos conceptos o hipótesis, entonces su utilidad resulta incuestionable, lo cual se ha confirmado una y otra vez no sólo en relación a este método sino a otras formas similares de portemonstrada comprensión cultural.

El investigador, una vez que ha hecho una distinción teórica relevante, como la que afecta al guiño y al tic, necesita evaluar la hipótesis que indica que están teniendo lugar guiños. Es en esta evaluación donde la lógica de la inferencia científica no tiene rival. En otras palabras, la mejor manera de determinar el significado de las contracciones de los párpados es la que se basa en los métodos sistemáticos descritos en este libro. Si lo crucial fuera

distinguir entre un tic nervioso y un guiño, resultaría fácil diseñar un procedimiento de investigación para hacerlo. Si, por ejemplo, creemos que ciertas contracciones de los párpados son guiños dotados de un significado político, será posible observar entonces otras circunstancias similares, ya que un elaborado mecanismo de señalización como éste (un «código público») probablemente se utilizará de nuevo una vez desarrollado. A la vista de esta posibilidad, podríamos registrar las ocasiones en las que se contrae el párpado de este actor, observar si el otro actor principal le mira en el momento adecuado y si el primero le responde. Podríamos incluso diseñar una serie de experimentos para ver si los individuos de esta cultura están acostumbrados a comunicarse de esta manera. Comprender la cultura, describir cuidadosamente el acontecimiento y estar muy familiarizados con situaciones similares nos ayudará a plantear las preguntas correctas y nos dará incluso más confianza en nuestras conclusiones. Pero sólo con los métodos de la inferencia científica podremos evaluar la hipótesis y determinar si es o no correcta.

La interpretación que da Geertz al guiño se expresa mejor como una hipótesis causal (que definiremos precisamente en el apartado 1 del capítulo 3): el hipotético efecto causal del guiño sobre el otro actor político es la respuesta de éste, dada la contracción del párpado menos su reacción a falta de otro movimiento (y de otros cambios). Si la contracción del párpado fuera un guiño, el efecto causal sería positivo; si sólo hubiera un tic nervioso, el efecto sería igual a cero. Si decidiéramos hacer una estimación de este efecto causal (y, por tanto, averiguar si se ha producido un guiño o un tic), para extraer la mejor inferencia respecto a la interpretación del comportamiento observado, habría que entender todos los problemas inferenciales que se analizan en profundidad en el resto del libro.

Si lo que interpretáramos como guiños fueran en realidad tics nerviosos involuntarios, nuestros intentos de extraer inferencias causales de las contracciones de los párpados, utilizando una teoría de la interacción social voltaria, fracasaría sistemáticamente: no podríamos generalizar, y lo sabríamos?

No es probable que un diseño de investigación que se proponga distinguir entre tics y guiños sea importante para la mayoría de las investigaciones de las ciencias políticas, pero el mismo problema metodológico aparece en gran parte de las áreas de trabajo de los politólogos. Con frecuencia se nos pide que tomamos las decisiones se envían mensajes unos a otros. Ante un determinado mensaje, cabe preguntarse si es una amenaza, un punto de negociación o una afirmación destinada a atraerse al público del propio país. Conocer las normas culturales, las convenciones de las comunicaciones internacionales y la historia de cada actor nos ayudará a hacer tal interpretación: del mismo modo que nos será útil observar de cerca otros rasgos secundarios de la comunicación. Pensemos también en el siguiente problema

cuantitativo: en los Estados Unidos los electores parecen estar enviando un mensaje al no acudir a las urnas, pero ¿qué significa esta escasa participación? ¿Acaso refleja una alienación respecto al sistema político, ¿representa un cálculo de los costes y beneficios de votar en el que los primeros son mayores que los segundos?, ¿muestra una decepción respecto a los últimos candidatos o campañas?, ¿podría ser consecuencia de un cambio en la edad mínima para votar?, ¿o una señal de que nada inquieta lo suficiente a los votantes como para llevarlos a las urnas? El hecho de que un ciudadano decida no votar, al igual que un guñón o un mensaje diplomático, puede significar muchas cosas. Un investigador inteligente siempre tendrá que esforzarse por plantear las preguntas adecuadas y después para diseñar cuidadosamente un estudio científico con el que averiguar qué significaba en realidad una determinada acción ambigua.

También nos gustaría ocuparnos brevemente de las afirmaciones extremas que hacen unos pocos partidarios de la interpretación cuando señalan que los objetivos de algunas investigaciones deberían ser sentimientos y significados sin consecuencias observables. En justicia, esta caracterización no afecta más que a una reducida minoría de los investigadores de esta tendencia, pero sus afirmaciones son lo suficientemente efusivas como para que valga la pena darles una contestación explícita. Al igual que las opiniones excesivamente entusiastas de los primeros positivistas, que mantenían la insostenible postura de que no había sitio en la investigación científica para los conceptos no observables, estos argumentos resultan inapropiados para la investigación empírica. Psathas (1968, p. 510), por ejemplo, señala:

cualesquier comportamiento que sólo se centre en lo evidente y en lo que se manifiesta en actos concretos, directamente observables, es, como mínimo, ingenuo. Por lo tanto, el desafío para el científico que pretende comprender la realidad social es captar el significado que, para el actor, tiene su propia acción.

Puede que Psathas tenga razón al señalar que los científicos sociales que sólo se centran en comportamientos evidentes u *observables* se están perdiendo mucho, pero ¿cómo habríamos de saber esto si no podemos ver más? Por ejemplo, si dos teorías sobre el concepto que uno tiene de sí mismo tienen idénticas manifestaciones observables, *ningún* observador tendrá información suficiente para distinguirlos. Esto es así para cualquier observador, independientemente de lo inteligente o sensible a factores culturales que sea, de sus habilidades para la interpretación, de lo bien que «clasifique» sus propias suposiciones o del empeño que ponga en su empresa. La interpretación, el presentamiento, la descripción minuciosa, la observación participante y la que no lo es, la entrevista en profundidad, la empatía, la cuantificación y el análisis estadístico, al igual que los demás procedimientos y métodos, no sirven para distinguir entre dos teorías que no tengan conse-

1.2 «SingULARIDAD», COMPLEJIDAD Y SIMPLIFICACIÓN

Algunos investigadores cualitativos rechazarían la idea de que un conocimiento general sea necesario o útil (ni siquiera posible) como punto de partida para comprender un determinado acontecimiento. Lo que sostienen es que los hechos o unidades que estudian son «mínicos», y, en cierto sentido, tienen razón. Sólo hubo una Revolución francesa, sólo hay una Tailandia y nadie que haya leído las biografías pertinentes o que viviera los años sesenta puede cuestionar el hecho de que sólo hubo un Lyndon B. Johnson. Pero aún van más lejos. La explicación, según su punto de vista, se limita a ese único acontecimiento o unidad: no al porqué de las revoluciones sino al de la Revolución francesa; no a por qué en ocasiones la democratización parece languidecer, sino a por qué ocurre así en Tailandia; no a por qué ganan los candidatos sino a por qué ganó Johnson en 1948 y 1964. Los investigadores de esta tradición creen que perderían su capacidad de explicar lo específico si intentarían ocuparse de lo general: de revoluciones, democratización o primarias del Senado estadounidense.

Sin embargo, el término «singularidad» es engañoso. La Revolución francesa, Tailandia y Johnson son realmente mínicos. En cierto sentido, lo son todos los fenómenos, todos los acontecimientos; pero también lo fueron las elecciones al Congreso en la séptima circunscripción de Pensilvania en 1988, así como la opción que tomó cada uno de los millones de electores que votaron en los comicios presidenciales de ese mismo año. Considerado de forma global, cada aspecto de la realidad social es infinitamente complejo y se relaciona de alguna manera con acontecimientos naturales y sociológicos anteriores. Por lo tanto, la singularidad de las cosas forma parte de la condición humana: no distingue entre las situaciones que son susceptibles de generalización científica y aquellas en las que no es posible generalizar. De hecho, como señalamos al analizar las teorías sobre la extinción de los dinosaurios en el capítulo 1, incluso los acontecimientos únicos pueden estudiarse de forma científica si se presta atención a las con-

ción a otra, aunque es probable que ambos estén presentes. Además, más que opuestos son complementarios. De hecho, puede que la mejor manera de entender un determinado acontecimiento sea utilizar también los métodos de la inferencia científica para estudiar pautas sistemáticas en acontecimientos paralelos similares.

1.3 Estudios de caso comparados

En gran medida, lo que hacen los politólogos es describir de forma sistemática acontecimientos que son importantes desde el punto de vista político. A la gente le interesa la caída de la Unión Soviética, la reacción popular en los países árabes ante la guerra autorizada por la ONU para expulsar a Irak de Kuwait y los resultados de las últimas elecciones al Congreso de los Estados Unidos. Consta en que los politólogos describan con un conocimiento más profundo que el de las informaciones periodísticas la relación que existe entre estos y otros acontecimientos relevantes, ya sean contemporáneos o históricos. Al describir los acontecimientos tenemos que ser lo más precisos y sistemáticos que sea posible. Esto supone que si logramos encontrar medidas cuantitativas válidas para lo que queremos saber, tenemos que utilizarlas: ¿qué proporción de los periódicos soviéticos crítica las políticas gubernamentales? ¿qué ponen de manifiesto las encuestas de opinión llevadas a cabo en Jordania y Egipto sobre la actitud de estos países hacia la Guerra del Golfo? ¿qué porcentaje de congresistas estadounidenses fue reelegido?

Aunque cuantificar produzca precisión, no favorece necesariamente la exactitud, ya que inventar índices cuantitativos que no se relacionen directamente con los conceptos o acontecimientos que pretendemos medir puede conducir a graves errores de medida y a problemas en las inferencias causales (véase el apartado 1 del capítulo 5). Del mismo modo, existen métodos más y menos precisos para describir acontecimientos no cuantificables. Hay disciplinados cuantitativistas que intentan analizar con cuidado las constituciones y leyes, en vez de informar únicamente de lo que los observadores dicen sobre ellas. Los investigadores, al realizar estudios de caso sobre políticas gubernamentales, plantean a los entrevistados preguntas agudas y bien definidas, cuyas respuestas serán relativamente claras, y sondean de forma sistemática todo comentario inesperado que pueda sugerir hipótesis relevantes. Los estudios de caso son esenciales para la descripción y, por tanto, para las ciencias sociales. No tiene sentido intentar explicar lo que antes no se ha descrito con un razonable grado de precisión. Describir con perspicacia acontecimientos complejos no es algo trivial. En áreas como la política comparada o las relaciones internacionales la labor de descripción es especialmente importante porque todavía hay muchas cosas que necesitamos saber, nuestra capacidad explicativa es débil y una

secuencias observables de las teorías que se han desarrollado para explicarlos.

Lo que realmente plantea la singularidad es el problema de la complejidad. La cuestión no es si los acontecimientos son de por sí únicos, sino el hecho de que sea posible o no extraer de un amasijo de acontecimientos las características principales de la realidad social que queremos comprender.

