



Programa de **Estudio** de los **Derechos Humanos**

Guía de estudio y antología de lecturas



Programa de Capacitación y Formación Profesional en Derechos Humanos

Fase de formación profesional

DIRECTORIO INSTITUCIONAL

PRESIDENTE

Emilio Álvarez Icaza Longoria

CONSEJO

Elena Azaola Garrido
Judit Bokser Misses
Daniel Cazés Menache
Santiago Corcuera Cabezut
Denise Dresser Guerra
Patricia Galeana Herrera
Ángeles González Gamio
Armando Hernández Cruz
Clara Jusidman Rapoport
Carlos Ríos Espinosa

SECRETARÍAS

Ejecutiva •Luis J. Vaquero Ochoa

Técnica •Ricardo Bucio Mújica

VISITADURÍAS

Primera •José Cruz Lavanderos Yáñez

Segunda •Patricia Colchero Aragonés

Tercera •Luis González Placencia

Cuarta •María Alejandra Nuño Ruiz Velasco

DIRECCIONES GENERALES

Administración

Román Torres Huato

Comunicación Social

Hugo Morales Galván

Educación y Prom. de los Derechos Humanos

Daniel Ponce Vázquez

Quejas y Orientación

Víctor Morales Noble

CONTRALORÍA INTERNA

Rosa María Cruz Lesbros

DIRECCIONES EJECUTIVAS

Investigación y Desarrollo Institucional

Sergio Santiago Galván, encargado de Despacho

Seguimiento

Tania Reneaum Panszi

COORDINACIONES

Asesores

Luis J. Vaquero Ochoa, encargado de Despacho

Asuntos Jurídicos

María del Rosario Laparra Chacón

Interlocución Institucional y Legislativa

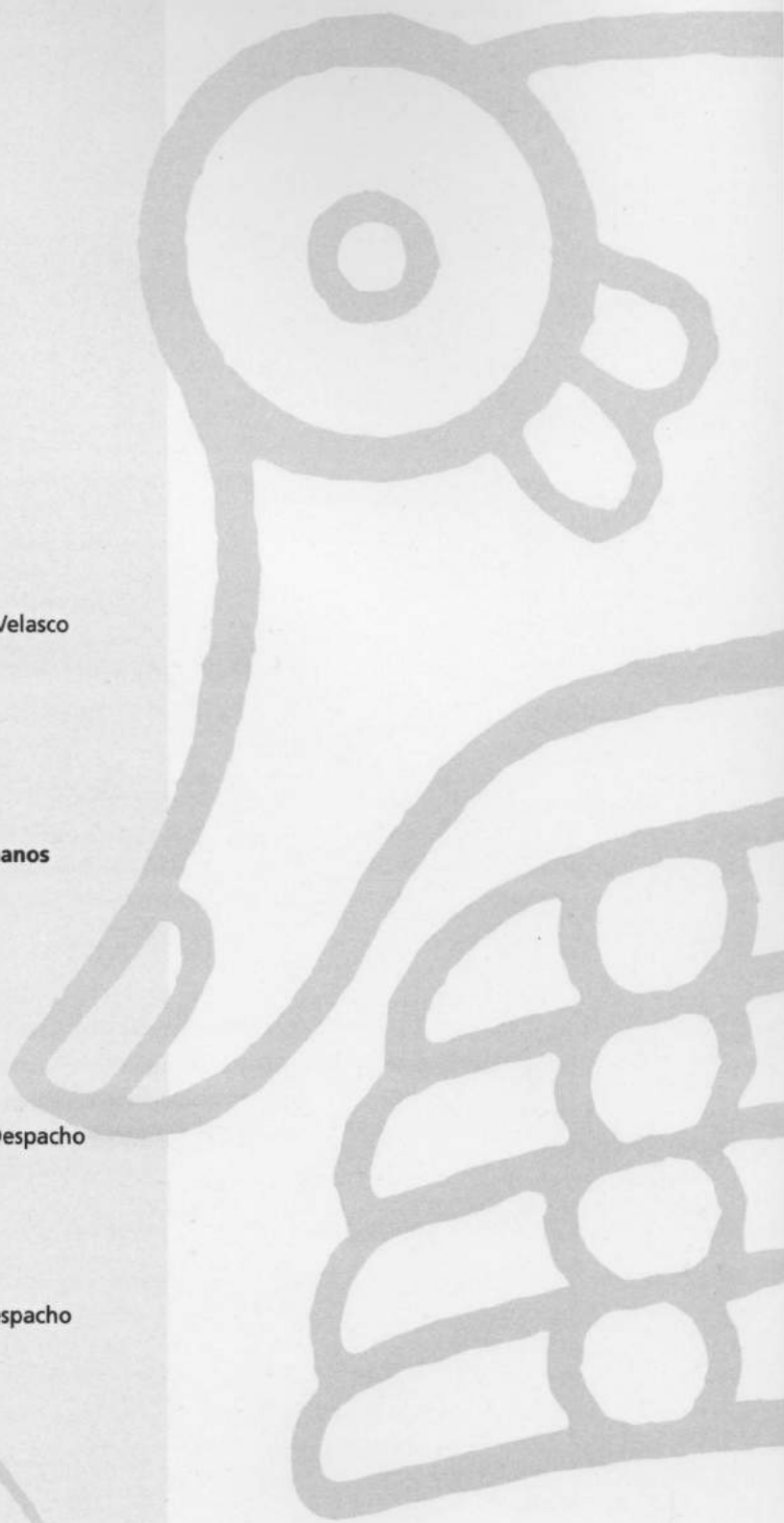
Víctor Brenes Berho

Relatoría para la Libertad de Expresión y Atención a Defensoras y Defensores de Derechos Humanos

Eréndira Cruzvillegas Fuentes

SECRETARÍA PARTICULAR DE LA PRESIDENCIA

Laura Gutiérrez Robledo







Programa de **Estudio** de los **Derechos Humanos**

Guía de estudio y antología de lecturas

Programa de Capacitación y Formación Profesional en Derechos Humanos

Fase de formación profesional

Esta guía fue elaborada bajo la coordinación de la Dirección Ejecutiva de Investigación y Desarrollo Institucional de la CDHDF con la colaboración del doctor Alejandro Monsiváis Carrillo.

Primera edición, 2008

D. R. © 2008, Comisión de Derechos Humanos del Distrito Federal
Av. Chapultepec 49, 06040 México, D. F.
www.cd hdf.org.mx

Material didáctico del Programa de Capacitación del Servicio Profesional en Derechos Humanos para uso exclusivo de la Comisión de Derechos Humanos del Distrito Federal. Prohibida su venta.

Impreso en México

Printed in Mexico

ÍNDICE

Presentación	9
Objetivos	9
Introducción	10
MÓDULO I. INTRODUCCIÓN A LA METODOLOGÍA DE LA INVESTIGACIÓN	
A. Síntesis	
I. Dimensiones de organización de la información	13
a. Sociedad, cultura y comunicación. Un punto de partida	13
II. Definiciones	14
a. Definición de <i>investigación científica</i> en ciencias sociales	14
b. Definición de <i>metodología</i>	14
c. Planteamiento del problema de investigación	14
III. Criterios para plantear el problema de investigación	15
IV. Elementos del planteamiento del problema de investigación	16
a. Objetivos de la investigación	16
b. Preguntas de investigación	16
c. Justificación de la investigación	17
V. Estrategias de investigación	17
VI. Indicadores y definiciones operativas	19
B. Material de estudio	
Lectura 1: "Planteamiento del problema: objetivos, preguntas de investigación y justificación del estudio", <i>Roberto Hernández, Carlos Fernández, Pilar Baptista</i>	21
Lectura 2: "La lucha de la luz y la sombra", <i>Luis Jesús Galindo</i>	42
Lectura 3: "La ciencia en las ciencias sociales", <i>Gary King, Robert O. Keohane, Sydney Verba</i>	59
Lectura 4: "La inferencia descriptiva", <i>Gary King, Robert O. Keohane, Sydney Verba</i>	90
Lectura 5: "Causalidad e inferencia causal", <i>Gary King, Robert O. Keohane, Sydney Verba</i>	131
Lectura 6: "Metodología de la investigación política", <i>Stefano Bartolini</i>	169
C. Ejercicios	208
D. Autoevaluación	211

MÓDULO II.	INVESTIGACIÓN CUANTITATIVA Y CUALITATIVA	
	A. Síntesis	
	i. Enfoques en la investigación social	217
	ii. La metodología de la encuesta	219
	iii. Grupos de discusión	220
	a. Selección de los participantes	221
	b. Integrantes del grupo de discusión	221
	c. Tamaño del grupo	221
	B. Material de estudio	
	Lectura 1: “El proceso de investigación y los enfoques cuantitativo y cualitativo: hacia un modelo integral”, <i>Roberto Hernández, Carlos Fernández, Pilar Baptista</i>	222
	Lectura 2: “La metodología de la encuesta”, <i>Heriberto López Romo</i>	247
	Lectura 3: “Grupos de discusión. De la investigación social a la investigación reflexiva”, <i>Bernardo Russi</i>	288
	C. Ejercicios	329
	D. Autoevaluación	332
MÓDULO III.	METODOLOGÍA DE LA INVESTIGACIÓN EN DERECHOS HUMANOS	
	A. Síntesis	
	i. El estado de las investigaciones en derechos humanos	337
	ii. Antecedentes de las investigaciones cuantitativas sobre derechos humanos	337
	iii. Utilidad de las investigaciones cuantitativas sobre derechos humanos	338
	iv. Conceptualización de los derechos humanos	338
	v. Aproximación a las fuentes cuantitativas sobre derechos humanos	339
	a. La situación actual de las fuentes de información	339
	b. Publicaciones periódicas con tratamiento particular de la información	340
	c. Estudios monográficos	341
	d. Fuentes con caracteres de periodicidad y generalidad en la información	341
	e. Los informes de Freedom House y Charles Humana	342
	vi. Objeciones a las fuentes cuantitativas y a las conclusiones teóricas	342
	a. Objeciones al sistema de obtención de datos	342
	b. Objeciones desde los criterios estadísticos de fiabilidad	343

c. Objeciones desde el multiculturalismo	
B. Material de estudio	
Lectura: "La investigación cuantitativa y estadística sobre los derechos humanos"	
<i>Ignacio Aymerich</i>	344
C. Ejercicios	526
D. Autoevaluación	529

MÓDULO IV. LA ESTADÍSTICA EN LA INVESTIGACIÓN EN DERECHOS HUMANOS

A. Síntesis	
i. Conceptos básicos de la estadística descriptiva	535
a. Selección de la muestra	535
b. Selección de la unidad de análisis	535
c. Tipos de muestra	535
ii. Tipos de muestreo	536
a. Muestreo aleatorio simple	536
b. Muestreo aleatorio sistemático	536
c. Muestreo aleatorio estratificado	537
d. Muestreo por cuotas	537
e. Tamaño de la muestra	538
iii. Procedimiento para el análisis cuantitativo de los datos	538
a. Distribución de frecuencias	539
b. Medida de tendencia central	539
c. Medidas de variabilidad	540
B. Material de estudio	
Lectura 1: "Tipos de muestreo",	
<i>Jacinto Rodríguez</i>	542
Lectura 2: "Tamaño de la muestra",	
<i>Jacinto Rodríguez</i>	567
Lectura 3: "Análisis de datos",	
<i>Roberto Hernández, Carlos Fernández, Pilar Baptista</i>	584
Lectura 4: "El uso de la computadora como auxiliar en el análisis de datos cualitativos",	
<i>Lilia Chernobilsky</i>	714
C. Ejercicios	746
D. Autoevaluación	749

Claves de respuesta	751
---------------------------	-----

PRESENTACIÓN

El presente curso forma parte de la Fase de Formación Profesional y está dirigido a las y los miembros del Servicio Profesional de Derechos Humanos (SPDH) del Programa de Estudio de la Comisión de Derechos Humanos del Distrito Federal (CDHDF). A través del mismo, se abordarán los conocimientos necesarios para desarrollar las habilidades prácticas que requiere el personal para el mejor desempeño de sus funciones.