Una de las primeras, y más difíciles, tareas de la investigación social es la de simplificar, la cual nos expone a ser criticados por practicarla en exceso y por omitir aspectos significativos de una situación. No obstante, ningún investigador puede evitar la simplificación, que ha sido parte integrante de todas las obras académicas conocidas —cuantitativas y cualitativas, antropológicas y económicas, de las ciencias sociales y de las naturales o físicas— y que probablemente siempre lo será. Ni siquiera la descripción más exhaustiva que hagan los mejores intérpretes de la cultura desde el conocimiento profundo de un determinado contexto podrá dejar de simplificar, objetualizar y reducir de forma drástica la realidad observada. De hecho, la diferencia que hay entre el grado de complejidad del mundo y el de la descripción más densa sigue siendo mucho más grande que la que existe entre dicha descripción y el análisis cuantitativo o formal más abstracto. Ninguna descripción —cuquiera que sea su densidad— y ninguna explicación —independientemente del número de factores explicativos que contemple— pueden llegar a captar, siquiera levemente, toda la «floreciente y agitada» realidad del mundo. No queda más remedio que simplificar. La simplificación sistemática es un paso crucial hacia el conocimiento útil. Como ha señalado un historiador económico, si la insistencia en la singularidad «se lleva hasta el extremo de hacer caso omiso de todas las regularidades, se negará a las ciencias sociales la posibilidad misma de que existan y los historiadores no serán más que baladistas sin objetivos» (Jones, 1981, p. 160).

El analista tiene que simplificar sus descripciones, siempre que sea posible, sólo después de comprender la riqueza histórica y cultural. Puede que los científicos sociales, para establecer diferencias, utilicen únicamente unos pocos elementos de la historia de ciertos grupos de acontecimientos. No obstante, un conocimiento profundo y no estructurado del contexto histórico y cultural de los fenómenos que quieren tratar de forma sintética y científica suele ser indispensable para evitar simplificaciones claramente equivocadas. Poco se podrá confiar en las generalizaciones de un científico-social que estudiara las revoluciones o las elecciones al Senado estadounidense si no supiera mucho sobre la Revolución francesa o las elecciones de 1948 en Tejas y no le importara su ignorancia.

En resumen, creemos que, siempre que sea posible, la investigación social tiene que ser general y específica al mismo tiempo: debe informarnos sobre los tipos de acontecimientos y también sobre hechos específicos en lugares concretos. Queremos aislarlos del tiempo sin perder el vínculo con él, y la insistencia en uno de estos objetivos puede variar de una investigación

buena descripción depende en parte de una buena explicación. Algunas de las razones que explican esta necesidad de conocer y la debilidad de nuestra capacidad explicativa son las mismas: en las relaciones internacionales, por ejemplo, las pautas de poder, de alineamiento y de interdependencia han cambiado mucho últimamente, por lo que, al alterarse el contexto sistémico en el que tienen lugar las relaciones que se observan entre los estados, cada vez resulta más necesario describir correctamente las nuevas situaciones. La causalidad suele ser difícil de determinar porque tanto los estados como otros actores pretenden prever y contrarrestar las acciones de los demás y, a la hora de explicar el comportamiento de dichos estados, las expectativas pueden tener un papel tan importante como las acciones que se observan. Pretender explicar algún aspecto de las relaciones internacionales presuponiendo que no hay interacciones estratégicas ni previsión de reacciones será mucho menos útil que describir con cuidado acontecimientos que tenemos razones para creer que son importantes y que están interconectados. Una buena descripción es mejor que una mala explicación.

Una de las ventajas del estudio de caso en profundidad, que suele pasar-se por alto, es que desarrollar una buena hipótesis causal *complementa* una buena descripción, no rivaliza con ella. Encuadrar un estudio de este tipo en una pregunta explicativa puede producir una descripción más centrada y relevante, aunque al final el estudio no logre extraer una sola inferencia causal válida.

Nosotros afirmamos que los estudios de caso comparados pueden generar inferencias causales válidas cuando se utilizan los procedimientos descritos en el resto del libro, aunque, tal como se practican en la actualidad, no suelen cumplir las normas de una inferencia válida (que explicaremos en el capítulo 3). De hecho, gran parte de lo que los científicos sociales con enfoque histórico o interpretativo denominan trabajo «explicativo» sigue siendo fundamentalmente descriptivo porque no cumple esas normas de aplicación universal. Desde esta perspectiva, resulta fundamental lo que señalan muchos investigadores cuando aconsejan que, al describir o explicar, los estudios de caso comparados sean más sistemáticos.

Alexander George, por ejemplo, recomienda un método de «comparación estructurada y centrada» que hace hincapié en la disciplina cuando se trata de recoger datos (George y McKeown, 1985; véase también Verba, 1967). George y sus colaboradores subrayan la necesidad de que haya una recogida sistemática de la misma información —de las mismas variables— en unidades seleccionadas cuidadosamente. También insisten en que, para que finalmente sea posible la inferencia causal, son necesarias directrices teóricas —plantearse preguntas explicativas cuidadosamente elaboradas— con el fin de describir de forma sistemática.

El método de la comparación estructurada y centrada es una forma sistemática de emplear lo que George y McKeown denominan el «procedimiento de la coherencia». A través de este sistema, el investigador «define y es-

tandanza las necesidades de los datos de los estudios de caso [...] formular preguntas relevantes desde el punto de vista teórico con el fin de dar pautas para el análisis de cada caso» (George y McKeown, 1985, p. 41). El problema que señalan George y McKeown está bien traido: «Para hacer una comparación controlada mediante un tamaño muestral pequeño hay que seguir un procedimiento sistemático de recogida de datos». Esa «comparación estructurada y centrada» precisa datos de las mismas variables en diferentes unidades. Por lo tanto, no es un método diferente al que subrayamos aquí, ya que constituye una forma de sistematizar la información en estudios de caso descriptivos de manera que sea factible utilizarla para hacer inferencias descriptivas o causales. Muchos de los valiosos consejos que, como éste, se dan para hacer estudios de caso comparados son elementales, pero se suelen pasar por alto.

2. La inferencia: el fin científico de la recogida de datos

La inferencia es un proceso en el que se utilizan hechos que conocemos para aprender sobre los que desconocemos. Los hechos conocidos se someten a las preguntas, teorías e hipótesis de nuestra investigación y conforman nuestros datos y observaciones de tipo cuantitativo o cualitativo.

Cuando buscamos un conocimiento general, en sí mismo o para comprender mejor determinados hechos, es preciso evitar de alguna manera que nos apabulle el enorme mare magnum de observaciones potenciales y reales que hay sobre el mundo. Por fortuna, la solución a este problema reside precisamente en la búsqueda del conocimiento general. Esto quiere decir que la mejor manera científica de organizar los hechos es hacer que sean consecuencias observables de alguna teoría o hipótesis. La simplificación científica entraña la provechosa elección de una teoría (o hipótesis) con el fin de evaluarla; después, la teoría nos lleva a seleccionar los hechos que son consecuencia de ella. Organizarlos en función de las consecuencias observables de una determinada teoría tiene importantes y beneficiosos resultados para el diseño y la realización de la investigación. En primer lugar, con este criterio de selección de los hechos podemos darnos cuenta rápidamente de que si existen más observaciones de las consecuencias de una teoría no podemos más que beneficiarnos de ello al evaluar la teoría en cuestión. Como más información de este tipo no puede ser perjudicial, esos datos nunca se desechan, y así mejora el proceso investigador.

En segundo lugar, no hay por qué poseer una teoría completa antes de recoger los datos, ni tampoco nuestra teoría tiene por qué mantenerse intactable durante el proceso. La teoría y los datos interactúan. Al igual que ocurre con la gallina y el huevo, siempre es necesaria alguna teoría antes de recoger los datos y también se precisan algunos datos antes de teorizar. Los manuales de investigación nos dicen que utilizemos los datos para compro-

que, a su vez, se compongan de «atributos», «variables» o «parámetros». La clase podría ser de «votantes»; las unidades, una muestra de esos «votantes» en diversas circunscripciones para las elecciones al Congreso estadounidense; y los atributos o variables podrían ser la renta, la identificación partidista o cualquier otra cosa que represente una consecuencia observable de la teoría que se evalúa. La clase también podría ser un determinado colectivo, como las comunidades o países; las unidades podrían ser una selección de éstos; y los atributos o variables podrían ser el tamaño de los mismos, el tipo de gobierno, su situación económica, composición étnica o cualquier otro elemento cuantificable y de interés para el investigador. Estos conceptos, así como elaboraciones teóricas como las tipologías, estructuras y todo tipo de clasificaciones, son útiles mecanismos temporales cuando estamos recogiendo datos, pero no plantean una hipótesis clara que evaluar. En general, no animamos a los investigadores a que organicen sus datos de esta manera, ya que sólo necesitamos el concepto organizador inherente a nuestra teoría. Es decir, si las observaciones no son consecuencia de nuestra teoría, resultan irrelevantes. Si son irrelevantes o no observables, hay que prescindir de ellas, pero si son relevantes, tenemos que utilizarlas. Nuestros datos no tienen por qué estar en el mismo nivel de análisis. Hay datos desagregados u observaciones sobre un período de tiempo diferente, o incluso sobre otra parte del mundo, que pueden proporcionar más consecuencias observables de una teoría. Quizá no nos interesen en absoluto estas consecuencias secundarias, pero si concuerdan con la teoría, tal como se predice, nos ayudarán a otorgarle más confianza a sus capacidades y aplicabilidad. Nuestros datos tampoco necesitan ser «simétricos»: siempre que todos sean consecuencias observables de nuestra teoría, podremos utilizar un detallado estudio de una provincia, una comparación de dos países, entrevistas personales con miembros de la administración pertenecientes a un único sector de políticas e, incluso, algún elemento cuantitativo. En este proceso vamos de lo particular a lo general, ya que caracterizar ciertas unidades a partir de rasgos comunes es un proceso generalizador. La consecuencia es que aprendemos mucho más acerca de las teorías generales y de los hechos concretos.

En general, lo que queremos es concentrar en nuestra hipótesis tanta información como sea posible, lo cual puede suponer la realización de más estudios de caso; pero esto suele ser demasiado difícil, requiere mucho tiempo o es caro. Es evidente que no debemos considerar información irrelevante. Por ejemplo, tratar el número de escafos conservadores en la Cámara de los Comunes británica como una variable mensual en vez de como algo que cambia en las elecciones nacionales aumentaría considerablemente el número de observaciones, pero no tendría sentido porque no aportaría mucha más información. Por otra parte, desagregar los resultados de las elecciones presidenciales estadounidenses por estados, e incluso por condados, aumenta el número de casos y también la cantidad de información que se aporta al problema.

bar nuestras teorías, pero aprender de ellos puede ser un objetivo tan importante como evaluar teorías e hipótesis previas. Para que se produzca este aprendizaje hay que reorganizar los datos como consecuencias observables de la nueva teoría. Esta reorganización es muy habitual al principio de muchas investigaciones, generalmente después de haber recogido algunos datos preliminares; después, prosigue la recogida de datos para evaluar la nueva teoría. Siempre se debe intentar seguir recogiendo datos, incluso después de la reorganización, para contrastar la nueva teoría y evitar así utilizar el mismo material para evaluar ésta y aquella de la que partió.

En tercer lugar, insistir en que se recojan datos que sean consecuencias observables de una hipótesis deja mucho más claro el territorio común que comparan los estilos de investigación cuantitativo y cualitativo. De hecho, una vez que dejamos de pensar en los casos, unidades o registros de la forma generalmente estrecha, e incluso ingenua, que es habitual, nos damos cuenta de que la mayoría de los estudios cualitativos podrían aportar una gran cantidad de consecuencias observables a las teorías que estamos evaluando, sin embargo, muchas de ellas pueden pasarle desapercibidas al investigador. En consecuencia, organizar los datos mediante una lista de las observaciones específicas de una teoría ayuda a poner de manifiesto el fin científico esencial de gran parte de la investigación cualitativa. En cierto sentido, estamos pidiéndole al investigador que estudia un acontecimiento concreto —quizá una decisión gubernamental específica— que se pregunte: «Si mi explicación de por qué la decisión resultó de una determinada manera es correcta, ¿que más podría observar en el mundo real?». Se pueden encontrar consecuencias observables adicionales en otras decisiones, pero también en otros aspectos de la que se estudia, por ejemplo: cuando se toma, cómo, o de qué manera se justificó. La máxima crucial que ha de guiar tanto la creación de una teoría como la recogida de datos es la búsqueda de más consecuencias observables de esa teoría.

Cada vez que desarrollamos una nueva teoría o hipótesis resulta provechoso hacer una lista de todas las consecuencias de la misma que, en principio, pudieran observarse. La lista, que podría después limitarse a los ejemplos, podría ser sobre los que se han recogido datos —o pudiera fácilmente hacerse—, constituye a continuación la guía básica operativa para el proyecto de estudio. Si recoger un dato más nos ayuda a considerar otra manera de evaluar la teoría, merece la pena hacerlo (siempre sujetos a los condicionantes de tiempo, dinero y esfuerzo). Cuando una entrevista u otro tipo de observación puede ser de interés, pero no es una consecuencia observable potencial de esta teoría (o de otra que sea relevante), entonces tiene que ser evidente que no nos ayudará a evaluarla.

Dentro del proceso de simplificación que se desarrolla al organizar nuestros datos como consecuencias observables de la teoría, necesitamos sistematizar dichos datos. Podemos pensar en convertir los fenómenos brutos del mundo real en «clases» que estén compuestas por «unidades» o en «casos»

?Podemos evaluar un modelo sin conocer qué características del objeto queremos evaluar? Esta clara que no. Por ejemplo, podríamos pensar que un modelo que se ocupa de la cantidad de desperdicios de un avión no sería de gran utilidad. Sin embargo, si que sería muy importante para fines educativos o para realizar pruebas de túneles aerodinámicos. Como incluso el polvo de las moquetas puede hacer que un avión pese más y que, por tanto, haya que pagar más combustible, los modelos de este tipo son importantes para la industria aeronáutica y se han realizado (ahorrandose así millones de dólares).

Hay versiones más o menos restrictivas de todos los modelos. Los más restrictivos son más claros, concisos y abstractos, pero también menos realistas (a menos que el mundo sea realmente conciso). Los poco restrictivos son detallados, contextuales y más realistas, pero también menos claros y más difíciles de calibrar con precisión (véase King, 1989: apartado 5 del capítulo 2). En qué punto de esta gradación decidimos construir un modelo depende de nuestro objetivo y de la complejidad del problema que estudiamos.

Existen modelos físicos, gráficos, verbales o algebraicos. Por ejemplo, la descripción cualitativa de los sistemas judiciales europeos en un libro que se ocupa de ese tema constituye un modelo del mismo. Independientemente de lo profunda que sea la descripción o del talento del autor, el contenido del libro siempre será una abstracción o una simplificación de los sistemas jurídicos reales. Como para comprender se necesita resumir, lo que indica que un libro es bueno es tanto lo que se deja fuera de él como lo que se incluye.

Mientras que los investigadores cualitativos a menudo emplean modelos verbales, nosotros utilizaremos los algebraicos en el análisis que sigue, con el fin de estudiar los de tipo verbal y mejorarlos. Al igual que ocurre con los juegos aerodinámicos y con los libros que se dedican en su totalidad a estudiar la Revolución francesa, nuestros modelos algebraicos cualitativos no tienen que confundirse con la propia investigación cualitativa. Su único propósito es señalar con especial claridad los problemas que hay que evitar y las oportunidades que pueden aprovecharse. Además, con frecuencia ocurre que nos ayudan a descubrir ideas que, de otro modo, no se nos habrían ocurrido.

Partimos de la base de que los lectores no han tenido experiencia previa con modelos algebraicos, aunque a los que conozcan los estadísticos algunos de los siguientes les resultarán familiares. Sin embargo, la lógica inferencial de estos modelos se aplica tanto a la investigación cualitativa como a la cualitativa. El hecho de que, probablemente, los cuantitativistas estén más familiarizados con nuestra terminología no significa que se les vaya a dar mejor aplicar la lógica de la inferencia científica. Además, estos modelos no se adaptan más directamente a la investigación cuantitativa que a la cualitativa, en ambos casos representan abstracciones útiles de la

Estos datos desagregados pueden parecer irrelevantes, ya que el objetivo es conocer las causas de la victoria de un determinado candidato en la carrera presidencial, que es una cuestión fundamentalmente agregada. Sin embargo, la mayoría de las explicaciones que se dan al resultado de las elecciones presidenciales tienen diferentes consecuencias observables en cada una de las unidades desagregadas. Si, por ejemplo, pronosticamos el resultado de esta votación en función de variables económicas como el índice de paro, el uso de este indicador en cada estado aportará muchas más observaciones sobre las consecuencias de nuestra teoría que el índice agregado del conjunto del país. Verificar que la teoría tiene sentido en otras situaciones —aunque éstas no tengan un interés directo— aumenta la confianza en que la teoría es correcta y en que explica adecuadamente la única consecuencia observable que nos interesa.

3. Modelos formales de investigación cualitativa

Un modelo es una simplificación de algún aspecto del mundo y una aproximación al mismo. Los modelos nunca son, netamente, «verdaderos» o «falsos», aunque los buenos sólo recopilan las características «correctas» de la realidad que representan.

Pensemos por ejemplo en un juguete de plástico y pegamento que sea un modelo de quince centímetros de un avión. Esta es una pequeña reproducción del auténtico aparato, que no puede volar ni contiene nada y cuyos componentes tampoco se mueven. Nadie la confundiría con el avión real; preguntaría si cada aspecto del modelo es verdadero es como plantearse si la modelo que posó para la *Cioconda* de Leonardo da Vinci tenía realmente una sonrisa tan seductora. Aunque la tuviera, no esperaríamos que el cuadro de Leonardo sea una representación exacta de nadie, ya sea una modelo real o la Virgen María, del mismo modo que no esperaríamos que el modelo de un avión refleje todas sus características reales. Sin embargo, nos gustaría saber si esa reproducción recopila las características correctas de un avión para un problema concreto. Si queremos enseñarle a un niño cómo es realmente un avión, este modelo podría ser adecuado. Si está hecho a escala, también podría ser útil para que los diseñadores aeronáuticos hicieran pruebas de túneles aerodinámicos. La característica principal de un avión real que se aprecia en este modelo es la forma, y, para ciertos propósitos, éste es realmente uno de los rasgos apropiados. Evidentemente, este modelo no refleja multitud de detalles de un avión, entre ellos el tamaño, el color, la sensación de estar en una aeronave, la solidez de sus diversos componentes, el número de plazas a bordo, la potencia de sus motores, el tejido de los cojines de los asientos, ni tampoco otros muchos sistemas esenciales como el eléctrico, el de aire o el de conducción de agua. Si queremos entender estos aspectos del avión, necesitaremos un conjunto de modelos completamente diferente.

investigación a la que se aplican. Para facilitar su presentación, todos los modelos algebraicos van precedidos de descripciones verbales a las que siguen recuadros con una notación matemática normalizada. Aunque no lo recomendamos, los recuadros pueden saltarse sin que se pierda el hilo argumental.

4. Un modelo formal sobre la recogida de datos

Antes de formalizar nuestra presentación de las inferencias descriptiva y causal —los dos objetivos principales de la investigación social—, desarrollaremos un modelo para la recogida y resumen de datos que es bastante sencillo, pero que resulta eficaz para analizar los problemas inferenciales. Nuestro modelo algebraico no es tan formal como los estadísticos, aunque sí clarifica nuestras ideas y facilita su transmisión. Con la expresión *recogida de datos* hacemos alusión a una gran variedad de métodos, entre los que se incluye la observación general y la participante, las entrevistas en profundidad, las encuestas a gran escala, la historia basada en fuentes secundarias, los experimentos aleatorios, la etnografía, el análisis de contenido y cualquier otro método que sirva para reunir datos fiables. *La regla más importante para toda recogida de datos es dejar claro como se han creado y de qué manera hemos accedido a ellos.* Toda información que recojamos tiene que concretar las consecuencias observables de nuestra teoría. Desarrollar una nueva pregunta de investigación puede ayudar-nos, pero no servirá para responder a la primera si no es una consecuencia observable de ella.

Hacemos modelos con los datos por medio de *variables, unidades y observaciones*. Un ejemplo sencillo es la renta anual de cuatro personas diferentes. Los datos pueden representarse simplemente mediante cuatro cantidades: 9.000, 22.000, 21.000 y 54.292 dólares. En el caso más general, podríamos etiquetar la renta de las cuatro personas (numeradas del 1 al 4) como y_1, y_2, y_3 e y_4 . Una variable codificada para dos entrevistados no «kinturagtas» podría utilizar los valores «participativo», «cooperador» o «intermedios», y se etiquetaría como y_1 e y_2 . En estos ejemplos la *variable* es y , las *unidades* son los individuos, y las *observaciones*, los valores de las variables en cada unidad (renta en dólares o grado de cooperación). El símbolo y se denomina variable porque su valor cambia en cada unidad y, en general, puede representar cualquier cosa cuyo valor se vaya alterando en un conjunto de unidades. Estas, dado que se puede recabar información en un período de tiempo o en diferentes secciones de un área, pueden ser personas, países, organizaciones, años, elecciones o décadas y, con frecuencia, alguna combinación de estas u otras unidades. Las observaciones pueden ser numéricas, verbales, visuales o cualquier otro tipo de datos empíricos.

Supongamos, por ejemplo, que nos interesa estudiar las organizaciones internacionales desde 1945. Antes de recoger los datos hay que decidir qué resultados se quiere explicar. Podríamos intentar comprender la influencia en 1990 del factor tamaño sobre la actividad de las organizaciones internacionales (que se clasificaría por su campo de acción o por el propio tipo de organización), los cambios en el tamaño agregado de la actividad de dichas organizaciones desde 1945 o las transformaciones ocurridas en su actividad desde esa fecha que hubieran tenido que ver con la distribución de ese factor. Entre las variables que miden la actividad de las organizaciones internacionales podríamos incluir el número de países que pertenecen a ellas en un momento dado, la cantidad de tareas que desempeñan o la magnitud de sus presupuestos y de su personal. En estos ejemplos, las unidades de análisis serían las organizaciones internacionales, sus campos de acción, el número de países miembros y períodos de tiempo como años, quinquenios o décadas. En la fase de recogida de datos no se aplica ninguna regla para determinar qué variables se recogen, cuántas unidades debe haber, si éstas deben ser más numerosas que las variables o cual es el mejor método para medir las últimas. El único criterio es que, en cada caso, nosotros pensemos que lo que hacemos es importante. Cuando tenemos una idea más clara de cómo van a utilizarse los datos, la regla es intentar encontrar tantas consecuencias observables de una teoría como sea posible. Como subrayamos en el capítulo 1, la investigación empírica puede utilizarse tanto para evaluar hipótesis previas como para plantear otras que no se habían barajado anteriormente, pero si sirve para el segundo propósito, hay que recoger nuevos datos para evaluar las nuevas hipótesis.