Este curso incluye cuatro módulos: Introducción a la metodología de la investigación, Investigación cuantitativa y cualitativa, Metodología de la investigación en derechos humanos, y Estadística en la investigación en derechos humanos.

El contenido de los módulos ha sido pensado de modo tal que se abarquen todos los pasos en el diseño y realización de la investigación en derechos humanos desde la definición del problema de investigación hasta el análisis de los datos. En este sentido, los temas incluidos en esta guía ofrecen a las y los miembros del SPDH diversas técnicas y estrategias de investigación en derechos humanos con el objetivo de que las y los servidores públicos tomen las decisiones metodológicas más pertinentes en cada investigación que lleven a cabo.

Esperamos que esta guía constituya un referente obligado que facilite a las y los miembros del SPDH la resolución de problemas prácticos, la realización de ejercicios que permitan enfrentar problemas metodológicos reales, la comunicación entre la y el servidor público y la y el docente en las asesorías, y la integración de los procesos involucrados en el estudio de los derechos humanos.

OBJETIVOS

Objetivo general

Contribuir a la formación profesional del personal que labora en las áreas de defensa, promoción y estudio de los derechos humanos, para desarrollar sus capacidades y habilidades de manera permanente y sistemática, a través de un esquema que brinde las herramientas necesarias que mejoren el desempeño de sus funciones.

Objetivo específico

Proporcionar los conocimientos y habilidades indispensables para el diseño y desarrollo de estudios, investigaciones y diagnósticos sobre el estado que guardan los derechos humanos en el Distrito Federal.

INTRODUCCIÓN

La presente guía de estudio y antología de lecturas es una herramienta pedagógica que tiene por objeto facilitar a las y los miembros del Servicio Profesional en Derechos Humanos (SPDH) el conocimiento requerido para la Fase de Formación Profesional del Programa de Estudio. Como tal, ha sido diseñada para su uso integral; cada una de las partes que la conforman tiene una función complementaria con el resto de los componentes.

Este documento contempla cuatro módulos que corresponden a cada uno de los temas generales a abordar durante el curso y para términos prácticos se conforma de un índice general de contenidos, y cuatro secciones principales: síntesis, material de estudio (antología de lecturas), ejercicios y autoevaluación.

SÍNTESIS

Con ella se busca facilitar el estudio de las y los miembros del SPDH, además de contribuir a situarse frente al tema y destacar la relevancia de los contenidos. Para tal fin, se ha realizado la síntesis del contenido sustantivo de los módulos, sustraída de las lecturas que conforman la bibliografía obligatoria del curso, y que destaca, básicamente, los conceptos elementales de cada tema. Esta síntesis no es exhaustiva del contenido de los módulos, su intención es dar un marco conceptual que permita a las y los servidores públicos comprender de mejor forma los temas contenidos en cada uno de los módulos de estudio.

MATERIALES DE LECTURA

Se incluye una antología de lecturas obligatoria correspondiente a los temas de estudio de cada módulo. La lectura de esta antología resulta ser la actividad más relevante y sustancial a desarrollar de este mecanismo de autoestudio, por ello, es necesario que las y los servidores públicos pongan particular interés en su desahogo.

EJERCICIOS

Las y los servidores públicos resolverán una serie de ejercicios para que, con base en las lecturas, su experiencia y la guía docente, fortalezcan sus habilidades para el estudio de los derechos humanos.

AUTOEVALUACIÓN

La última parte de cada uno de los módulos está integrada por una autoevaluación. Esta sección tiene un doble objetivo: por una parte, tener una idea clara del nivel de comprensión alcanzado sobre los temas tratados o, en su caso, identificar las áreas que requieren un estudio particular. Por otra, permitir a las y los miembros del SPDH familiarizarse con el tipo de evaluación que presentarán al concluir los trabajos del curso. Esta sección incluye las claves de respuesta que permitirán a las y los miembros del SPDH revisar los resultados de sus autoevaluaciones y así verificar los avances obtenidos a partir del estudio dedicado a los contenidos del curso.

Módulo I.
Introducción a la metodología
de la investigación

OBJETIVOS ESPECÍFICOS

- Proporcionar los conocimientos básicos para construir planteamientos de investigación en derechos humanos.
- Identificar los componentes de un proyecto de investigación, su coherencia interna y la relación entre éstos.
- Generar capacidades analíticas para el desarrollo de argumentos.

A. SÍNTESIS

I. Dimensiones de la organización de la información

a. Sociedad, cultura y comunicación. Un punto de partida

El mundo contemporáneo, cada vez más complejo y diferenciado, exige esfuerzos de comprensión desde distintas aristas. La economía, la política, el arte, la religión observan al mundo desde su propia perspectiva y según sus propios criterios. La ciencia social, por su parte, tiene una manera especial de aproximarse a la complejidad actual. Este tipo de observación, el de segundo orden, es el campo de la reflexión metodológica.

Sin embargo, es necesario tener en cuenta los rasgos generales de una apuesta teórico-metodológica que permita ordenar los diversos paquetes técnicos. Ese orden se divide en tres dimensiones de organización de información sobre lo social, lo que corresponde a la sociedad, a la cultura y a la ecología. Cada dimensión se ordena en niveles de complejidad, los cuales son útiles para ubicar objetos particulares de observación y configuración reflexiva (Galindo, p. 46). De acuerdo con Galindo estas dimensiones se muestran en el siguiente cuadro:

Cuadro 1.1. Dimensiones de organización de la información

Sociedad	Se refiere a lo más evidente de la composición, lo que aparece a la mirada del observador que inicia una indagación. Es el plano de la observación que se ordena según el objeto cognitivo exploración y/o descripción.
Cultura	Se encuentra más allá de lo evidente, la estructura que configura a la diversidad y heterogeneidad de lo social.
Comunicación	Es lo más amplio de toda la observación reflexiva, lo que permite relacionar lo social-cultural con lo no social ni cultural, lo que marca los ámbitos generales de configuración de la escena humana, lo que marca los límites de lo humano y lo no humano.

Una manera de observar y sintetizar la complejidad del mundo contemporáneo es mediante la investigación social. A continuación se describirán sus principales características y las reglas generales propias de cualquier procedimiento de investigación científica.

II. Definiciones

a. Definición de investigación científica en ciencias sociales

Para definir qué es una buena investigación se debe recurrir a la palabra *científica* como calificativo. En cuanto al diseño, la investigación tiene cuatro características:

1. **El objetivo es la inferencia.** 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.
2. **Los procedimientos son públicos.** Para generar y analizar datos la investigación científica utiliza métodos explícitos, codificados y públicos que, por lo tanto, pueden evaluarse.
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. Por supuesto, es imposible llegar a conclusiones perfectamente ciertas utilizando datos inciertos.
4. **El contenido es el método.** La investigación científica propugna un conjunto de normas inferenciales de las que depende su validez (King, Keohane y Verba, p. 64 y 65).

Estas son cuatro características generales que debe poseer cualquier diseño de investigación científica. Más que el tema seleccionado, lo que le da validez a una investigación es el procedimiento para llegar a determinados resultados, es decir, la metodología de la investigación.

b. Definición de *metodología*

El concepto *metodología* se usa en las ciencias humanas con diversos significados. Por una parte, con frecuencia se incluye en su esfera de competencia el estudio del fundamento filosófico del conocimiento de tipo científico. Por otra, se entiende por metodología el estudio de técnicas específicas de investigación. Está un tercer nivel del término *metodología* que se ubica de manera intermedia entre la filosofía y las técnicas de investigación. Esta tercera acepción de la metodología de investigación incluye aquellos procedimientos lógicos que se refieren a la formulación de los problemas de investigación, la formación y el tratamiento de conceptos, la elección de los casos y las variables, los procedimientos de control de los resultados (Bartolini, p. 169).

A continuación se enuncian los componentes básicos de la metodología de la investigación comenzando por la selección y formulación del problema de investigación.

c. Planteamiento del problema de investigación

Una vez que se ha concebido el tema de la investigación (por ejemplo, los derechos humanos de las mujeres) y el investigador ha leído la literatura al respecto y ha profundizado en el tema en cuestión, se encuentra en condiciones de plantear el problema de investigación (Hernández, p. 23).

En realidad, plantear el problema de investigación no es sino afinar y estructurar más formalmente la idea de investigación. El paso de la idea al planteamiento del problema en ocasiones puede ser inmediato, casi automático, o bien llevar una considerable cantidad de tiempo; ello depende de cuán familiarizado esté el investigador con el tema, la complejidad de la idea, la existencia de estudios antecedentes, el empeño del investigador, el enfoque elegido y sus habilidades personales. Seleccionar un tema o una idea no lo coloca inmediatamente en la posición de considerar qué información habrá de recolectar, con cuáles métodos y cómo analizará los datos que obtenga. Antes necesita formular el problema específico en términos concretos y explícitos, de manera que pueda investigarse con procedimientos científicos (Hernández, p. 23).

Ahora bien, un problema correctamente planteado está casi resuelto; a mayor exactitud corresponden más posibilidades de obtener una solución satisfactoria. De cualquier manera, el investigador debe ser capaz no sólo de conceptualizar el problema, sino también de verbalizarlo en forma clara, precisa y accesible. En algunas ocasiones sabe lo que desea hacer pero no cómo comunicarlo a los demás y es necesario que realice mayor esfuerzo para traducir su pensamiento en términos comprensibles, pues en la actualidad la mayoría de las investigaciones requieren la colaboración de muchas personas (Hernández, pp. 23 y 24).

III. Criterios para plantear el problema de investigación

De acuerdo con Roberto Hernández Sampieri y otros autores, los criterios para plantear adecuadamente el problema de investigación son:

- Establecer relación entre variables. Es decir, el problema debe expresar una relación entre las propiedades del fenómeno que se está estudiando; por ejemplo, la relación entre violaciones a los derechos humanos y políticas públicas.
- El problema debe estar formulado claramente y sin ambigüedad como pregunta (por ejemplo, ¿qué efecto?, ¿en qué condiciones?, ¿cómo se relaciona con?). Cuando se llegue al punto de plantear el problema de investigación, éste debe formularse con claridad y evitando la ambigüedad, lo que es muy diferente a la dispersión de datos o apertura en la información.
- El planteamiento debe implicar la posibilidad de realizar una prueba empírica o una recolección de datos, es decir, la factibilidad de observarse en la realidad o en un entorno. Las ciencias trabajan con aspectos observables y medibles en la realidad o en determinado ambiente (Hernández, p. 24).

En el mismo tenor, Stefano Bartolini señala que la formulación o planteamiento del problema debe tener las siguientes características:

- **Explícita.** La formulación de un problema empírico para la investigación debería ser, en primer lugar, lo más *explícita* posible; es decir, explicitar al máximo la interrogante central de la investigación, separándola y distinguiéndola de las interrogantes periféricas o secundarias.
- **Clara.** En segundo lugar, la formulación del problema debe ser *clara* en el sentido de que los términos utilizados en ella no deben generar ambigüedad acerca de su significado, ni deben encontrarse en ellos asertos o respuestas implícitas.

- **Debe dar respuestas empíricas.** En tercer lugar, el problema debe formularse de tal modo que pueda tener una respuesta empírica, excluyendo, pues, aquellas cuestiones a las que, indirectamente o de modo incompleto, no se les puede dar tales respuestas.
- **Debe tener un valor teórico.** Por último, una correcta formulación del problema debe estar dotada de *valor teórico*, debe estar formulada de tal modo que se inserte en y contribuya al desarrollo de los conocimientos ya acumulados (Bartolini, p. 172).

Además, cualquier planteamiento del problema debe servir para generar inferencias válidas sobre la vida social y política. En las ciencias sociales, la investigación, ya sea cuantitativa o cualitativa tiene dos objetivos: describir y explicar. La inferencia descriptiva es un proceso mediante el cual se comprende un fenómeno no observado a partir de un conjunto de observaciones (King, Keohane y Verba, pp. 59, 90, 111).

Un planteamiento del problema con estas características fortalece todo el procedimiento de la investigación científica porque evita la dispersión de los datos y la búsqueda de información innecesaria. Estas cualidades permitirán que se elaboren correctamente cada uno de los elementos del planteamiento del problema.

iv. Elementos del planteamiento del problema de investigación

Los elementos para plantear un problema son tres y están relacionados entre sí:

- Los objetivos de la investigación
- Las preguntas de investigación
- La justificación del estudio (Hernández, p. 25)

a. Objetivos de la investigación

En primer lugar es necesario establecer qué se pretende con la investigación, es decir, *cuáles son sus objetivos*. Hay investigaciones que intentan, ante todo, contribuir a resolver un problema especial, en ese caso deberá mencionarse cuál es y de qué manera se piensa que el estudio ayudará a resolverlo (Hernández, p. 25).

Los objetivos tienen que expresarse con claridad para evitar posibles desviaciones en el proceso de investigación y deben ser alcanzables. Son las guías que deberán tenerse presentes durante toda la investigación. Por supuesto, los objetivos que se especifiquen deberán ser congruentes entre sí (Hernández, p. 25).

b. Preguntas de investigación

Plantear el problema en forma de preguntas tiene la ventaja de presentarlo de manera directa, minimizando la distorsión. Las preguntas generales tienen que aclararse y delimitarse para esbozar el área-problema y sugerir actividades pertinentes para la investigación (Hernández, p. 26).

Hay preguntas generales que no conducirían a una investigación concreta, como ¿por qué se violan los derechos humanos? Esta es una pregunta tan general y abierta que con-

duciría a muchos problemas todos ellos diversos. Cuanto más precisas sean las preguntas, más servirán de orientación hacia las respuestas que se buscan.

c. Justificación de la investigación

Además de los objetivos y las preguntas de investigación, es necesario justificar el estudio exponiendo sus razones. Se tiene que explicar por qué es conveniente llevar a cabo la investigación y cuáles son los beneficios que se derivarán de ella (Hernández, p. 30).

Una investigación puede ser conveniente por diversos motivos. Algunos criterios para valorar su potencial son los siguientes:

- Conveniencia, ¿para qué sirve la investigación?
- Relevancia social, ¿cuál es su trascendencia para la sociedad?
- Implicaciones prácticas, ¿ayudará a resolver un problema real?

Desde luego, es muy difícil que una investigación pueda responder positivamente a todas estas interrogantes; algunas veces sólo cumplen un criterio (Hernández, p. 31).

Además de los tres elementos que conforman el planteamiento del problema, es necesario considerar otro aspecto importante: la viabilidad o factibilidad del estudio; para ello, debemos tomar en cuenta la disponibilidad de recursos financieros, humanos y materiales que determinarán en última instancia los alcances de la investigación (Hernández, p. 32).

Una vez que se tienen definidos los objetivos, las preguntas y la justificación del planteamiento del problema de la investigación se encuentran dadas las condiciones para definir las estrategias de la investigación que se deberán tomar en consideración.

v. Estrategias de investigación

Una formulación correcta del problema en el sentido expuesto ofrece la ventaja de indicar con suficiente claridad qué unidades y qué propiedades o características de tales unidades deberá tomar en consideración la investigación. Las *unidades* son los tipos de objetos o acontecimientos de que se ocupa una determinada investigación científica (Bartolini, p. 173) por ejemplo, instituciones de gobierno, individuos, derechos humanos, u otros. Las *propiedades*, en cambio, son las características o las dimensiones de las unidades que se considerarán importantes para la investigación (Bartolini, p. 173).

Por ejemplo, para un estudio que tiene como unidad de análisis los derechos políticos, una determinada formulación del problema podría considerar importantes ciertas propiedades de dichos derechos como las elecciones libres y no manipuladas, la existencia de una oposición real, la pluralidad partidista, la posibilidad de la ciudadanía de organizarse en agrupaciones políticas, la descentralización del poder político, entre otros.

La existencia de diversas estrategias de investigación estará en función *el número de unidades y número de propiedades* que se consideren. Puede haber muchas unidades y propiedades; sin embargo, es necesario realizar una selección importante de ellas para un estudio concreto (Bartolini, p. 174).

Según el número de unidades y propiedades se pueden configurar dos estrategias de investigación. La primera tiene un carácter *intensivo*; es decir, considerando muy pocas

unidades y muchas propiedades. La segunda estrategia tiene un carácter *extensivo*, pues considera muy pocas propiedades en un gran número de casos (Bartolini, p. 174).

Veamos las diferencias entre los tipos de estrategias con un ejemplo: supongamos que se quiere hacer una investigación sobre el derecho a un medioambiente sano en el Distrito Federal. Las estrategias aplicables se describen en los siguientes cuadros:

Cuadro 1.2. Estrategia intensiva

Tema	Derecho a un medioambiente sano en el Distrito Federal.
Unidad de análisis	Quejas de violaciones al derecho a un medioambiente sano.
Número de casos	3
Propiedades	Datos sociodemográficos de las personas que presentaron la queja (sexo, edad, ocupación, nivel de ingresos, domicilio, escolaridad, etcétera). Descripción específica del derecho violado (agua, aire, etcétera). Proceso de la denuncia (a qué instancia acudió, cuándo, cuál fue la respuesta de las autoridades). Temporalidad del caso (cuándo se presentó el problema, cuándo presentó denuncia, cuándo acudió a la Comisión). Papel de las autoridades al respecto (cuáles son las instancias responsables, qué respuestas han dado, cuál es su posición al respecto). Información de contexto (cuáles son las características de la zona, qué tan extendido está el problema, a cuánta población afecta, cuáles parecen las causas estructurales).

En este ejemplo se están considerando sólo tres casos de la unidad de análisis y más de veinte propiedades para cada caso. Cuánto mayor sea el número de propiedades menor será el número de casos, pues de ese modo la investigación será menos costosa.

Uno de los principales objetivos de esta estrategia es profundizar en la complejidad de los casos seleccionados. En el siguiente cuadro se presenta el mismo ejemplo con una estrategia extensiva:

Cuadro 1.3. Estrategia extensiva

Tema	Derecho a un medio ambiente sano en el Distrito Federal.
Unidad de análisis	individuos que ven afectado su derecho a un medioambiente sano.
Número de casos	5000
Propiedades	Sexo Edad Domicilio Si consideran que se ha violado su derecho a un medioambiente sano Si han presentado denuncia al respecto Si han presentado una queja ante la CDHDF

Estas propiedades podrían extenderse hasta formar un instrumento con 15 o 20 preguntas aplicable en todos los casos. Como puede verse, con la estrategia intensiva no se trata de profundizar en la complejidad de cada caso, sino abarcar muchos de ellos con sólo unas cuantas propiedades claramente delimitadas.

vi. Indicadores y definiciones operativas

Las propiedades y las unidades hacen referencia a conceptos que se observan y se miden en la realidad. Por ejemplo, el “derecho a un medioambiente sano” puede entenderse como un valor, como una norma positiva o como un hecho social concreto. La manera en que se entienden los conceptos influye en la manera en que se estudian.

Los conceptos pueden ser tan generales y de tan elevado nivel de abstracción que podría ser difícil obtener de ellos criterios válidos para identificar referentes empíricos. Para colmar esta distancia entre intención de un significado conceptual y sus referentes empíricos hay que dar algunos pasos intermedios. Estos pasos son la identificación de indicadores de los conceptos y de definiciones operativas (Bartolini, p. 181).

La definición operativa de un concepto es un tipo de definición que incorpora en su interior la especificación del campo de referentes empíricos del concepto. Está dotada en su interior de una serie de instrucciones que indican cómo se propone etiquetar, medir o identificar el concepto dado (Bartolini, p. 182). Véase el cuadro siguiente como ejemplo.

Cuadro 1.4. Definición operativa de derechos políticos

Concepto	Derechos políticos
Definición operativa	Aquellos que posibilitan a la población a participar libremente en los procesos políticos

La definición operativa de un concepto puede desagregarse en una serie de atributos más detallados que ofrezcan la posibilidad de operativizar de manera directa al concepto principal. A estos atributos más específicos se les llama indicadores. En el siguiente cuadro se ejemplifica este caso.

Cuadro 1.5. Indicadores del concepto *derechos políticos*

Concepto	Derechos políticos
Definición operativa	Aquellos que posibilitan a la población a participar libremente en los procesos políticos
Indicadores	Elecciones libres y no manipuladas Porcentaje significativo de voto a las opciones de la oposición Administraciones locales dirigidas por autoridades electas libremente Porcentaje del voto a las opciones de la oposición Leyes electorales equitativas

En general, es necesaria una serie de indicadores para expresar completamente el significado y los referentes de un concepto complejo. Se trata, pues, de identificar las dimensiones o las características relevantes de un concepto y establecer un indicador para cada una de ellas. La operación implica un proceso de segmentación del concepto complejo y los indicadores escogidos para cada segmento del concepto deberían corresponder en conjunto bastante bien al universo más amplio de características que tenemos en mente cuando utilizamos el concepto originario (Bartolini, p. 182).

Para concluir, hay que señalar que un planteamiento del problema claro, coherente y explícito en forma de objetivos y preguntas, así como la definición de las estrategias de investigación, son pasos básicos en la elaboración de cualquier proyecto. La elaboración de un

proyecto de investigación no se agota con lo dicho hasta aquí, sin embargo, con esto se han planteado las condiciones mínimas que se deben considerar al elaborar un proyecto de investigación.

B. MATERIAL DE ESTUDIO

Bibliografía obligatoria

Hernández, Roberto *et al.*, "Planteamiento del problema: objetivos, preguntas de investigación y justificación del estudio", en *Metodología de la investigación*, México, McGraw-Hill, 2003, pp. 41-60 **[Lectura 1]**.

Galindo, Luis Jesús, "La lucha de la luz y la sombra", en Luis Jesús Galindo (coord.), *Técnicas de investigación en sociedad, cultura y comunicación*, México, Pearson, 1998, pp. 9-25 **[Lectura 2]**.

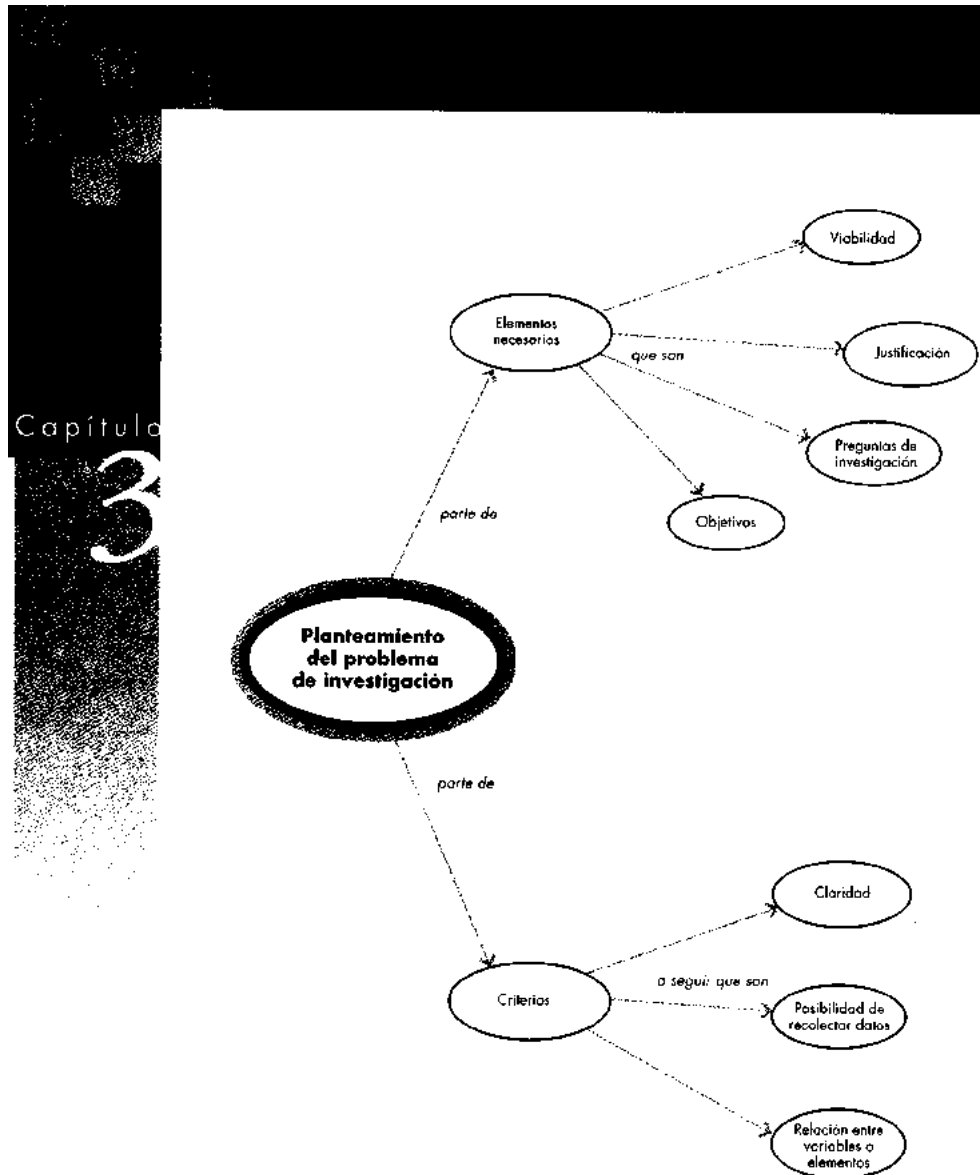
King, Gary *et al.*, "La ciencia en las ciencias sociales", en *El diseño de la investigación social. La inferencia científica en los estudios cualitativos*, Madrid, Alianza, 2007, pp. 13-43 **[Lectura 3]**.