A partir de nuestro análisis debe quedar claro que en la mayoría de las obras que se denominan «estudios de caso» se miden numerosas variables de muchos tipos diferentes de unidades. Aunque estas investigaciones no suelen utilizar más de un puñado de casos, el número total de observaciones es, generalmente, inmenso. Por lo tanto, es esencial distinguir entre números de casos y de observaciones. El primero puede tener cierto interés para ciertos fines, pero el último es importante para evaluar la cantidad de información que un estudio aporta a la resolución de una pregunta teórica. Por consiguiente, reservamos la n habitual para referirnos únicamente al número de observaciones y no al de casos. Sólo de forma ocasional, como cuando las observaciones individuales sean parcialmente dependientes, distinguiremos entre información y número de observaciones. El uso de esta última expresión procede del muestreo de encuestas, en el que n es el número de personas que hay que entrevistar, sin embargo, nosotros la utilizaremos de forma más general. En realidad, nuestra definición de «observación» coincide exactamente con lo que Harry Eckstein (1975, p. 85) denomina «casos». Como señala este autor: «Un estudio de seis elecciones generales en Gran Bretaña puede ser, pero no tiene por qué, un estudio en el que $n = 1$. Podría ser otro en el que $n = 6$ o también $n = 120.000.000$ ».

vo es exponer características apropiadas de éstos mediante un formato útil? Por ejemplo, la *media muestral*, o promedio, es un estadístico:

$$\bar{y} = \frac{1}{n} (y_1 + y_2 + \dots + y_n)$$

en el que $\sum_{i=1}^n y_i$ es una manera cómoda de escribir $y_1 + y_2 + y_3 + \dots + y_n$. Otro estadístico es el *máximo muestral*, que se etiqueta como y_{\max} :

$$y_{\max} = \text{Máximo } (y_1, y_2, \dots, y_n) \quad (2.1)$$

La media muestral de las cuatro cifras de renta del ejemplo del apartado 4 (9.000, 22.000, 21.000 y 54.292 dólares) es 26.573 dólares, mientras que el máximo muestral es 54.292. Podemos resumir los datos originales que contienen cuatro cifras mediante las dos cantidades que representen, respectivamente, la media y el máximo muestrales. También podemos calcular otras características de la muestra, como el mínimo, la mediana, la moda o la varianza.

Cada resumen de este modelo reduce todos los datos (cuatro cifras en este ejemplo simple o lo que conocemos de algún aspecto de la historia europea en el anterior) a un único número. Comunicarse mediante resúmenes suele ser más fácil y tiene más sentido para el lector que utilizar todos los datos originales. Evidentemente, si en nuestro conjunto de datos sólo hubiera cuatro números, no tendría mucho sentido utilizar cinco resúmenes diferentes y sería más sencillo presentar las cuatro cantidades originales. Interpretar un estadístico suele ser más fácil que comprender todo un conjunto de datos, pero siempre se pierde información al describir una gran cantidad de números utilizando unos pocos.

¿Qué normas rigen el resumen de los pormenores históricos? La primera es que los resúmenes deben centrarse en los resultados que queremos describir o explicar. Si lo que nos interesa es el crecimiento de la organización internacional media, no sería muy inteligente centrarse en las Naciones Unidas, pero sí lo que nos importa es cómo influye el mayor o menor tamaño en la distribución de dichas organizaciones, las Naciones Unidas serían seguramente una de las unidades en las que tendríamos que centrarnos. La ONU no es una organización representativa, pero sí es importante. Para estudiar la típica organización internacional desde el punto de vista estadístico examinaremos valores medios (de presupuestos, comités, número de miembros, etc.), pero para comprender el abanico de actividades tendramos que examinar la varianza. El segundo precepto, igualmente evidente, es que un resumen tiene que simplificar la información de que disponemos. En términos cuantitativos, esto significa que siempre hay que utilizar menos estadísticos sintetizadores que unidades hay en los datos originales, porque, de no ser así, podría ser más fácil presentar los datos originales sin

Todo depende de si el objeto de estudio son los sistemas electorales, las elecciones o los votantes». La «ambigüedad sobre lo que constituye un "individuo" (de ahí el "caso") sólo puede evitarse si nos ocupamos de las medidas de entidades concretas, y no de las propias entidades. De este modo, un "caso" puede definirse técnicamente como un fenómeno del que sólo señalamos e interpretamos una única medida en cada variable pertinente». La única diferencia con nuestro uso es que, desde que Eckstein escribió su artículo, los investigadores han continuado utilizando la palabra «caso» para referirse a un estudio de caso completo, que todavía tiene una definición bastante imprecisa. Por lo tanto, siempre que sea posible utilizaremos la palabra «caso» como lo hacen la mayoría de los autores y reservaremos el término «observación» para hacer referencia a las medidas de una o varias variables en una sola unidad.

En el resto del capítulo queremos mostrar de qué manera conceptos como el de variable y unidad pueden hacer que abordemos de forma más clara el diseño de una investigación, incluso cuando no sea apropiado utilizar medidas cuantitativas para resumir la información de que dispongamos. La cuestión que planteamos es: ¿cómo podemos hacer inferencias descriptivas sobre «la historia tal como fue» sin perdernos en un mar de datos irrelevantes? Dicho de otro modo, ¿cómo separar lo esencial de lo efímero?

5. El resumen de los pormenores históricos

Después de recoger los datos, el primer paso de cualquier análisis es resumirlos. Los resúmenes pueden describir una gran cantidad de datos pero no están directamente relacionados con la inferencia. Como, en realidad, lo que nos interesa es generalizar y explicar, un resumen de los datos que van a explicarse suele ser un buen punto de partida, pero no constituye un objetivo suficiente para la investigación en las ciencias sociales.

Resumir es necesario, ya que, como nunca podemos decir «todo lo que sabemos» sobre un conjunto de acontecimientos, no tendría sentido intentar hacerlo. Los buenos historiólogos comprenden cuáles son los datos cruciales y, por lo tanto, en vez de perderse en digresiones, realizan estudios que subrayan lo esencial. Para entender la historia europea de los primeros quince años del siglo XIX, quizá necesitemos comprender ciertos principios de estrategia militar tal como lo hizo Napoleón e, incluso, saber qué comía su ejército si su avance dependía de los viveres que encontraba a su paso; pero puede que sea irrelevante conocer el color del pelo de Napoleón o si prefería los nuevos tritos a los cocidos. Los buenos textos históricos suelen incluir, aunque no se limiten a ello, un sucinto resumen verbal de gran capacidad de pormenores históricos.

Nuestro modelo para el proceso de resumen de los detalles históricos es un *estadístico*, con el que se expresan datos de forma abreviada. Su objeto-

ningún tipo de resumen⁷. La síntesis que utilizamos también ha de ser lo suficientemente simple como para que nuestro público pueda entenderla. Ningún fenómeno puede resumirse de manera perfecta, de manera que las normas que dictan lo que es adecuado dependen de nuestros propósitos y del público. Por ejemplo, un artículo científico sobre guerras y alianzas podría incluir datos sobre 10 000 observaciones. En este trabajo, estaría justificado que hubiera resúmenes que utilizaran cincuenta números; sin embargo, incluso para un experto, cincuenta indicadores separados podrían resultar incomprensibles si no existe otro resumen. Para impartir una clase sobre esta materia a uno de los primeros cursos de licenciatura, tres cifras serían más que suficientes.

6. La inferencia descriptiva

La inferencia descriptiva es un proceso mediante el cual se comprende un fenómeno no observado a partir de un conjunto de observaciones. Por ejemplo, en las elecciones británicas de 1979 nos podría interesar como cambia en cada circunscripción el voto a los partidos Conservador, Laborista y Socialdemócrata. Lo más probable es que tengamos varias hipótesis que evaluar, sin embargo, lo que observamos realmente son las elecciones a la Cámara de los Comunes de ese año en 650 circunscripciones.

Podríamos pensar ingenuamente que, al registrar qué porcentaje de votos y escanos recibieron los conservadores en cada circunscripción, estamos observando de forma directa su fuerza electoral. Sin embargo, la política siempre conlleva un cierto margen para lo impredecible, al igual que el conjunto de la vida social y toda investigación científica⁷. Supongamos que por un súbito descaído (o como detención hacia las ciencias sociales) el Parlamento Británico hubiera decidido celebrar elecciones cada semana en 1979 y supongamos también (de forma contraria) que estas fueran independientes unas de otras. Aunque el apoyo subyacente a los conservadores se mantuviera constante, cada repetición semanal no registraría el mismo número de votos por partido en cada circunscripción. Podría cambiar la situación climatológica, surgir epidemias, la gente podría tomar sus vacaciones, y todo ello afectaría a la participación en las elecciones y a su resultado. Además, podrían producirse acontecimientos fortuitos en el contexto internacional o aparecer escándalos en los medios de comunicación, aunque estas contingencias no tuvieran importancia a largo plazo, si podrían influir en los resultados semanales. De este modo, numerosos acontecimientos transitorios podrían producir resultados electorales ligeramente diferentes. Después de todo, nuestra observación de cualquier día de estas elecciones no sería una medida perfecta de la fortaleza de los conservadores.

Por poner otro ejemplo, supongamos que nos interesa la intensidad del conflicto entre los israelíes (policía y residentes) y los palestinos en las co-

municipalidades de los territorios ocupados por Israel en Cisjordania. Los informes oficiales de ambos bandos parecen sospechosos o están censurados, de manera que decidimos llevar a cabo nuestro propio estudio. Quizá podamos verificar la intensidad general del conflicto en diferentes comunidades mediante entrevistas en profundidad o participando en actividades familiares o de otros colectivos. Si hacemos ambas cosas durante una semana en cada una de las comunidades, nuestras conclusiones sobre la intensidad del conflicto en ellas reflejarán en parte los acontecimientos que, por casualidad, sucedan en el período en que visitamos esos enclaves. Aunque nuestro estudio se llevara a cabo durante un año, tampoco podríamos conocer perfectamente el verdadero nivel de conflictividad, aunque si se redujera la incertidumbre sobre él.

En estos ejemplos, puede decirse que la varianza del voto conservador en diferentes circunscripciones o la del conflicto en las comunidades cisjordanas surge de dos factores separados: de diferencias sistemáticas y no sistemáticas. En el ejemplo electoral, las primeras incluyen características fundamentales y predecibles de las circunscripciones, como las diferencias ideológicas, de renta, de organización de la campaña o las referencias al apoyo tradicional a cada partido. En hipótesis repeticiones semanales de las mismas elecciones se mantendrían las mismas diferencias sistemáticas, pero variarían las no sistemáticas, como son los cambios producidos en la participación por las condiciones climatológicas. En el ejemplo de Cisjordania, se considerarían diferencias sistemáticas las profundas divergencias culturales entre israelíes y palestinos, su conocimiento mutuo y las pautas de segregación geográfica de la vivienda. Si pudiéramos comenzar nuestra semana de observación en doce ocasiones distintas, estas diferencias sistemáticas entre comunidades seguirían influyendo en la intensidad del conflicto observado. Sin embargo, las no sistemáticas, como los atentados terroristas o los ejemplos de brutalidad policial israelí, no serían predecibles y sólo afectarían a la semana en que ocurrieran. En general, con las técnicas inferenciales apropiadas, podemos conocer la naturaleza de las diferencias sistemáticas, aunque sea con la ambigüedad que está presente en un conjunto de datos de la vida real por las diferencias no sistemáticas o aleatorias que contiene.

Por lo tanto, uno de los objetivos fundamentales de la inferencia es distinguir entre el componente sistemático y el no sistemático dentro de los fenómenos que estudiamos. El sistemático no es más importante que el que no lo es, y nuestra atención no debería centrarse en uno a costa del otro. Sin embargo, en las ciencias sociales resulta fundamental diferenciarlos. Una de las formas de abordar la inferencia es considerar el conjunto de datos recogido como uno más de los que son posibles, al igual que los resultados reales de las elecciones británicas de 1979 son sólo uno de los muchos conjuntos de resultados posibles en diferentes días hipotéticos en los que las elecciones podrían haber tenido lugar, o del mismo modo que una semana

Uno de los objetivos de la inferencia es conocer las características sistemáticas de las variables aleatorias X_1, \dots, X_n (hay que señalar lo contradictorio de esta terminología estándar: aunque, en general, nuestros datos, lo que pretendemos en un caso determinado es tomar una variable aleatoria y extraer sus características sistemáticas). Por ejemplo, podríamos querer saber el número de votos que obtenga el Partido Laborista en la circunscripción 5 (el promedio de votos X_5 para este partido en un gran número de elecciones hipotéticas en esa circunscripción). Al ser esta una característica sistemática del sistema electoral subyacente, su valor esperado tiene un interés considerable para los científicos sociales. Por el contrario, los votos para el Partido Laborista en unas elecciones observadas, y_5 , tienen, a largo plazo, un valor bastante más reducido, ya que son el resultado de características sistemáticas y de errores aleatorios.¹⁰

El valor esperado (una de las características del componente sistemático) en la quinta comunidad de Cisjordania, El-Bireh, se expresa formalmente como sigue:

$$E(X_5) = \mu_5$$

donde $E(\cdot)$ es la operación del valor esperado, de la que surge la media en un número infinito de reproducciones hipotéticas de la semana de observación en la comunidad 5, El-Bireh. El parámetro μ_5 (la letra griega mu con el subíndice 5) representa la respuesta al cálculo del valor esperado (un nivel de conflictividad entre palestinos e israelíes) para la comunidad 5. Este parámetro es parte de nuestro modelo de características sistemáticas para una variable aleatoria X_5 . Se podría utilizar el nivel de conflictividad observado, y_5 , como un cálculo aproximado de μ_5 , pero como y_5 contiene muchos elementos imprevistos, junto con información acerca de esta característica sistemática, suele haber mejores estimadores (véase el apartado 7).

También podríamos querer conocer otra característica sistemática de estas variables aleatorias: el nivel de conflictividad en la comunidad cisjordana *na media*:

$$\frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n \mu_i = \mu \quad (2.2)$$

Un estimador de μ podría ser la media de los niveles de conflictividad observados en todas las comunidades que se estudian, y , pero también existen otros para esta característica sistemática (hay que señalar que el mismo resumen de datos que aparece en el análisis del proceso de síntesis de los pormenores históricos del apartado 5 se utiliza para hacer la estimación de una inferencia descriptiva). Entre las características sistemáticas de las va-

de observación en una pequeña comunidad es sólo una de las muchas sermanas posibles.

Con la inferencia descriptiva pretendemos comprender en qué medida nuestras observaciones reflejan fenómenos típicos o atípicos. Si las elecciones británicas de 1979 hubieran tenido lugar durante una epidemia de gripe que se hubiera extendido por los hogares de clase obrera sin afectar a los de los ricos, quizá nuestras observaciones habrían calibrado bastante mal la fuerza subyacente de los conservadores, precisamente porque el elemento no sistemático e imprevisto de los datos tendería a revelar sobre el sistema matemático o a distorsionarlo. Del mismo modo, si nuestra semana de observación en Cisjordania hubiera tenido lugar inmediatamente después de la invasión israelí del sur del Líbano, no cabría esperar resultados indicativos de lo que ocurre habitualmente en la zona.

En teoría, el mundo político puede producir múltiples conjuntos de datos para cada problema, pero no siempre atendemos las necesidades de los científicos sociales. Lo normal es que sólo tengamos la suerte de observar un conjunto de datos. Para construir un modelo dejaremos que este único conjunto lo represente una variable y (los votos recibidos por los laboristas, digamos) que se mida en todas las unidades ($n = 650$), que son las circunscripciones: y_1, y_2, \dots, y_n (y_1 , por ejemplo, podría representar a las 23,562 personas que votaran a los laboristas en la circunscripción 1). El conjunto de observaciones que denominamos y es una *variable realizada* cuyo valor varía en cada unidad. Además, definiremos X como *variable aleatoria* porque cambia aleatoriamente en todas las reproducciones hipotéticas de las mismas elecciones. En consecuencia, y_5 es el número de personas que han votado a los laboristas en la circunscripción 5 e X_5 es la variable aleatoria que representa los votos en muchas elecciones hipotéticas que podrían haberse celebrado en esa circunscripción en condiciones fundamentalmente similares. Los votos que obtiene el Partido Laborista en la única muestra que consideramos, y_1, y_2, \dots, y_n , varían en cada circunscripción por factores sistemáticos y aleatorios. Esto quiere decir que, para distinguir los dos tipos de «variables», solemos utilizar la expresión *variable realizada* para referirnos a y y *variable aleatoria* para referirnos a X .

El mismo criterio se aplica a nuestro ejemplo cualitativo. No podríamos albergar esperanza o deseo de cuantificar el grado de tensión entre israelíes y palestinos, en parte porque el «conflicto» es un asunto complicado que tiene que ver con los sentimientos de numerosos individuos, la oposición de ciertas organizaciones, las tensiones ideológicas y muchas otras cosas. En esta situación, y_5 es una variable realizada que representa la totalidad del conflicto que se ha observado en la quinta comunidad, El-Bireh. La variable aleatoria X_5 representa tanto lo que observamos en esta localidad como lo que podríamos haber observado; la aleatoriedad proviene de la variación en los acontecimientos imprevistos durante las posibles semanas que podríamos haber elegido para hacer la observación.¹¹

riables aleatorias también se incluye la varianza y diversos parámetros cau- sales que se presentan en el apartado 1 del capítulo 3.

Todavía hay otra característica de estas variables aleatorias que podría ser de interés: la varianción del nivel de conflictividad dentro de una comu- nidad aun cuando las características sistemáticas no cambien, es decir, en hipótesis de la misma variable aleatoria) producen resultados divergentes. Dicho de otro modo, éste es el tamaño del componente no sistemático que, formalmente, se calcula en una única comunidad utilizando la varianza (y no las expectativas):

$$V(X) = \sigma_1^2 \quad (2.3)$$

en donde σ^2 (la letra griega sigma) indica lo que se obtiene al aplicar el operador de varianza a la variable aleatoria X_i . Vivir en una comunidad de Cisjordania en la que hubiera un alto grado de conflictividad entre israelíes y palestinos no sería agradable, pero vivir en un lugar en el que los valores del conflicto registrarán una elevada varianza y, por tanto, fueran muy im- predecibles, podría ser peor. En cualquier caso, ambas situaciones pueden tener un interés considerable para los investigadores académicos.

Para comprender mejor estas cuestiones distinguimos dos puntos de vis- ta fundamentales sobre la varianción aleatoria¹¹, que constituyen los dos ex- tremos de una gradación. Aunque hay un numero significativo de investiga- dores que se sienten cómodo en cada uno de esos extremos, la mayoría de los politólogos tienen puntos de vista que, de alguna manera, se sitúan entre ambos polos.

Punto de vista 1: un mundo probabilístico

La varianción aleatoria existe tanto en la naturaleza como en los mundos social y politi- co, y nunca podrá eliminarse. Aunque midiéramos todas las variables sin cometer erro- res, registraríamos todos los datos (en vez de sólo una muestra) e incluiríamos toda va- riable explicativa concebible, nuestros análisis no podrían hacer nunca predicciones perfectas. Un investigador puede dividir el mundo en componentes aparentemente siste- máticos y no sistemáticos y mejorar con frecuencia sus predicciones, pero nada de lo que haga para analizar los datos podrá reducir de manera significativa el grado de varia- ción no sistemática que existe en diversas partes del mundo empírico.

Punto de vista 2: un mundo determinado

La varianción aleatoria sólo representa la parte del mundo que no podemos explicar. La división entre varianción sistemática y estocástica la impone el investigador, y depende

de que variables explicativas se encuentren disponibles y de cuales se incluya en el análisis. Con las variables explicativas apropiadas el mundo es completamente predecible.

Estos puntos de vista diferentes producen diversas ambigüedades en las inferencias de cada área de investigación¹². Sin embargo, en la mayoría de ellas *ambas perspectivas pueden considerarse equivalentes desde el punto de vista de la observación*. Esto es especialmente cierto si, con el punto de vista 2, partimos de la base de que al menos ciertas variables explicativas seguirán sin conocerse. De este modo, la equivalencia desde el punto de vista de la observación tiene lugar cuando dichas variables desconocidas del punto de vista 2 se convierten, en el punto de vista 1, en la interpretación de la varianción aleatoria. Al carecerse de cualquier consecuencia ob- servable con la que distinguir entre ellos, elegir uno u otro depende más de la fe o de la creencia que de la verificación empírica.

Por poner otro ejemplo, en ambos puntos de vista determinar si un acontecimiento político o social concreto es el resultado de un proceso sistema- tico o no sistemático depende de las opciones que tome el investigador. Usando el punto de vista 1 podemos clasificar provisionalmente un efecto como sistemático o como no sistemático, pero, a menos que logremos en- contrar otro conjunto de datos (o incluso un solo caso más) con el que com- probar la persistencia de un efecto o pauta, será muy difícil hacer una apre- ciación correcta.

Si usamos la versión más extrema del punto de vista 2, no podremos más que describir los datos, ya que juzgar «incorrectamente» que un aconteci- miento es estocástico o sistemático es imposible o irrelevante. Una versión más realista de esta perspectiva acepta que el punto de vista 1 puede consi- derar de forma correcta o incorrecta que una pauta es aleatoria o sistemáti- ca, pero deja cierta libertad para decidir cual será objeto de examen en cada estudio concreto y cual seguirá sin explicarse. En este sentido, al comenzar cualquier análisis, todas las observaciones son consecuencia de fuerzas «no sistemáticas». Por tanto, a nosotros nos corresponde demostrar que esos acontecimientos o procesos son el resultado de fuerzas sistemáticas. Deci- dir si un acontecimiento o proceso no explicado es un hecho realmente alea- torio o si procede simplemente de variables explicativas aun no identificadas queda para futuras investigaciones.