_____, "La inferencia descriptiva", en *El diseño de la investigación social, op. cit.*, pp. 45-85 **[Lectura 4]**.

_____, "Causalidad e inferencia causal", en *El diseño de la investigación social, op.cit.*, pp. 87-124 **[Lectura 5]**.

Bartolini, Stefano, "Metodología de la investigación política", en *Manual de ciencia política*, Madrid, Alianza, 1996, pp. 39-79 **[Lectura 6]**.

Lectura 1



Tomada de: Roberto Hernández *et al.*, "Planteamiento del problema: objetivos, preguntas de investigación y justificación del estudio, en *Metodología de la investigación*, México, McGraw-Hill, 2003, pp. 41-60.



PROCESO DE INVESTIGACIÓN

Segundo paso

PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN

- Establecer objetivos de investigación
- Desarrollar las preguntas de investigación
- Justificar la investigación y analizar su viabilidad

Síntesis

En el presente capítulo se mostrará la manera en que la idea se desarrolla y se transforma en el planteamiento del problema de investigación científica. Es decir, el capítulo trata sobre cómo plantear un problema de investigación científica. Los elementos resultan fundamentales para plantear un problema: objetivos de investigación, preguntas de investigación y justificación de la investigación. En el capítulo se analizan estos elementos.

Asimismo, se vincula el planteamiento del problema con los enfoques cuantitativo, cualitativo y mixto de la investigación.

Capítulo

3

OBJETIVOS DE APRENDIZAJE

Al terminar este capítulo, el alumno será capaz de:

- Formular de manera lógica y coherente problemas de investigación científica.
- Redactar objetivos y preguntas de investigación científica.
- Comprender los criterios para evaluar un problema de investigación científica.
- Relacionar el planteamiento del problema con investigaciones bajo los enfoques cuantitativo, cualitativo y mixto.

¿QUÉ ES PLANTEAR EL PROBLEMA DE INVESTIGACIÓN?

Una vez que se ha concebido la idea de investigación y el científico, estudiante o experto han profundizado en el tema en cuestión, se encuentran en condiciones de plantear el problema de investigación.

En realidad, *plantear el problema no es sino afinar y estructurar más formalmente la idea de investigación*. El paso de la idea al planteamiento del problema en ocasiones puede ser inmediato, casi automático, o bien llevar una considerable cantidad de tiempo; ello depende de cuán familiarizado esté el investigador con el tema a tratar, la complejidad misma de la idea, la existencia de estudios antecedentes, el empeño del investigador, el enfoque elegido (cuantitativo, cualitativo o mixto) y sus habilidades personales. Seccionar un tema o una idea no lo coloca inmediatamente en la posición de considerar qué información habrá de recolectar, con cuáles métodos y cómo analizará los datos que obtenga. Antes necesita formular el problema específico en términos concretos y explícitos, de manera que sea susceptible de investigarse con procedimientos científicos (Selltriz *et al.*, 1980).

En el caso del enfoque cualitativo en la investigación, o cuando éste prevalece sobre el cuantitativo, el planteamiento del problema llega a tener lugar en diferentes momentos de la investigación: 1. en este segundo paso que sigue a la generación de la idea de investigación, 2. durante el proceso de investigación (ya sea al consultar la literatura, elaborar el marco teórico, al recolectar información o al analizarla) y 3. al final del proceso investigativo (al redactar el reporte de resultados o el informe terminal).

Bajo el enfoque cualitativo no siempre se requieren términos concretos y explícitos; incluso, a veces es deseable que no sea así.

Cuando mezclamos los enfoques cuantitativo y cualitativo, puede ser que sí necesitemos de tales términos o no: depende del fenómeno estudiado, la manera de enfocar el estudio, el entorno de la investigación y la formación del

investigador, además del modo en que se mezclen los dos enfoques.

Ahora bien, como señala Ackoff (1967): un problema correctamente planteado está parcialmente resuelto; a mayor exactitud corresponden más posibilidades de obtener una solución satisfactoria (aunque, como ya se ha mencionado, no se trata de una regla inflexible, pues en algunos estudios de corte cualitativo lo que se busca precisamente es no tener ideas preconcebidas sobre el fenómeno estudiado, ni definiciones exactas, las cuales se van desarrollando y obteniendo durante el proceso de investigación). De cualquier manera, el investigador debe ser capaz

Planteamiento del problema: significa afinar y estructurar más formalmente la idea de investigación.

no sólo de conceptualizar el problema, sino también de verbalizarlo en forma clara, precisa y accesible. En algunas ocasiones sabe lo que desea hacer pero no cómo comunicarlo a los demás y es necesario que realice un mayor esfuerzo por traducir su pensamiento a términos comprensibles, pues en la actualidad la mayoría de las investigaciones requieren la colaboración de muchas personas.

Criterios para plantear el problema (básicamente para un enfoque cuantitativo o mixto; también para estudios cualitativos que en cualquier parte del proceso de investigación derivan un planteamiento)

Según Kerlinger (2002), los criterios para plantear adecuadamente el problema de investigación son:

- El problema debe expresar una relación entre dos o más variables (recordando que en los estudios cualitativos éste no es un requisito).
- El problema debe estar formulado claramente y sin ambigüedad como pregunta (por ejemplo, ¿qué efecto?, ¿en qué condiciones...?, ¿cuál es la probabilidad de...? ¿cómo se relaciona... con...?) Aunque en los estudios cualitativos la formulación del problema no necesariamente precede a la recolección y al análisis de datos, cuando se llega al punto de plantear el problema de investigación, éste debe formularse con claridad y evitando la ambigüedad (que es muy diferente a la dispersión de datos o apertura en la información).
- El planteamiento debe implicar la posibilidad de realizar una prueba empírica (enfoque cuantitativo) o una recolección de datos (enfoque cualitativo). Es decir, la factibilidad de observarse en la realidad o en un entorno. Por ejemplo, si alguien piensa estudiar cuán sublime es el alma de los adolescentes, está planteando un problema que no puede probarse empíricamente, pues “lo sublime” y “el alma” no son observables. Claro que el ejemplo es extremo, pero nos recuerda que las ciencias trabajan con aspectos observables y medibles en la realidad o en determinado ambiente.

Recordemos que los estudios cualitativos también son empíricos (que es diferente de “empiricistas”).

Una investigación cualitativa predominantemente inductiva recolecta datos en un ambiente, una situación o un evento; y aunque el planteamiento del problema de investigación surge en cualquier fase del proceso, debe existir dicho planteamiento. Lo cualitativo no debe confundirse con lo no científico o con el desorden total en la investigación, pues hay procedimientos y orden, aunque exista apertura y variedad.

¿QUÉ ELEMENTOS CONTIENE EL PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN?

Los elementos para plantear un problema son tres y están relacionados entre sí: *los objetivos que persigue la investigación, las preguntas de investigación y la justificación del estudio*. Esto es independiente del momento en que se desarrolle y se obtenga bajo un esquema deductivo o inductivo.

Objetivos de investigación

En primer lugar, es necesario establecer qué pretende la investigación, es decir, *cuáles son sus objetivos*. Hay investigaciones que buscan, ante todo, contribuir a resolver un problema en especial (en este caso debe mencionarse cuál es y de qué manera se piensa que el estudio ayudará a resolverlo), y otras que tienen como objetivo principal probar una teoría o aportar evidencia empírica en favor de ella. También existen estudios que como resultado final pretenden generar un planteamiento del problema o inducir el conocimiento (en especial los cualitativos).

Los objetivos tienen que expresarse con claridad para evitar posibles desviaciones en el proceso de investigación y deben ser susceptibles de alcanzarse (Rojas, 2001); *son las guías del estudio* y hay que tenerlos presentes durante todo su desarrollo. Evidentemente, los objetivos que se especifiquen requieren ser congruentes entre sí.

Continuando con el ejemplo de la joven interesada en realizar una investigación acerca de los factores que intervienen en el desarrollo del noviazgo. Una vez que se ha familiarizado con el tema encuentra que, según algunos estudios, los factores más importantes son la atracción física, la confianza, la proximidad física, el grado en que cada uno de los novios refuerza positivamente la autoimagen del otro y la similitud entre ambos. Entonces los objetivos de su estudio se plantearían de la siguiente manera:

- Determinar si la atracción física, la confianza, la proximidad física, el reforzamiento de la autoestima y la similitud tienen una influencia importante en el desarrollo del noviazgo entre jóvenes guatemaltecos.
- Evaluar cuáles de los factores mencionados tienen mayor importancia en el desarrollo del noviazgo entre jóvenes guatemaltecos.
 - Analizar si hay o no diferencia entre los hombres y las mujeres con respecto a la importancia atribuida a cada uno de los factores mencionados.
 - Analizar si hay o no diferencias entre las parejas de novios de distintas edades, en relación con la importancia asignada a cada uno de los mismos factores.

■ **Objetivos de investigación:** tienen la finalidad de señalar a lo que se aspira en la investigación y deben expresarse con claridad, pues son las guías del estudio.

También es conveniente comentar que durante la investigación es posible que surjan objetivos adicionales, se modifiquen los objetivos iniciales o incluso se sustituyan por nuevos objetivos, según la dirección que tome la investigación.

En una investigación que se inicie dentro de un esquema cualitativo, la joven puede comenzar con entrevistas a parejas de novios, sin un objetivo particular ni concepciones previas, y, después de entrevistar a varias parejas, comenzar a visualizar lo que le interesa investigar, construyendo inductivamente sus objetivos de estudio.

Preguntas de investigación

Además de definir los objetivos concretos de la investigación, es conveniente plantear, a través de una o varias preguntas, el problema que se estudiará. Plantearlo en forma de preguntas tiene la ventaja de presentarlo de manera directa, minimizando la distorsión (Christensen, 2000).

No siempre en la *pregunta* o las *preguntas* se comunica el problema en su totalidad, con toda su riqueza y contenido. A veces se formula solamente el propósito del estudio, aunque las preguntas deben resumir lo que habrá de ser la investigación. Al respecto, no podemos decir que haya una forma correcta de expresar todos los problemas de investigación, pues cada uno de ellos requiere un análisis particular. Las preguntas generales tienen que aclararse y delimitarse para esbozar el área-problema y sugerir actividades pertinentes para la investigación (Ferman y Levin, 1979), especialmente dentro del enfoque cuantitativo o mixto. Aunque también ello suele ocurrir en algunos estudios cualitativos, ya sea al inicio o durante su desarrollo.

Bajo un esquema deductivo y cuantitativo, hay preguntas demasiado generales que no conducirían a una investigación concreta como: ¿por qué algunos matrimonios duran más que otros?, ¿por qué hay personas más satisfechas con su trabajo que otras?, ¿en qué programas de televisión hay muchas escenas sexuales?, ¿cambian con el tiempo las personas que van a psicoterapia?, ¿los gerentes se comprometen más con su empresa que los obreros?, ¿cómo se relacionan los medios de comunicación colectiva con el voto? Las preguntas no deben utilizar términos ambiguos ni abstractos. Tales preguntas constituyen más bien ideas iniciales que es necesario refinar y precisar para que guíen el comienzo de un estudio.



Una investigación cualitativa respecto de los factores que integran el noviazgo, puede iniciarse con entrevistas generales a varias parejas.

La última pregunta, por ejemplo, habla de “medios de comunicación colectiva”, término que implica la radio, la televisión, los periódicos, las publicaciones, el cine, los anuncios publicitarios en exteriores y otros más. Asimismo, se menciona “voto”, sin especificar el tipo, el contexto ni el sistema social (si se trata de una votación política de nivel nacional o local, sindical, religiosa, para elegir al representante de una cámara industrial o a otro funcionario). Y aun pensando que el voto fuera para una elección presidencial, la relación expresada no lleva a diseñar actividades pertinentes para desarrollar una investigación, a menos que se piense en “un gran estudio” que analice todas las posibles vinculaciones entre ambos términos (medios de comunicación colectiva y voto).

En efecto, como se formula la pregunta, origina una gran cantidad de dudas tales como: ¿se investigarán los efectos que la difusión de propaganda, a través de dichos medios, tiene en la conducta de los votantes?, ¿se analizará el papel de estos medios como agentes de socialización política en cuanto al voto?, ¿se investigará en qué medida se incrementa el número de mensajes políticos en los medios de comunicación masiva durante épocas electorales?, ¿acaso se estudiará cómo los resultados de una votación afectan lo que opinan las personas que manejan esos medios? Es decir, no queda claro qué se va a hacer en realidad.

Lo mismo ocurre con las otras preguntas, son demasiado generales. En lugar de ellas deben plantearse preguntas mucho más específicas como: ¿el tiempo que un matrimonio dedica diariamente a platicar sobre su relación tiene que ver con cuánto tiende a perdurar ésta?, ¿cómo se vinculan la satisfacción laboral y la variedad en el trabajo en la gestión gerencial de grandes empresas industriales en Venezuela?, ¿las series televisivas estadounidenses traducidas al español contienen mayor cantidad de sexo, que las series de telenovelas mexicanas?, ¿conforme se desarrollan las psicoterapias, aumentan o declinan las expresiones verbales de discusión y exploración de planes futuros personales que manifiestan los pacientes?, ¿existe alguna relación entre el nivel jerárquico y la motivación intrínseca en el trabajo en las empresas gubernamentales de Buenos Aires?, ¿cuál es el promedio de horas diarias de televisión que ven los niños colombianos de áreas urbanas?, ¿la exposición por parte de los votantes a los debates televisivos de candidatos a la presidencia de Brasil está correlacionada con la decisión de votar o de abstenerse?

Dentro de un esquema cualitativo es posible que en un primer momento las preguntas sean generales y que paulatinamente vayan precisándose.

Las preguntas pueden ser más o menos generales, como se mencionó anteriormente, pero en la mayoría de los casos es mejor que sean más precisas, sobre todo en el caso de estudiantes que se inician dentro de la investigación. Desde luego, hay macroestudios que investigan muchas dimensiones de un problema y que inicialmente llegan a plantear preguntas más generales. Sin embargo, casi todos los estudios versan sobre cuestiones más específicas y limitadas.

Asimismo, como sugiere Rojas (2001), es necesario establecer los límites temporales y espaciales del estudio, y esbozar un perfil de las unidades de observación (personas, periódicos, viviendas, escuelas, etcétera), perfil que, aunque es tentativo, resulta muy útil para definir el tipo de investigación que habrá de llevarse a cabo. Desde luego, es muy difícil que todos estos aspectos se incluyan en las preguntas de investigación; pero pueden plantearse una o varias preguntas, y acompañarlas de una breve explicación del tiempo, el lugar y las unidades de observación del estudio.

Un asesor en cuestiones de organización decide llevar a cabo un estudio sobre los medios de comunicación que utilizan los altos ejecutivos y plantea las siguientes preguntas de investigación: ¿cuáles son los medios de comunicación que utilizan con mayor frecuencia los niveles gerenciales, o similares, en su trabajo?, ¿qué tipo de información se transmite a través de dichos medios?, ¿con qué propósitos se utiliza cada medio? En estas preguntas no se han especificado diversas cuestiones que es necesario aclarar mediante una breve explicación. En el ejemplo, ésta podría ser la siguiente:

La investigación incluirá las siguientes formas de comunicación en las organizaciones: la interacción diádica "cara a cara", las reuniones en grupos pequeños, el teléfono, la comunicación a través de terceras personas, la correspondencia (cartas, memoranda, avisos, notas, oficios), las reuniones en grandes grupos, los tableros de avisos, las terminales de computadora, el boletín y otras publicaciones de la empresa, y las grabaciones. Se abarcarán solamente los tres niveles jerárquicos más altos de las empresas que cuentan con más de mil trabajadores del área metropolitana de la ciudad de Bogotá.

Un ejemplo de ingeniería sería averiguar cuáles son los factores que inciden en la rotura de envases de vidrio para refrescos embotellados.

Aplicando lo anterior al ejemplo de la investigación sobre el noviazgo, las preguntas de investigación podrían ser:

- ¿La atracción física, la confianza, la proximidad física, el reforzamiento de la autoestima y la similitud ejercen una influencia significativa sobre la evaluación que hacen los novios de su relación, el interés mostrado por ésta y la disposición de continuar la relación?

- ¿Cuál de estos factores ejerce mayor influencia sobre la evaluación de la relación, el interés mostrado por ésta y la disposición de continuar la relación?
- ¿Están vinculados entre sí la atracción física, la confianza, la proximidad física, el reforzamiento de la autoestima y la similitud?
- ¿Existe alguna diferencia entre los hombres y las mujeres con respecto al peso que le asignan a cada factor en la evaluación de la relación, el interés mostrado por ésta y la disposición de continuar la relación?
- ¿La edad está relacionada con el peso asignado a cada factor con respecto a la evaluación de la relación, el interés mostrado por ésta y la disposición de continuar la relación?

Ahora bien, con una simple ojeada al tema nos daríamos cuenta de que se pretende abarcar demasiado en el problema de investigación y, a menos que se cuente con muchos recursos y tiempo, se tendría que limitar el estudio, por ejemplo, a la similitud. Entonces se preguntaría: ¿la similitud ejerce alguna influencia significativa sobre la elección de la pareja en el noviazgo y la satisfacción dentro de él?

■ **Preguntas de investigación:** orientan hacia las respuestas que se buscan con la investigación.

En el estudio que hemos comentado acerca del desarrollo municipal las preguntas serían: ¿con qué infraestructura cuentan ya los municipios estudiados?, ¿qué se puede ofrecer al inversionista en cuanto a infraestructura, ventajas impositivas y disposición de recursos humanos?, ¿qué normas las leyes locales que regulan el desarrollo municipal?, ¿de qué

recursos locales, regionales y nacionales se dispondría para impulsar el desarrollo municipal?

Al igual que en el caso de los objetivos, durante el desarrollo de la investigación pueden modificarse las preguntas originales o agregarse otras nuevas; y como se ha venido sugiriendo, la mayoría de los estudios plantean más de una pregunta, ya que de este modo se cubren diversos aspectos del problema a investigar.

El enfoque cualitativo y las preguntas de investigación

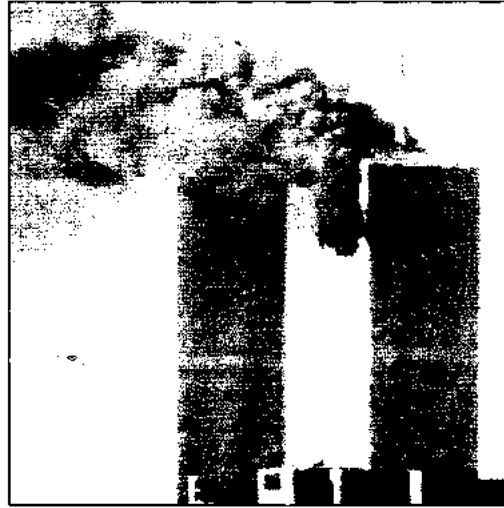
Algunas veces los estudios cualitativos siguen el esquema ya comentado para generar las preguntas de investigación, buscando que acoten lo que se pretende inquirir; aunque en otras, la recolección y el análisis de datos pueden utilizarse

para descubrir cuáles son las preguntas de investigación relevantes, para más tarde “refinarlas” y agregar precisión a dichas preguntas.

Se han desarrollado algunos estudios cualitativos, porque no están de acuerdo con las investigaciones antecedentes ni con sus preguntas de investigación. Esto puede ser válido y en ocasiones ha sido benéfico para el avance del conocimiento. Por ejemplo, un investigador puede creer que una explosión o un desastre provocado (como el choque de los aviones de pasajeros contra las Torres Gemelas de Nueva York, el 11 de septiembre de 2001; o la explosión de pólvora en la Ciudad de Celaya, México, en 1999) generará secuelas psicológicas entre los sobrevivientes, y en lugar de basarse en la literatura sobre ansiedad posteventos, traumas específicos y miedos detectados por estudios antecedentes, iniciaría su estudio sin predeterminedar variables ni hipotetizar efectos, sino tan sólo conviviendo con los sobrevivientes y analizando su conducta, incluso entrevistándolos y, después de revisar cuidadosamente sus observaciones, desarrollar el planteamiento del problema (objetivos y preguntas) para una segunda etapa del estudio, o para futuras investigaciones.

Justificación de la investigación

Además de los objetivos y las preguntas de investigación, *es necesario justificar el estudio exponiendo sus razones*. La mayoría de las investigaciones se efectúan con un propósito definido, no se hacen simplemente por capricho de una persona; y ese propósito debe ser lo suficientemente fuerte para que se justifique su realización. Además, en muchos casos se tiene que explicar por qué es conveniente llevar a cabo la investigación y cuáles son los beneficios que se derivarán de ella: el pasante deberá explicar a un comité escolar el valor de la tesis que piensa realizar, el investigador universitario hará lo mismo con el grupo de personas que en su institución aprueban proyectos de investigación e incluso con sus colegas, el asesor tendrá que explicar a su cliente las recompensas que se obtendrán de un estudio determinado, el subordinado que propone una investigación a su superior deberá dar razones de la utilidad de ella. Lo mismo ocurre en casi todos los casos. Trátase de estudios cuantitativos, cualitativos o mixtos, siempre es importante dicha justificación.



Se han realizado muchas investigaciones respecto de las secuelas psicológicas entre los sobrevivientes de los atentados del 11 de septiembre en Nueva York.

Criterios para evaluar el valor potencial de una investigación

Una investigación llega a ser conveniente por diversos motivos: tal vez ayude a resolver un problema social, a construir una nueva teoría o a generar preguntas de investigación. Lo que algunos consideran relevante para investigarse puede no serlo para otros. Respecto de ello, suele diferir la opinión de las personas. Sin embargo, es posible establecer una serie de criterios para evaluar la utilidad de un estudio propuesto, los cuales, evidentemente, son flexibles y de ninguna manera son exhaustivos. A continuación se indican algunos de estos criterios formulados como preguntas, los cuales fueron adaptados de Ackoff (1967) y Miller (2002). Y afirmaremos que cuanto mayor número de respuestas se contesten positiva y satisfactoriamente, la investigación tendrá bases más sólidas para justificar su realización.

Justificación de la investigación: indica el porqué de la investigación exponiendo sus razones.