Este argumento se aplica con igual fuerza a la investigación cualitativa que a la cuantitativa. La primera suele ser histórica, pero es de extrema uti- lidad para las ciencias sociales cuando también es explícitamente interen- cial. Para conceptualizar las variables aleatorias de las que proceden las observaciones e intentar hacer una estimación de sus características siste- máticas —en vez de limitarse a resumir los pormenores históricos— no es preciso recoger datos a gran escala. De hecho, uno de los rasgos que defi- nen a un buen historiador es su capacidad para distinguir, en la situación que describe, los aspectos sistemáticos de los circunstanciales. Por lo tanto,

esta defensa de la inferencia descriptiva no es en absoluto una crítica a los estudios de caso o al trabajo del historiador. Lo que queremos decir es que, en las ciencias sociales, cualquier tipo de investigación tendría que cumplir los principios inferenciales básicos que se analizan en este libro. Demostrar que existen características sistemáticas será más difícil con algunos tipos de datos, pero no deja de ser importante hacerlo.

Como ejemplo de los problemas de la inferencia descriptiva en la investigación histórica, supongamos que nos interesaran los resultados de los encuentros en la cumbre que mantuvieron los Estados Unidos y la Unión Soviética entre 1955 y 1990. Nuestro fin último es responder a una pregunta de tipo causal: ¿en qué condiciones y en qué medida condujeron esas reuniones a una mayor cooperación? Para contestar a esta pregunta hay que resolver ciertas cuestiones difíciles del análisis causal, especialmente las relacionadas con la dirección que toma la causalidad dentro de un conjunto de variables relacionadas de forma sistemática¹⁷. Sin embargo, en este apartado, nos limitamos a tratar los problemas de la inferencia descriptiva.

Supongamos que hemos ideado una forma de calibrar —mediante análisis históricos, expertos en encuestas, recuento de acontecimientos «cooperativos» o «conflictivos», o combinando estas técnicas de medición— en qué medida se registró después de cada cumbre un aumento de la cooperación entre las superpotencias y que tenemos ciertas hipótesis sobre las condiciones en las que aumenta la cooperación, que se relacionan con cambios de poder, con los ciclos electorales en los Estados Unidos, con la situación económica en cada país y con el grado de cumplimiento de las expectativas previas de ambos bandos. Supongamos también que esperamos poder explicar el nivel subyacente de cooperación en cada año, relacionándolo de alguna manera con la presencia o ausencia de una cumbre en el período anterior, así como con nuestros propios factores explicativos.

Lo que observamos (aunque nuestros índices de cooperación sean perfectamente) es únicamente el grado de cooperación que se da *realmente* en cada año. Cuando se observen niveles de cooperación altos en los años posteriores a la celebración de las cumbres, no sabremos, sin un estudio más profundo, si hay una relación sistemática entre ambos factores. Con un pequeño número de observaciones, podría ser que asociar las cumbres a la cooperación reflejara una aleatoriedad debida a una incertidumbre fundamental (buena o mala suerte, según el punto de vista 1) o a variables explicativas no identificadas por el momento (según el punto de vista 2). Entre estas últimas variables se incluyen las fluctuaciones climatológicas que podrían producir malas cosechas en la Unión Soviética, así como los cambios en el equilibrio militar o en los mandatos, circunstancias que podrían explicar las transformaciones en el grado de cooperación. Si se identifican, estas variables serán explicaciones alternativas —variables omitidas que podrían recogerse o examinarse para evaluar su influencia en el resultado de la cumbre. Si no se identifican, pueden considerarse acontecimientos no

sistemáticos que podrían explicar el alto grado de cooperación observado entre las superpotencias. Para aportar pruebas que contradigan la posibilidad de que haya acontecimientos aleatorios (variables explicativas no identificadas) que expliquen la cooperación observada, podríamos considerar muchos otros años. Será extremadamente improbable que los acontecimientos y procesos aleatorios, al ser por definición no persistentes, produzcan una cooperación diferente en los años en que haya cumbres y en los que no las haya. De nuevo, llegamos a la conclusión de que sólo mediante pruebas repetidas en diferentes contextos (años, en este caso) podremos decidir si es mejor definir una pauta como sistemática o como resultante de las consecuencias pasajeras de procesos aleatorios.

Suele ser difícil distinguir entre procesos sistemáticos y no sistemáticos. Desde el punto de vista de las ciencias sociales, una epidemia de gripe que ataque con mayor virulencia a los votantes de clase obrera que a los de clase media es un acontecimiento impredecible (no sistemático) que, en una reproducción hipotética de las elecciones de 1979, disminuiría los votos laboristas. Sin embargo, la persistencia de un componente diferenciador de clase en la incidencia de una discapacidad sería un efecto sistemático que reduciría la media de votos laboristas en muchas reproducciones.

En las elecciones estadounidenses, el hecho de que un candidato venza a otro por su personalidad o por un lapsus linguae ocurrido en un debate televisado podría ser un factor aleatorio que quizá hubiera influido en las probabilidades de cooperación entre la URSS y los Estados Unidos durante la Guerra Fría. Sin embargo, si la principal promesa electoral hubiera sido la de reducir la tensión con los soviéticos, las constantes victorias de los candidatos conciliadores hubieran sido un factor sistemático que explicaría las probabilidades de cooperación.

Los factores sistemáticos son persistentes y sus consecuencias se repiten cuando tienen un valor determinado. Los no sistemáticos son transitorios: no podemos predecir su impacto. Sin embargo, esto no significa que los sistemáticos representen constantes. Las promesas electorales de una campaña pueden ser un factor sistemático a la hora de explicar el comportamiento electoral, pero esto no significa que dichas promesas no cambien. En el resultado de unas elecciones, la constante de las promesas electorales es su *efecto* y, si es variable, cambia de forma predecible. Cuando las relaciones soviético-estadounidenses eran buenas, puede que prometer políticas conciliadoras ganara votos en las elecciones de los Estados Unidos; cuando eran malas, quizá ocurriera lo contrario. Del mismo modo, la climatología puede ser un factor aleatorio (si los cambios intermitentes e impredecibles tienen consecuencias también impredecibles) o una característica sistemática (si el mal tiempo produce siempre menos votos para los candidatos que están a favor de políticas conciliadoras).

temos que hacer inferencias descriptivas que distinguan los fenómenos aleatorios de los sistemáticos. Saber qué ocurrió en una ocasión concreta no es suficiente. Si no nos esforzamos por determinar las características sistemáticas de un asunto, las lecciones de la historia se perderán y nada aprenderemos de los aspectos de nuestro objeto de estudio que pueden mantenerse o ser relevantes para estudios o acontecimientos futuros.

7. Criterios para juzgar las inferencias descriptivas

En este último apartado presentamos tres criterios explícitos que suelen utilizarse en estadística para juzgar los métodos inferenciales: falta de sesgo, eficiencia y coherencia. Todos se basan en el marco de las variables aleatorias que señalamos en el apartado 6 de este mismo capítulo, pero tenemos consecuencias directas y profundas para la evaluación y mejora de la investigación cualitativa. Con el fin de clarificar estos conceptos, en el presente apartado sólo damos ejemplos muy sencillos, todos ellos tomados de la inferencia descriptiva. Una versión simple de la inferencia conlleva la estimación de ciertos parámetros, incluyendo en ellos el valor esperado o la varianza de una variable aleatoria (μ o σ^2) en una inferencia descriptiva. También utilizamos estos mismos criterios para evaluar inferencias causales en el siguiente capítulo (véase su apartado 4). Reservamos para capítulos posteriores los consejos específicos que implican estos criterios para la investigación cualitativa y, en el resto del apartado, nos centramos únicamente en conceptos.

7.1 Inferencias no sesgadas

Si aplicamos una y otra vez un método inferencial obtendremos estimaciones que unas veces serán demasiado grandes y otras demasiado pequeñas. En un número elevado de aplicaciones, ¿obtenemos, en promedio, la respuesta correcta? Si la contestación es sí, entonces se dice que este método, o «estimador», no está sesgado. En un estimador, esta propiedad no nos dice en qué medida cualquiera de las aplicaciones del método está lejos del promedio, pero lo deseable es que la media sea correcta.

Se obtienen estimaciones no sesgadas cuando la variación que hay entre la reproducción de una medida y la que le sigue no es sistemática y empuja ese cálculo a veces en una dirección y a veces en otra. Se produce sesgo cuando, en un conjunto de reproducciones, hay un error sistemático de medida que lleva la estimación más en una dirección que en otra. Si en nuestro estudio de la conflictividad en las comunidades de Cisjordania los líderes políticos hubieran creado tensiones con el fin de influir en el resultado de la investigación (quizá para impulsar sus fines), el promedio de conflictivi-

dad observado en cada comunidad estaría sesgado hacia niveles altos. Si las reproducciones de nuestras hipótesis elecciones de 1979 se hubieran hecho todas en sábado (cuando se podrían haber celebrado cualquier otro día), las estimaciones estarían sesgadas si ese hecho ayudara sistemáticamente a uno de los partidos y no al otro (sí, por ejemplo, los conservadores fueran más reacios a votar en sábados por razones religiosas). Las estimaciones reproducidas también podrían haberse basado en informes de contabilizadores de votos corruptos que favorecerían a un partido más que a otro. Sin embargo, si las elecciones reproducidas se celebraran en diversos días elegidos de una manera que no estuviera relacionada con la variable que nos interesa, los posibles errores de medida no producirían resultados sesgados, aunque uno u otro día favorecería a algún partido. Por ejemplo, si hubiera fallos en el recuento de votos por descuidos aleatorios de los encargados de contabilizarlos, el conjunto de estimaciones no estaría sesgado.

Si, por ley, las elecciones británicas se celebraran siempre en domingo o uno de los partidos (por la utilización de un determinado procedimiento o quizá incluso por una persistente corrupción), necesitaríamos un estimador que variara según la media de votos que pudiera esperarse en las condiciones de estas características sistemáticas. De este modo, el sesgo depende de la teoría que se investiga, y no sólo existe en los datos. No tiene mucho sentido decir que un determinado conjunto de datos está sesgado, aunque este plagado de errores individuales.

En este ejemplo, puede que queramos distinguir entre nuestra definición de «sesgo estadístico» en un *estimador* y la de «sesgo esencial» en un *sistema electoral*. Un ejemplo del segundo tipo es un horario de votación que dificulte el ejercicio del voto a los trabajadores; sesgo esencial bastante común en varios sistemas electorales. Como investigadores, quizá queramos hacer una estimación de la media de votos en el presente sistema electoral (el del sesgo esencial), pero quizá queramos también hacer un cálculo aproximado de este indicador en un hipotético sistema que no presente un sesgo esencial debido a los horarios de los colegios electorales. Esto haría posible que calculáramos el grado de sesgo esencial del sistema. Cualquiera que sea la media que estemos calculando, lo que queremos es un estimador estadístico sin sesgos.

Los datos de las ciencias sociales pueden sufrir un sesgo importante del tipo que hay que ser consciente: las personas que nos proporcionan los datos que utilizamos en nuestras inferencias descriptivas suelen tener razones para hacer estimaciones que son, sistemáticamente, demasiado altas o demasiado bajas. Puede que los funcionarios quieran exagerar las consecuencias de un nuevo programa para pedir una mayor financiación o calcular el índice de paro a la baja para demostrar que están haciendo un buen trabajo. Quizá sea necesario buscar con más ahínco estimaciones menos sesgadas. En el estudio cualitativo de Myron Weiner (1991) acerca de la

educación y el trabajo infantil en la India encontramos un expresivo ejemplo. Al intentar explicar el reducido nivel de aceptación de la educación obligatoria en la India, en comparación con otros países, el autor primero tuvo que comprobar si dicho nivel era realmente bajo. Se dio cuenta de que, en uno de los estados del país, las estadísticas oficiales indicaban que el 98 por ciento de los niños en edad escolar acudía al colegio. Sin embargo, un examen más atento puso de manifiesto que la asistencia sólo se medía en una ocasión, la primera vez que los niños iban a la escuela. Después, figuraban en las listas de asistencia durante siete años, aunque sólo hubieran ido un día! Un escrutinio más minucioso mostró que la cifra de asistencia real era mucho más baja.