- *Conveniencia* ¿Qué tan conveniente es la investigación?; esto es, ¿para qué sirve?
- *Relevancia social* ¿Cuál es su trascendencia para la sociedad?; ¿quiénes se beneficiarán con los resultados de la investigación?; ¿de qué modo? En resumen, ¿qué alcance social tiene?
- *Implicaciones prácticas* ¿Ayudará a resolver algún problema real?; ¿tiene implicaciones trascendentales para una amplia gama de problemas prácticos?
- *Valor teórico* Con la investigación, ¿se llenará algún hueco de conocimiento?; ¿se podrán generalizar los resultados a principios más amplios?; ¿la información que se obtenga puede servir para comentar, desarrollar o apoyar una teoría?; ¿se podrá conocer en mayor medida el comportamiento de una o de diversas variables o la relación entre ellas?; ¿ofrece la posibilidad de una exploración fructífera de algún fenómeno o ambiente?; ¿qué se espera saber con los resultados que no se conociera antes?; ¿puede sugerir ideas, recomendaciones o hipótesis a futuros estudios?
- *Utilidad metodológica* ¿La investigación puede ayudar a crear un nuevo instrumento para recolectar o analizar datos?; ¿ayuda a la definición de un concepto, ambiente, contexto variable o relación entre variables?; ¿pueden lograrse con ella mejoras en la forma de experimentar con una o más variables?; ¿sugiere cómo estudiar más adecuadamente una población?; ¿puede ayudar a mezclar los enfoques cuantitativos y cualitativos para enriquecer la búsqueda de conocimiento?

Desde luego, es muy difícil que una investigación pueda responder positivamente a todas estas interrogantes; algunas veces sólo cumple un criterio.

Viabilidad de la investigación

Además de los tres elementos que conforman propiamente el planteamiento del problema, es necesario considerar otro aspecto importante: la *viabilidad* o *factibilidad* misma del estudio; para ello, debemos tomar en cuenta la disponibilidad de recursos financieros, humanos y materiales que determinarán, en última instancia, los alcances de la investigación (Rojas, 2001). Es decir, tenemos que preguntarnos realísticamente: ¿es factible llevar a cabo esta investigación? y ¿cuánto tiempo tomará realizarla? Dichos cuestionamientos son particularmente importantes cuando se sabe de antemano que se dispondrá de pocos recursos para efectuar la investigación.

Viabilidad de la investigación: factibilidad en la realización de un estudio en cuanto a la disponibilidad de recursos.

UN EJEMPLO DE INVIABILIDAD

Un caso ilustrativo de este hecho ocurrió hace algunos años, cuando un grupo de estudiantes de Ciencias de la Comunicación decidió realizar su tesis de licenciatura sobre el impacto que tendría introducir la televisión en una comunidad donde no existía. El estudio buscaba, entre otras cosas, analizar si los patrones de consumo cambiaban, las relaciones interpersonales se modificaban, y las actitudes y los valores centrales de los habitantes (religión; actitudes hacia el matrimonio, la familia, la planificación familiar, el trabajo) se transformaban con la introducción de la televisión. La investigación resultaba interesante porque había pocos estudios similares y éste aportaría información útil para el análisis de los efectos de tal medio, la difusión de innovaciones y otras muchas áreas de conocimiento. Sin embargo, el costo de la investigación era muy elevado (había que adquirir muchos televisores y obsequiarlos a los habitantes o rentarlos, hacer llegar a la comunidad las transmisiones, contratar a bastante personal, hacer considerables erogaciones en viáticos, etcétera), y superaba, por mucho, las posibilidades económicas de los estudiantes, aun cuando consiguieran financiamiento. Además, llevaría bastante tiempo realizarlo (cerca de tres años), tomando en cuenta que se trataba de una tesis. Posiblemente para un investigador especializado en el área, este tiempo no resultaría un obstáculo. El factor "tiempo" varía en cada investigación; a veces se requieren los datos en el corto plazo, mientras que en otras ocasiones el tiempo no es relevante. Hay estudios que duran varios años porque su naturaleza así lo exige.

Consecuencias de la investigación

Aunque no sea con fines científicos, es necesario que el investigador se cuestione acerca de *las consecuencias de su estudio*. En el ejemplo anterior, suponiendo que

se hubiera efectuado la investigación, resultaría conveniente preguntarse antes de realizarla: ¿cómo va a afectar a los habitantes de esa comunidad?

Imaginemos que se piensa realizar un estudio sobre el efecto de una droga muy fuerte, que se usa en el tratamiento de alguna clase de esquizofrenia. Cabría reflexionar sobre la conveniencia de efectuar o no la investigación. Y ello no contradice el postulado de que la investigación científica no estudia aspectos morales ni formula juicios de este tipo. No lo hace, pero ello no significa que un investigador no pueda decidir si realiza o no un estudio porque ocasionaría efectos perjudiciales para otros seres humanos. Aquí se está hablando de suspender una investigación por cuestiones de ética personal, y de no llevar a cabo un estudio por cuestiones éticas o estéticas. La decisión de realizar o no una investigación por las consecuencias que ésta pueda acarrear es una decisión personal de quien la concibe. Desde el punto de vista de los autores, también es un aspecto del planteamiento del problema que debe ventilarse, y la responsabilidad es algo muy digno de tomarse en cuenta siempre que se va a realizar un estudio.

La investigación sobre clonación plantea retos interesantes a este respecto.

Consecuencias de la investigación:
repercusiones positivas o negativas que el estudio implica en los ámbitos ético y estético.



- Plantear el problema de investigación consiste en afinar y estructurar más formalmente la idea de investigación, desarrollando tres elementos de la investigación: objetivos, preguntas y justificación.
- En la investigación cuantitativa los tres elementos deben ser capaces de conducir hacia una investigación concreta y con posibilidad de prueba empírica.
- En el enfoque cuantitativo el planteamiento del problema de investigación precede a la revisión de la literatura y al resto del proceso de investigación; aunque esta revisión pueda modificar el planteamiento original.
- En el enfoque cualitativo, el planteamiento del problema llega a surgir en cualquier momento de la investigación, incluso al principio o al final.
- En ocasiones, en el enfoque cualitativo, la recolección y el análisis de los datos contribuyen a plantear las preguntas de investigación.
 - Los objetivos y las preguntas de investigación deben ser congruentes entre sí e ir en la misma dirección.
 - Los objetivos establecen qué pretende la investigación, las preguntas nos dicen qué respuestas deben encontrarse mediante la investigación y la justificación nos indica por qué debe hacerse la investigación.
 - Los criterios principales para evaluar el valor potencial de una investigación son: conveniencia, relevancia social, implicaciones prácticas, valor teórico y utilidad metodológica. Además, debe analizarse la viabilidad de la investigación y sus posibles consecuencias.

- El planteamiento de un problema de investigación científica no puede incluir juicios morales ni estéticos. Pero debe cuestionarse si es o no ético llevarlo a cabo.

Consecuencias de la investigación	Enfoque cuantitativo	Planteamiento del problema
Criterios para evaluar una investigación	Justificación de la investigación	Preguntas de investigación
Enfoque cualitativo	Objetivos de investigación	Viabilidad de la investigación



1. Vea una película sobre estudiantes (de nivel medio o superior) y su vida cotidiana, deduzca una idea, después consulte algunos libros o artículos que hablen sobre esa idea y, finalmente, plantee un problema de investigación en torno a dicha idea (objetivos, preguntas y justificación de la investigación).
2. Seleccione un artículo de una revista científica que contenga los resultados de una investigación y responda las siguientes preguntas: ¿cuáles son los objetivos de esa investigación?, ¿cuáles son las preguntas?, ¿cuál es su justificación?
3. Visite una comunidad rural y observe qué sucede en tal comunidad, platique con sus habitantes y recolecte información sobre un asunto que le interese. Tome notas y analícelas. De esta experiencia, plantee un problema de investigación.
4. Respecto de la idea que eligió en el capítulo 2, transfórmela en un planteamiento del problema de investigación. Pregúntese: ¿los objetivos son claros, precisos y llevarán a la realización de una investigación en la "realidad"? ¿son ambiguas las preguntas?, ¿qué va a lograrse con este planteamiento?, ¿es posible realizar esa investigación? Además, evalúe su planteamiento de acuerdo con los criterios expuestos en este capítulo.
5. Compare los siguientes objetivos y preguntas de investigación. ¿Cuál de ambos planteamientos es más específico y claro?, ¿cuál piensa que es mejor?

Planteamiento 1

Objetivo: Analizar el efecto de utilizar a un profesor autocrático *versus* un profesor democrático, en el aprendizaje de conceptos matemáticos elementales en niños de escuelas públicas ubicadas en zonas rurales. El estudio se realizaría con niños que asisten a su primer curso de matemáticas.

Pregunta: ¿El estilo de liderazgo (democrático-autocrático) del profesor se encuentra relacionado con el nivel de aprendizaje de conceptos matemáticos elementales?

Planteamiento 2

Objetivo: Analizar las variables que se relacionen con el proceso de enseñanza-aprendizaje de los niños en edad preescolar.

Pregunta: ¿Cuáles son las variables que se relacionan con el proceso de enseñanza-aprendizaje?

¿Cree que el segundo planteamiento es demasiado global? Y respecto al primero, ¿podría mejorarse? Y si es así, ¿de qué manera?

6. Algunos calificativos que no se aceptan en un planteamiento de un problema de investigación son:

Ambiguo	Vago
Confuso	Ininteligible
General	Incomprensible
Vasto	Desorganizado
Injustificable	Incoherente
Irracional	Inconsistente
Prejuicioso	

¿Qué otros calificativos no puede aceptar un problema de investigación?



HURTADO, J. (1998), *Metodología de la investigación holística*, (2a. ed.), Caracas: SYPAL.
KERLINGER, F. N., y LEE, H. B. (2002), *Investigación del comportamiento: métodos de investigación en ciencias sociales*, México: McGraw-Hill Interamericana Editores, pp. 16-28.

LA TELEVISIÓN Y EL NIÑO

OBJETIVOS

1. Describir el uso que los niños de la ciudad de México hacen de los medios de comunicación colectiva.
2. Indagar el tiempo que los niños de la ciudad de México dedican a ver la televisión.
3. Describir cuáles son los programas preferidos de los niños de la ciudad de México.
4. Determinar las funciones y gratificaciones de la televisión para el niño de la ciudad de México.
5. Conocer el tipo de control que ejercen los padres sobre la actividad de ver televisión de sus hijos.
6. Analizar qué tipos de niños ven más televisión.

PREGUNTAS DE INVESTIGACIÓN

¿Cuál es el uso que los niños de la ciudad de México hacen de los medios de comunicación colectiva?

¿Cuánto tiempo dedican a ver la televisión diferentes tipos de niños?

¿Cuáles son los programas preferidos de dichos niños?

¿Cuáles son las funciones y gratificaciones de la televisión para el niño?

¿Qué tipo de control ejercen los padres sobre sus hijos en relación con la actividad de ver televisión?

JUSTIFICACIÓN

Para la mayoría de los niños ver televisión, dormir e ir a la escuela constituyen sus principales actividades. Asimismo, la televisión es el medio de comunicación preferido por los pequeños. Se estima que, en promedio, diariamente el niño ve televisión tres horas, y se calculó en un reporte de la agencia de investigación Nielsen que, al cumplir los 15 años, un niño ha visto cerca de 15 000 horas de contenidos televisivos. Este hecho ha generado diversos cuestionamientos de padres, maestros, investigadores y, en general, de la sociedad sobre la relación niño-televisión, y los efectos de ésta sobre el infante. Así, se ha considerado trascendente estudiar dicha relación, con el propósito de analizar el papel que en la vida del niño desempeña un agente de socialización tan relevante como la televisión. El estudio planteado ayudará, entre otros aspectos, a conocer la relación niño-televisión, sus implicaciones para el desarrollo del niño, y proporcionará información que será útil para padres y maestros sobre cómo manejar de modo más provechoso la relación del niño con la televisión.

Por otra parte, la investigación contribuiría a contrastar, con datos de México, los datos sobre usos y gratificaciones de la televisión en el niño encontrados en otros países.

La investigación es viable, pues se dispone de los recursos necesarios para llevarla a cabo.

EL CONTAGIO DEL SIDA

OBJETIVOS

Conocer la evolución y evaluar la situación actual de los individuos que recibieron sangre o derivados con la posibilidad de estar contaminados (factor de riesgo) adquiridos de "Transfusiones y Hematología, S.A." por los Servicios Médicos de Pemex, en el periodo de enero de 1984 a mayo de 1987 y de sus contactos, con el propósito de tomar las medidas preventivas necesarias para interrumpir la cadena de transmisión y propagación del virus de inmunodeficiencia humana (VIH), así como para fundamentar el manejo administrativo y laboral en los casos de trabajadores que hubieran resultado afectados (*SIDA con factor de riesgo postransfusión, receptores y contactos*, Hernández Galicia, Gerencia de Servicios Médicos de la Subdirección Técnica Administrativa de Petróleos Mexicanos, 1989).

CONTEXTO

"Un número por determinar de pacientes atendidos entre enero de 1984 a mayo de 1987, en las unidades hospitalarias de Petróleos Mexicanos, así como de sus contactos directos, son portadores del virus de inmunodeficiencia humana (VIH) debido a que por requerimientos propios de su padecimiento recibieron transfusiones de sangre o sus derivados, posiblemente contaminados, provenientes del banco particular Transfusiones y Hematología, S.A."

PREGUNTAS DE INVESTIGACIÓN

¿Habrá diferencia en el tiempo que se tarda en desarrollar el SIDA entre las personas contagiadas por transmisión sexual contra las contagiadas por transfusión sanguínea?

¿Se encontrarán diferencias en el tiempo que tardan en desarrollar la enfermedad personas que pertenecen a grupos con diferentes edades?

¿Existen síntomas que evidencien de manera diferente el desarrollo de la enfermedad entre los contagiados por transfusión y los contagiados por transmisión sexual?

JUSTIFICACIÓN

Es necesario estudiar los efectos que tienen las diferentes formas de adquirir el VIH para establecer medidas que limiten y controlen las posibilidades de transmisión del SIDA. Lo anterior llevará a implantar acciones en hospitales y clínicas que eviten la transfusión de sangre sin el control requerido.

La investigación es conveniente desde el punto de vista médico y contribuirá al conocimiento del SIDA. Además, permitirá un manejo justo y equitativo en casos de infección por el VIH.

Ejemplo de enfoque mixto (cuantitativo y cualitativo)

ESTUDIO DE LA MODA EN LAS MUJERES MEXICANAS

CONTEXTO

Un grupo de mercadólogos fue contratado por una empresa para realizar un estudio sobre las tendencias de la moda entre mujeres latinoamericanas. Básicamente, la empresa (una gran cadena de tiendas departamentales con un área dedicada a la ropa femenina para adolescentes y adultos) deseaba conocer cómo define la moda la mujer mexicana, qué elementos implica la moda desde su perspectiva, cómo evalúan las secciones del departamento de ropa para damas y qué es importante que la tienda haga por sus clientas.

Los investigadores, con un conocimiento mínimo sobre la moda femenina, decidieron iniciar la investigación de manera inductiva y cualitativa; sin un planteamiento del problema definido ni estructurado, y mucho menos con una hipótesis.

El primer paso consistió en enviar a un grupo de mujeres entrenadas para observar de manera no obstrusiva a las personas que llegaban al departamento de ropa para damas (haciéndose pasar por clientas). No se estructuró una guía de observación, tan sólo se les indicó que registraran el comportamiento que percibieran de las clientas (lo que ellas vieran). Las observadoras tomaron nota de una amplia variedad de comportamientos verbales y no verbales (desde el tiempo que permanecían en dicho departamento hasta qué objetos, tipo de ropa y partes del área les llamaban más la atención; qué les emocionaba; los colores y modelos que se probaban y compraban); perfiles físicos (aproximadamente de qué edades, tipo de vestimenta, si venían solas o acompañadas y, en este último caso, de quién). La observación se prolongó durante una semana.

Tales registros y observaciones les sirvieron a los investigadores para comenzar a definir las áreas temáticas que podía contener el estudio y para elaborar una guía general de entrevistas (con preguntas amplias como: ¿qué es la moda?, ¿cómo se define estar a la moda?, ¿qué es lo más importante para ser una mujer que se vista a la moda?, etcétera). Posteriormente, el grupo de observadoras capacitadas realizó entrevistas a las personas que visitaban el departamento de damas y otras investigadoras se hicieron pasar por vendedoras de la tienda. No se determinó ningún tamaño de muestra, sólo se les pidió que entrevistaran de manera informal a las clientas. (Finalmente se llevaron a cabo 213 entrevistas.)

Después se realizaron entrevistas en profundidad con mujeres de diferentes edades (desde los 14 hasta los 65 años) en sus propios hogares, para conversar sobre moda, gustos, marcas favoritas y, de manera general, sobre cómo percibían a la tienda, entre otras cuestiones.

En primera instancia, todo el cúmulo de información obtenido se analizó de forma individual, por cada investigador, y después en grupo (material producto de observaciones, entrevistas y pláticas que tuvo el personal de campo en su papel de vendedoras). Tal análisis siguió las técnicas cualitativas que se revisarán más adelante

dentro del libro. A raíz de dichas experiencias, finalmente se planteó el problema de investigación (tras haber recolectado información y efectuado el análisis)

El problema abarcó como objetivos:

- Obtener las definiciones y percepciones de la moda para las mujeres mexicanas.
- Determinar qué factores componen la definición de moda para las mujeres mexicanas.
- Conocer el significado de "estar a la moda" entre las mujeres mexicanas.
 - Precisar qué características tienen las prendas y los accesorios que se consideran "a la moda" para dichas mujeres.
 - Evaluar qué comportamientos manifiestan tales mujeres al comprar ropa.
 - Obtener un perfil ideal (naturaleza, características y atributos) de un departamento o una tienda de ropa femenina.
 - Conocer qué tiendas prefieren las mujeres mexicanas para comprar ropa.
 - Evaluar el departamento de damas de una tienda (incluyendo sus secciones).

Entre algunas de las preguntas de investigación que se establecieron estaban: ¿qué es la moda para las mujeres mexicanas?, ¿qué significa "estar a la moda" para ellas?, ¿qué dimensiones integran dicho concepto de moda?, ¿qué marcas, tipo de prendas, colores y estilos prefieren las mexicanas?, ¿qué atributos debe tener un departamento o una tienda de ropa para damas?, ¿cómo evalúan al departamento de ropa para damas?

La justificación incluyó la necesidad que tenía la cadena de tiendas departamentales de conocer mejor el pensamiento de sus clientes femeninos y así mantenerse a la vanguardia, ante la creciente competencia local e internacional en el mercado de ropa para mujer.

Así, un estudio se planteó con dos vertientes: cuantitativa y cualitativa, situación que se desarrollará a lo largo del texto.



Creo que debemos hacerles ver a los estudiantes que comprender el método científico no es difícil y que, por lo tanto, investigar la realidad tampoco lo es. La investigación bien utilizada es una valiosa herramienta del profesional en cualquier área; no hay mejor forma de plantear soluciones eficientes y creativas para los problemas que tener conocimientos profundos acerca de la situación. También hay que hacerles comprender que la teoría y la realidad no son polos opuestos, sino que están totalmente relacionados.

Un problema de investigación bien planteado es la llave de la puerta de entrada al trabajo en general, pues de esta manera permite la precisión en los límites de la investigación, la organización adecuada del marco teórico y las relaciones entre

las variables; en consecuencia, es posible llegar a resolver el problema y generar datos relevantes para interpretar la realidad que se desea aclarar.

En un mismo estudio es posible combinar diferentes enfoques; también estrategias y diseños, puesto que se puede estudiar un problema cuantitativamente y, a la vez, entrar a niveles de mayor profundidad por medio de las estrategias de los estudios cualitativos. Se trata de un excelente modo de estudiar las complejas realidades del comportamiento social.

En cuanto a los avances que se han logrado en investigación cuantitativa, destaca la creación de instrumentos para medir una serie de fenómenos psicosociales que hasta hace poco se consideraban imposibles de abordar científicamente. Por otro lado, el desarrollo y uso masivo de la computadora en la investigación ha propiciado que se facilite el uso de diseños, con los cuales es posible estudiar múltiples influencias sobre una o más variables. Lo anterior acercó la compleja realidad social a la teoría científica.

La investigación cualitativa se ha consolidado, enmarcando sus límites y posibilidades; asimismo, han avanzado sus técnicas para recopilar datos y manejar situaciones propias. Al mismo tiempo, con este modelo se logra estudiar cuestiones que no es factible analizar por medio del enfoque cuantitativo.

Aunque resulta difícil precisar los parámetros de una buena investigación, es claro que se caracteriza por la relación armónica entre los elementos de su estructura interna; además, por su novedad, importancia social y utilidad. Lo único que no es recomendable en la actividad científica es que el investigador actúe en forma negligente.

EDWIN SALUSTIO SALAS BLAS
Facultad de Psicología
Universidad de Lima
Lima, Perú

La información más completa es la más cercana; por lo tanto, los estudiantes que no saben por dónde iniciar su proyecto de investigación deben recurrir a las fuentes más próximas a ellos, como son los problemas sociales, el desarrollo tecnológico interno e, incluso, los puntos de vista de latinoamericanos acerca del arte o la cultura.

La investigación no tiene que ser obligatoria ni realizarse sin entusiasmo, de ahí la importancia de que el tema elegido sea de interés para el estudiante y que, al mismo tiempo, haya en él un compromiso de iniciar y concluir con la misma fuerza.

El planteamiento del problema es el aspecto más importante para comenzar un proyecto, ya que no es posible obtener un buen resultado si antes no se determina lo que se pretende conseguir. Efectuar bien esta primera actividad hará de la investigación una labor más fluida y objetiva.

FELIPE ORTIZ VÁZQUEZ
Profesor de tiempo completo
Departamento de Contaduría
Tecnológico de Estudios Superiores de Ecatepec
Estado de México, México

Lectura 2

Introducción

La lucha de la luz y la sombra

Luis Jesús Galindo Cáceres

Sociedad, cultura y comunicación. Un punto de partida

Por una perspectiva teórico-metodológica. Heurística, ciencia y comunicación

¡OBSERVAR AL MUNDO CONTEMPORÁNEO SUPONE UN esfuerzo de visión total de una ambición soberbia. Tanta información sobre una complejidad organizada en variados planos y niveles, multidimensional, polifónica y ante ella, la pretensión de saberlo todo de alguna manera. ¡Esta es una de las puntas del conocimiento sobre lo social, la que pretende estar por encima del mundo para observarlo más allá del tiempo y la historia, del espacio y la geografía, por fuera de la vida, desde un no lugar que una supuesta ciencia permite, el ejercicio máximo y definitivo de la alteridad, el programa de distanciamiento supremo que es coronado con el conocimiento de la globalidad. Mucho talento y energía se han invertido en este suspiro teológico de conocer la verdad final y determinante. Y por otra parte, el dentro de la vida, la resolución de lo que la percepción puede tocar y sentir desde el plano inmediato de la vida cotidiana. Aquí se presenta un proceso distinto al distanciamiento, casi el mimetismo, la búsqueda de la identidad con el entorno y sus aristas que lastiman y pueden ser limadas o evitadas. Otro programa, en el que los mundos se multiplican, porque el primero es pequeño y cercano y permite distinguir otros muchos, más allá de sus familiares fronteras. Y entre ambos programas, una multitud de paradojas y entusiasmos, de deseos y miedos, de aspiraciones y frustración.

Las miradas que observan a los mundos contemporáneos son muy diversas, las formas que las agrupan son parte del ejercicio de la mirada que mira a las miradas. En este escenario las apuestas también son variadas y urgentes. Un ramillete de opciones aparece ante el observador en busca de alternativas de atención. Este tipo de observación, de segundo orden, es el campo estricto de la reflexión metodológica. Y desde ahí opera el vuelo hacia un tercero o cuarto orden, la mirada que mira a la mirada que mira a la mirada que mira, y un aterrizaje en el primer orden, la pregunta por la mirada directa y sus condiciones de operación. Hoy en

Tomada de: Luis Jesús Galindo, "La lucha de la luz y la sombra", en *Técnicas de investigación en sociedad, cultura y comunicación*, México, Pearson, 1998, pp. 9-25.

día el campo está abierto, las alternativas reflexivas sobre la observación se configuran y reconfiguran, la labor constructiva sobre la mente que elabora mundos es el corazón de la acción teórico-metodológica. Y entonces, hoy como nunca antes, la intención que indaga se accecha a sí misma en forma permanente, creando, imaginando, ajustando, aprendiendo, volviendo a empezar. Pensar la metodología es hoy más complejo y profundo que antes, más diverso, intenso, estético, lúdico.