7.2 Eficiencia

Generalmente, no tenemos una oportunidad para aplicar nuestro estimador a un gran número de aplicaciones esencialmente idénticas. En realidad, excepto en algunos cuidadosos experimentos, sólo lo aplicamos una vez. En este caso, la falta de sesgo es importante, pero quisieramos estar más seguros de que la *única* estimación que hacemos se acerca al valor correcto. La eficiencia proporciona una manera de distinguir entre estimadores no sesgados. De hecho, este criterio también puede ayudarnos a diferenciar estimadores alternativos con un reducido nivel de sesgo (en general, un estimador muy sesgado debe descartarse sin siquiera evaluar su eficiencia).

La eficiencia es un concepto relativo que se mide calculando la varianza del estimador en las reproducciones hipotéticas. En los estimadores no sesgados, cuanto menor es la varianza, más eficiente (mejor) es el estimador. Una varianza pequeña es mejor porque nuestra única estimación probablemente estará más cerca del verdadero valor del parámetro. No nos interesa la eficiencia de un estimador muy sesgado porque, en esta situación, la presencia de una varianza reducida hará improbable que la estimación se acerque al valor verdadero (porque la mayoría de tales estimaciones consistirán en apretado conglomerado cercano al valor equivocado). Como discutimos posteriormente, nos interesa la eficiencia si el sesgo es pequeño, y quizá a veces estemos dispuestos a admitir un pequeño sesgo a cambio de una mayor eficiencia.

Supongamos de nuevo que nos interesara hacer una estimación del nivel medio de conflictividad entre palestinos e israelíes en Cisjordania y que estuvieramos evaluando dos métodos: una sola observación de una única comunidad, elegida como típica, y observaciones similares de, por ejemplo, 25 comunidades. Parece evidente que 25 observaciones son mejores que una sola, siempre que se haga el mismo esfuerzo en la primera que en cada una de las 25. Vamos a demostrar precisamente por qué es así. Este resultado explica por qué debemos observar tantas consecuencias de nues-

Un ejemplo formal de falta de sesgo

Supongamos, por ejemplo, que quisieramos hacer una estimación de μ en la ecuación (2.2) y que decidieramos utilizar la media como estimador. $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. En un único conjunto de datos, \bar{y} es la proporción de votos laboristas en la media de todas las circunscripciones, $n = 650$ (o el nivel medio de conflictividad en las comunidades de Cisjordania). Sin embargo, la media muestral se convierte en una función de 650 variables aleatorias, $Y = \frac{1}{n} \sum_{i=1}^n Y_i$, si se considera en un número infinito de reproducciones hipotéticas de las elecciones en cada circunscripción. De este modo, la media muestral también se convierte en una variable aleatoria. En algunas reproducciones hipotéticas, Y producirá resultados electorales que se acerquen a μ y en otras ocasiones éstos se alejarán de ella. La cuestión es si Y será acertada, es decir, igual a μ en la media de estas reproducciones hipotéticas. Para hallar la respuesta utilizamos de nuevo la operación del valor esperado, que nos permite determinar la media del infinito número de hipotéticas elecciones. Las reglas de la expectativa nos permiten realizar los siguientes cálculos:

$$E(Y) = E\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) \quad (2.4)$$

$$= \frac{1}{n} \sum_{i=1}^n E(Y_i)$$

$$= \frac{\mu}{1}$$

$$= \mu$$

Así, Y es un estimador no sesgado de μ (este ejemplo es un poco menos formal que los que aparecen en los manuales de estadística propiamente dichos, pero las características principales son las mismas).

La teoría como sea posible, pero también demuestra el concepto general de eficiencia estadística, que es igualmente relevante cuando estamos decidiendo cómo se pueden evaluar mejor las diferentes maneras de combinar en una inferencia las observaciones recogidas.

La eficiencia nos permite comparar el estimador μ del estudio de caso en el que sólo hay una observación ($n = 1$) con el gran estimador que tiene

ción que tiene con las demás. Podríamos preguntar a unos pocos residentes o leer informes de prensa para comprobar si es una comunidad media o si algún factor no sistemático ha hecho que esta observación sea atípica; después podríamos ajustar el nivel de confiabilidad observado para hacer una estimación del nivel medio, μ , que se registra en Cisjordania. Esta sería la parte más difícil del estimador del estudio de caso, y habría que tener cuidado de que no se nos colara el sesgo. Una vez que estamos razonablemente seguros de haberlo minimizado, podríamos centrarnos en aumentar la eficiencia. Para hacerlo, podríamos pasar muchas semanas en la comunidad, llevando a cabo numerosos estudios independientes. Podríamos entrevistar a líderes comunitarios, leer los periódicos, acompañar a una familia en su vida diaria y utilizar otras técnicas para recabar información. Mediantes tales procedimientos, podríamos conseguir muchas más de 25 observaciones dentro de esta única comunidad y producir un estudio de caso que, al igual que el de las 25 comunidades, no estuviera sesgado y que fuera más eficiente que éste.

Consideremos otro ejemplo: supongamos que estamos realizando un estudio sobre el problema internacional que plantean las drogas y que necesitamos calcular el porcentaje de tierra cultivable de una determinada región del mundo en el que se ha plantado cocaína. Supongamos, además, que podemos elegir entre dos métodos: un estudio de caso de un solo país o una investigación a gran escala de tipo estadístico que se ocupe de todos los países de la región. Parece mejor estudiar toda la región, pero hay que señalar que, para optar por esta alternativa, es necesario (por razones prácticas) utilizar los datos que proporcionan los gobiernos de la zona a un organismo de la ONU. Se sabe que estas cifras tienen poca relación con las pautas reales de cultivo, ya que las preparan los ministerios de asuntos exteriores y se basan en consideraciones relativas a las relaciones públicas. Supongamos también que, mediante visitas que observarían de cerca un país, podríamos corregir las estimaciones de su gobierno para acercarnos más a las cifras verdaderas. ¿Qué método elegiríamos? Quizá decidamos estudiar sólo un país, o quizá dos o tres; también podríamos analizar uno en profundidad y utilizar nuestros resultados para reinterpretar, y por tanto mejorar, los datos gubernamentales de los otros países. La elección tendría que basarse en qué datos responden mejor a nuestras preguntas.

Por poner un ejemplo más, supongamos que estamos estudiando la Comunidad Europea y queremos hacer una estimación del grado de regulación que se espera produzcan en un sector de toda la Comunidad las acciones de un gran número de normativas adoptadas formalmente por el sector industrial en cuestión, codificarlas según su severidad y, posteriormente, hacer una estimación de la severidad media de una normativa. Si recogemos datos sobre 100 normativas que muestren, en principio, un rigor similar, la varianza de nuestro cálculo será el resultado de dividir la de cualquiera de las

un n grande ($n = 25$) y que representa el nivel medio de confiabilidad encontrado en 25 estudios separados de una semana de duración en diversas comunidades de Cisjordania. Ambos estimadores, si se utilizan apropiadamente, no están sesgados. Si el mismo modelo es cierto, el estimador de la observación única tendrá una varianza de $W(Y_{\text{única}}) = \sigma^2$. Esto quiere decir que habríamos elegido una circunscripción que consideramos «típica», la cual, no obstante, se veía afectada por variables aleatorias. La varianza del gran estimador n es $W(\bar{Y}) = \sigma^2/25$, o sea, la de la media muestral. De este modo, el estimador de la observación única es veinticinco veces más variable (menos eficiente) que el que se obtiene cuando $n = 25$. De ahí que obtengamos el resultado evidente de que es mejor tener más observaciones. Más interesantes son las condiciones en las que un estudio muy detallado de nuestra única comunidad pudiera arrojar resultados tan buenos como los de la investigación que utiliza un n grande. Esto quiere decir que, aunque siempre son preferibles los estudios con más observaciones (si se dispone de los medios para recogerlas), hay situaciones en las que un único estudio de caso (como siempre, con muchas observaciones) es mejor que otro que se basa en más observaciones que, por separado, no son ni tan detalladas ni tan ciertas.

En igualdad de condiciones, nuestro análisis demuestra que cuantas más observaciones haya, mejor, porque así disminuye la varianza (y, por tanto, la ineficiencia). De hecho, la *coherencia* es una propiedad que consiste en que a medida que aumenta el número de observaciones, la varianza disminuye y se acerca a cero y la estimación es igual al parámetro que estamos intentando calcular.¹⁴

Sin embargo, es frecuente que las condiciones no sean las mismas. Sin embargo, por ejemplo, que cualquier medida individual de un fenómeno que estamos estudiando se vea influida por factores que aumenten la probabilidad de que se aleje del valor verdadero (o sea, que la varianza del estimador es alta), y supongamos que tenemos alguna idea —quizá por otros estudios— de cuáles pueden ser dichos factores. Supongamos, además, que nuestra capacidad de observar y corregir esos factores disminuye sustancialmente a medida que aumenta el número de comunidades estudiadas (aunque sólo sea porque carecemos del tiempo y del conocimiento suficientes para corregir esos factores en un número elevado de observaciones). Nos enfrentamos entonces al dilema de elegir entre un estudio de caso que tiene observaciones adicionales dentro del propio caso y 25 casos que contienen cada uno sólo una observación.

Si nuestro caso único tiene una observación, será evidentemente inferior al estudio de 25 observaciones. Sin embargo, los investigadores que utilizan estudios de caso disfrutan de considerables ventajas, que se entienden mejor al formalizarlas. Por ejemplo, podríamos comenzar por seleccionar cuidadosamente nuestra comunidad con el fin de asegurarnos de que es especialmente representativa del resto del país o de que entendemos la rela-

normativas por 100 ($\sigma^2/100$), o por menos si las normativas están relacionadas. Sin duda, la medida resultante será mejor que la que obtendríamos si utilizáramos datos sobre una sola normativa como estimador de la severidad reguladora para el conjunto del sector.

No obstante, para utilizar este procedimiento es preciso aceptar que la normativa formal es equivalente a la auténtica actividad reguladora en el sector que se estudia. Sin embargo, una investigación más profunda de la aplicación de las normativas podría poner de manifiesto una importante variación en su grado de implantación real. En consecuencia, puede que las mediciones de las normativas formales estén sistemáticamente sesgadas y que, por ejemplo, exageren la severidad reguladora. En tal caso, nos enfrentáramos de nuevo a la necesidad de alcanzar un equilibrio entre sesgo y eficiencia, y quizá hubiera sentido realizar tres o cuatro estudios de caso en profundidad sobre la implantación de normativas con el fin de investigar la relación que existe entre las normativas formales y la actividad reguladora real. Una de las posibilidades sería sustituir el estimador basado en 100 casos por otro, menos sesgado y también menos eficiente, que procediera de tres o cuatro. Sin embargo, sería más creativo, si es factible, utilizar el estudio en profundidad de los tres o cuatro casos para corregir el sesgo de nuestro estimador. En este procedimiento estaríamos combinando las aportaciones de nuestros estudios de caso en profundidad con las técnicas que manejan un n grande, práctica que pensamos debería realizarse con mucha más frecuencia de lo que es habitual en las ciencias sociales con-

temporáneas. La defensa de los estudios de caso que hacen aquellos que conocen bien una determinada parte del mundo suele ser la que está implícita en el ejemplo anterior. Puede que las investigaciones a gran escala dependan de cifras que no comprenda del todo el ingenuo investigador que trabaja con una base de datos (que quizá no sea consciente de cómo se elaboran las estadísticas electorales de un determinado lugar y presuponga, equivocadamente, que tienen una relación real con los votos emitidos). Quien trabaje de cerca con los materiales y comprenda de dónde proceden puede hacer las correcciones necesarias. En los apartados siguientes intentaremos explicar cómo se puede elegir de forma sistemática en relación a este asunto.

Precisamente, el análisis formal de este problema, que aparece en el recuadro posterior, muestra cómo se puede determinar cuál es el resultado de alcanzar un equilibrio entre eficiencia y sesgo, cuestión que surgía en el ejemplo de las circunscripciones electorales británicas. En cualquier ejemplo concreto la decisión siempre será mejor si se utiliza la lógica que aparece en el análisis formal siguiente. Sin embargo, para tomar este tipo de decisión casi siempre se necesitarán también consideraciones cualitativas. Para terminar, merece la pena plantearse de forma más específica las concesiones que a veces hay que realizar para hallar un equilibrio entre ses-

Comparaciones de eficiencia formales

La varianza de la media muestral \bar{Y} se expresa como $V(\bar{Y})$, y las reglas para calcular varianzas de variables aleatorias en el caso sencillo del muestreo aleatorio permiten que ocurra lo siguiente:

$$V(\bar{Y}) = V\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) = \frac{1}{n} \sum_{i=1}^n V(Y_i)$$

Además, si suponemos que la varianza en las reproducciones hipotéticas de las elecciones en cada circunscripción es la misma que en una de cada dos y que se expresa como σ^2 , entonces la de la media muestral será:

$$V(\bar{Y}) = \frac{1}{n} \sum_{i=1}^n V(Y_i) \quad (2.5)$$

$$= \frac{1}{n} \sum_{i=1}^n \sigma^2$$

$$= \frac{1}{n} n \sigma^2$$

$$= \sigma^2/n$$

En el ejemplo anterior $n = 650$, de manera que la varianza del estimador de este n grande es $\sigma^2/650$, y la del estimador del estudio de caso es σ^2 . La estimación estadística será preferible por su eficiencia, a menos que podamos utilizar correcciones cualitativas del error aleatorio para reducir la varianza del estimador del estudio de caso mediante un factor que valga, al menos, 650.

En cualquier conjunto grande de observaciones que no esté sesgado, la media muestral de las dos primeras tampoco lo estará, y lo mismo ocurre con la media muestral de todas ellas. Sin embargo, al utilizar sólo dos observaciones se prescinde de información sustancial; esto no afecta a la falta de sesgo, pero sí reduce considerablemente la eficiencia. Si no utilizamos también el criterio de eficiencia, careceremos de reglas formales para elegir entre un estimador y otro.

Comparaciones formales entre sesgo y eficiencia

Pensemos en dos estimadores, uno el de un estudio con un n grande que, al haber sido realizado por alguien que tiene un prejuicio, está ligeramente sesgado y otro con un n muy pequeño que creemos que no está sesgado pero que es relativamente menos eficiente y ha sido elaborado por un investigador imparcial. Como modelo formal de este ejemplo, suponemos que quisiéramos hacer una estimación de μ y que el estudio de n grande produce un estimador d :

$$d = \frac{1}{n} \sum_{i=1}^n Y_i - 0,01$$

Hacemos un modelo del estudio pequeño con un estimador diferente de μ , que es c :

$$c = \left(\frac{Y_1 + Y_2}{2} \right)$$

en el que las circunscriptpciones 1 y 2 son de tipo medio, de manera que $E(Y_1) = \mu$ y $E(Y_2) = \mu$

¿Qué estimador sería preferible? Nuestra primera respuesta es que no deberíamos utilizar ninguno de ellos y que sería mejor optar por la media muestral y , o sea, por un estudio de n grande realizado por un investigador imparcial. Sin embargo, no siempre puede aplicarse el estimador obvio o el mejor. Para solucionar este problema, recurrimos a una evaluación del sesgo y de la eficiencia.

En primer lugar, calibramos el sesgo. Podemos señalar que el primer estimador d está ligeramente sesgado según el cálculo habitual:

$$E(d) = E\left(\frac{1}{n} \sum_{i=1}^n Y_i - 0,01\right) \\ = E\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) - E(0,01) \\ = \mu - 0,01$$

También podemos indicar, mediante un cálculo similar, que el segundo estimador c no está sesgado:

$$E(c) = E\left(\frac{Y_1 + Y_2}{2}\right)$$

$$= \frac{E(Y_1) + E(Y_2)}{2}$$

$$= \frac{\mu + \mu}{2}$$

$$= \mu$$

Utilizando sólo estos cálculos, elegiríamos el estimador c , que representa el resultado del estudio con un n pequeño de nuestro investigador imparcial, porque no está sesgado. El investigador prejuicioso obtendría con d una respuesta que, en promedio, estaría equivocada, aunque sólo fuera levemente, en un número infinito de reproducciones. Con el estimador c se obtendría una media correcta.

El criterio de eficiencia presenta una situación diferente. Comenzamos por calcular la varianza de cada estimador:

$$V(d) = V\left(\frac{1}{n} \sum_{i=1}^n Y_i - 0,01\right)$$

$$= V\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) - V(0,01)$$

$$= \sigma^2/n$$

$$= \sigma^2/650$$

Esta varianza es la misma que la de la media muestral porque 0,01 no cambia (tiene una varianza cero) en las diferentes muestras. Del mismo modo, se calcula la varianza de c como sigue¹⁵:

$$V(c) = V\left(\frac{Y_1 + Y_2}{2}\right)$$

$$= \frac{1}{2} [V(Y_1) + V(Y_2)]$$

$$= \frac{1}{2} 2\sigma^2$$

$$= \sigma^2/2$$

En consecuencia, c es bastante menos eficiente que d porque $V(c) = \sigma^2/2$ es 325 veces más grande que $V(d) = \sigma^2/650$. Esto también debería quedar claro intuitivamente, ya que c desecha gran parte de la información que hay en la matriz de datos.

¿Cuál tendríamos que elegir? El estimador d está sesgado pero es más eficiente que c , mientras que éste no tiene sesgo pero es menos eficiente. En este caso concreto, probablemente sería preferible el estimador d . Por tanto, estaríamos dispuestos a sacrificar la falta de sesgo, ya que el sacrificio es bastante pequeño (0,01), con el fin de lograr un estimador bastante más eficiente. Sin embargo, habrá un momento en el que la mayor eficiencia no compensará el pequeño sesgo porque acabaremos garantizando que los cálculos aproximados están lejos de la verdad. Formalmente, la manera de calibrar el equilibrio entre sesgo y eficiencia es calcular el *error medio cuadrático* (EMC), que es una combinación de ambos factores. Si g es el estimador de algún parámetro γ (la letra griega gamma), EMC se definirá así:

$$\text{EMC}(g) = V(g) + E(g - \gamma)^2 \quad (2.6)$$

Por tanto, el error medio cuadrático es la suma de la varianza y del sesgo al cuadrado (véase Johnston, 1984, pp. 27-28). La idea es elegir el estimador que tenga el mínimo error cuadrático medio, ya que éste muestra

Supongamos que nos interesa saber si el Partido Demócrata va a ganar las próximas elecciones presidenciales en los Estados Unidos y que preguntáramos a 20 estadounidenses adultos elegidos al azar a qué partido piensan votar (en este sencillo ejemplo de selección aleatoria elegimos a los encuestados entre todos los estadounidenses adultos y cada uno de ellos tiene las mismas posibilidades de ser seleccionado). Supongamos que alguien ha realizado un estudio similar con 1.000 ciudadanos; ¿acaso tendríamos que incorporar esas observaciones de ser seleccionado). Supongamos que alguien ha realizado un estudio similar con 1.000 ciudadanos; ¿acaso tendríamos que incorporar esas observaciones adicionales a las nuestras para constituir una única estimación que se basara en 1.020 encuestados? Si las nuevas observaciones hubieran sido seleccionadas aleatoriamente, al igual que las primeras veinte, sería fácil decidir incorporar esos datos a los nuestros: con las nuevas observaciones el estimador sigue sin estar sesgado y ahora es mucho más eficiente. Sin embargo, supongamos que sólo 990 de las 1.000 nuevas observaciones fueran resultado de una selección aleatoria de la población de Estados Unidos y que las otras diez fueran congrasistas demócratas que hubieran sido incluidos, sin querer, en los datos, después de haberse hecho la muestra aleatoria. Supongamos además que nos damos cuenta de que se han incorporado dichas observaciones a nuestros datos pero que desconocemos cuáles son y que, por tanto, no podemos eliminarlas. Ahora sabemos que

precisamente por qué un estimador con cierto sesgo puede ser preferible

si su varianza es pequeña.

En nuestro ejemplo, los dos EMCs son:

$$\text{EMC}(d) = \frac{\sigma^2}{650} + (0,01)^2 \quad (2.7)$$

$$= \frac{\sigma^2}{650} + 0,0001$$

y

$$\text{EMC}(c) = \frac{\sigma^2}{2} \quad (2.8)$$

De este modo, en la mayoría de los valores de σ^2 , $\text{EMC}(d) < \text{EMC}(c)$, y d sería preferible como estimador de c .

En teoría, siempre hay que optar por estimadores no sesgados que sean lo más eficientes posible (o sea, que utilicen la mayor cantidad posible de información). Sin embargo, en las investigaciones reales que analizamos en capítulos posteriores, este equilibrio entre sesgo y eficiencia es bastante llamativo.

en principio, un estimador que se basara en los 1.020 encuestados exageraría ligeramente la posibilidad de que un demócrata ganara las elecciones nacionales. De este modo, incluir las 1.000 observaciones adicionales sería ligeramente mejor que la estimación global pero también aumentaría su eficiencia de forma considerable. En consecuencia, el que incluyamos o no estas observaciones depende de si el aumento del sesgo se compensa con el incremento de la eficiencia estadística. De una manera intuitiva, parece claro que el estimador que se base en 1.020 observaciones producirá estimaciones bastante más cercanas a la respuesta correcta con mucha más frecuencia que el que utiliza sólo 20. El sesgo resultante será lo suficientemente pequeño como para que sea preferible el estimador de la muestra más grande, aunque en la práctica probablemente aplicaríamos los dos (además, en este caso, conocemos la dirección del sesgo e incluso podríamos corregirlo parcialmente).

Si se dispone de datos cuantitativos adecuados y es posible formalizar este tipo de problemas, generalmente se puede tomar una decisión clara. Sin embargo, aunque la naturaleza cualitativa de la investigación haga que evaluar este equilibrio sea difícil o imposible, entenderlo debería ayudarnos a hacer inferencias más fiables.