Los caminos para indagar dependen de la intención particular del sujeto de la acción, y del contexto de referencia en que ese sujeto se inscribe consciente o inconscientemente, voluntaria o involuntariamente. Así, algunos caminos se configuran en un acuerdo denominado ciencia, otros en uno llamado arte, religión, magia, o algún otro. Cada acuerdo tiene sus reglas del juego, su forma básica de acción y variantes cercanas y lejanas, estilos, ortodoxias y heterodoxias, comportamientos subordinados a las normas, y comportamientos de ruptura o herejía. Como sea el caso, todos ellos suponen un principio de autoridad, una cierta ley, y un ámbito de intersubjetividad y de puesta en escena de las normas y valores.

Es difícil asumir de pronto una postura ante el campo metodológico de la investigación de lo social sin el riesgo de un cierre excesivo de reflexividad. De hecho, eso es lo que sucede en cualquier acción concreta de indagación. Es una regla que la acción significa el cierre momentáneo de información y reflexividad que, de inmediato, puede abrir o cerrar posibilidades según el tipo de operación efectuada. La relación entre ciencia y heurística es una gran ayuda para configurar esta situación.

La ciencia aparece en este ducto con el papel del cierre, la imagen que el positivismo le ha ayudado a construir. Su puesta en escena es muy formal y requiere ser observable en todo momento, para ser refutada y replicada. Este ejercicio es muy importante, permite que la reflexividad componga operaciones de alteridad observable, el que indaga necesita darse cuenta de lo que está haciendo todo el tiempo, lo logre o no. Esta configuración conforma un edificio de nociones y operaciones visibles a los ojos del investigador y de cualquier otro. De ahí que todo curso de acción reflexivo sea ajustable, mejorable, explicitable. El método es el camino general de este movimiento, y las técnicas son las operaciones concretas en cada paso. La teoría sería la proveedora de las metáforas que permiten conducir al camino y poner en forma la información obtenida técnicamente.

La heurística es abierta, busca indagar de un modo más libre. En cierto sentido es la base de la ciencia y su continuación. Es menos evidente en todas sus operaciones, respeta la intuición. Supone un ejercicio de creatividad semejante al arte. Es la que ensaya los nuevos caminos. Es una exploradora constante. Ciencia y heurística se necesitan, una más jurídica, otra más artística, y ambas enlazadas en la interacción entre indagadores y procesos, entre intenciones y resultados, entre descubrimientos y catálogos, inniscuidas en la trama y la urdimbre de la comunicación.

*Exploración del espacio conceptual y el árbol de búsqueda.
La investigación como proceso creativo*

El reto básico en la investigación es la creatividad, la capacidad de configurar posibilidades a partir de posibilidades. El pensamiento metodológico y tecnológico tradicional tiene más funciones de administración y control que de movimiento y desarrollo de trayectorias. Bajo la forma conservadora de la investigación, el mundo aparece como único y percible en aproximaciones sucesivas que acercan a su conocimiento total y definitivo. El orden preexiste al acto perceptivo y la ciencia lo único que hace es descubrirlo, explicitarlo. Todo está configurado en forma fija y estable con ciertas variantes que pueden ser previsibles bajo estrictas condiciones de manipulación e intervención en las causas descubiertas. El punto es que no importa en sentido estricto que el asunto sea así o no, lo importante son las consecuencias de esta visión. Si lo que aparece es percibido como estático, será estático en las operaciones y sentidos que se le atribuyan y con los cuales sea concebido. La perspectiva perceptual configura una estructura cognitiva que construye lo observado según sus parámetros de observación. De la misma forma, una perspectiva distinta también opera sobre la percepción y la acción en forma consistente. Dime cómo está configurada tu percepción y te diré cómo actúas y observas.

El peligro de la forma conservadora de investigación es que inhibe a la percepción de su potencial configurador creativo. Y de este modo, lo alterno no sólo no aparece sino que es impertinente a la imaginación. Todo el referente perceptual se estabiliza en patrones de memoria que fijan y reproducen formas de relación mente-mundo. En el caso de las ciencias sociales el efecto es tremendo, la tendencia es a mantener ciertos patrones de configuración y a empobrecer el cambio.

La investigación es un proceso de creatividad reflexivo. Es decir, al investigar acontece lo mismo que en el arte o en cualquier otra actividad creadora, pero con una diferencia sustantiva, el autor creador se observa con atención durante el movimiento de su intención a través del espacio conceptual e imaginario durante la acción indagadora. El investigador es un creador altamente reflexivo, un observador que nunca pierde detalle de lo que le sucede a su interior y de lo que acontece en su exterior.

La creatividad consiste, en pocas palabras, y según una imagen configurada en la indagación sobre el tema de la ciencia de la inteligencia artificial, en la exploración de un espacio conceptual a través de cierto árbol de búsqueda. El espacio conceptual es el ámbito de las imágenes y conceptos de la cultura que configuran la ecología de información y comunicación del sujeto explorador. El investigador social explora en diversos ámbitos del espacio conceptual posible, lo hace desde el mundo semicerrado del campo académico-científico, y se mueve hacia la percepción y construcción cognitiva de mundos sociales construidos por diversos actores y ámbitos colectivos. En este sentido, aunque su punto de partida es el

*Exploración del espacio conceptual y el árbol de búsqueda.
La investigación como proceso creativo*

El reto básico en la investigación es la creatividad, la capacidad de configurar posibilidades a partir de posibilidades. El pensamiento metodológico y tecnológico tradicional tiene más funciones de administración y control que de movimiento y desarrollo de trayectorias. Bajo la forma conservadora de la investigación, el mundo aparece como único y percible en aproximaciones sucesivas que acercan a su conocimiento total y definitivo. El orden preexiste al acto perceptivo y la ciencia lo único que hace es descubrirlo, explicitarlo. Todo está configurado en forma fija y estable con ciertas variantes que pueden ser previsibles bajo estrictas condiciones de manipulación e intervención en las causas descubiertas. El punto es que no importa en sentido estricto que el asunto sea así o no, lo importante son las consecuencias de esta visión. Si lo que aparece es percibido como estático, será estático en las operaciones y sentidos que se le atribuyan y con los cuales sea concebido. La perspectiva perceptual configura una estructura cognitiva que construye lo observado según sus parámetros de observación. De la misma forma, una perspectiva distinta también opera sobre la percepción y la acción en forma consistente. Dime cómo está configurada tu percepción y te diré cómo actúas y observas.

El peligro de la forma conservadora de investigación es que inhibe a la percepción de su potencial configurador creativo. Y de este modo, lo alterno no sólo no aparece sino que es impertinente a la imaginación. Todo el referente perceptual se estabiliza en patrones de memoria que fijan y reproducen formas de relación mente-mundo. En el caso de las ciencias sociales el efecto es tremendo, la tendencia es a mantener ciertos patrones de configuración y a empobrecer el cambio.

La investigación es un proceso de creatividad reflexivo. Es decir, al investigar acontece lo mismo que en el arte o en cualquier otra actividad creadora, pero con una diferencia sustantiva, el autor creador se observa con atención durante el movimiento de su intención a través del espacio conceptual e imaginario durante la acción indagadora. El investigador es un creador altamente reflexivo, un observador que nunca pierde detalle de lo que le sucede a su interior y de lo que acontece en su exterior.

La creatividad consiste, en pocas palabras, y según una imagen configurada en la indagación sobre el tema de la ciencia de la inteligencia artificial, en la exploración de un espacio conceptual a través de cierto árbol de búsqueda. El espacio conceptual es el ámbito de las imágenes y conceptos de la cultura que configuran la ecología de información y comunicación del sujeto explorador. El investigador social explora en diversos ámbitos del espacio conceptual posible, lo hace desde el mundo semicerrado del campo académico-científico, y se mueve hacia la percepción y construcción cognitiva de mundos sociales construidos por diversos actores y ámbitos colectivos. En este sentido, aunque su punto de partida es el

técnicos. La propuesta de relaciones entre estos componentes del proceso de investigación social que aquí se presente, también requiere un apunte teórico que vincule los juicios del orden del método y la técnica con imágenes del campo perceptivo sobre lo social.

Ahora, el punto es presentar los rasgos generales de una apuesta teórico-metodológica que permite ordenar los diversos paquetes técnicos. Ese orden se corresponde con una mirada que se descompone en tres dimensiones de organización de la información sobre lo social, lo que corresponde a la sociedad, a la cultura y a la ecología. Cada dimensión se ordena en niveles de complejidad, los cuales son útiles para ubicar objetos particulares de observación y configuración reflexiva.

La dimensión de la sociedad se refiere a lo más evidente de la composición, lo que aparece a la mirada del observador que inicia una indagación. Es el plano de la observación de lo que se ordena según el objeto cognitivo exploración y/o descripción. La dimensión de la cultura es lo que se encuentra más allá de lo evidente, la estructura que configura en poco a la diversidad y heterogeneidad de lo social. La dimensión de la ecología es lo más amplio de toda la observación reflexiva, lo que permite relacionar lo social-cultural con lo no social ni cultural, lo que marca los ámbitos generales de configuración de la escena humana, lo que marca los límites de lo humano y lo no humano. Las tres dimensiones se componen de niveles de abstracción y de complejidad en la configuración y la trayectoria.

La dimensión de la sociedad supone la extensión fenomenológica, una superficie de información que se corresponde en forma isomórfica con el plano de referencia en observación. Por tal motivo, supone una metodología que permita desplazarse a la observación en el tiempo y el espacio sociales con la mayor velocidad, como en el caso de la encuesta estadística, o con la mayor precisión y detalle, como en el caso de la etnografía. La imagen es la siguiente, un investigador inicia su observación desde un lugar y un tiempo, requiere moverse físicamente para observar y registrar por sí mismo el espacio-tiempo sociales, o necesita medios instrumentales que le permitan obtener observaciones semejantes con la misma eficiencia. O requiere conocer lo que sucede simultáneamente, lo cual le es imposible mediante sus propios ojos sin ningún medio técnico extra. Esta dimensión fenomenológica es la que se cubre con los paquetes técnicos de superficie.

El investigador no sólo requiere conocer lo que pasa, aunque ese es siempre su punto de partida. En cuanto obtiene un registro en superficie inicia la organización de información que le permitirá hacer conjeturas sobre por qué aparece *así* y no *asado* lo registrado. Esta es la dimensión en la cultura. El supuesto es que hay formas, en un nivel de patrón, que cubren una extensión de fenómenos que permite separar lo estructural de lo estrictamente fenoménico. El investigador busca entender lo que sustenta, lo que aparece, lo que organiza y compone lo diverso y único en ciertas formas generales. Esto implica paquetes técnicos que ordenan la información fenoménica de forma tal que se jerarquiza e inducen imágenes de orden.

La cibernética de segundo orden propone que en el primer nivel de complejidad, el que corresponde a la sociedad, tenemos relaciones de contigüidad, básicas, de presencia. En el segundo nivel configuramos relación de relaciones, algo así como un árbol que dé las ramas fundamentales y dé sostén estructural. Hay un tercer nivel, es lo contrario al primero, no es superficial, no es fenomenológico evidente, no se configura en extensión. El tercer nivel llega al tronco del árbol, identifica el centro organizador de lo diverso, y además lo relaciona con otros árboles y el contexto del bosque y la región, hasta llegar al posible cosmos así configurado. Esta es la dimensión de la ecología, la más ambiciosa y hermosa del plan, la que permite configurar redes de relación de relaciones. La complejidad es enorme, incluso mayor en número que la extensión uniforme y finita del primer nivel. Esta dimensión supone paquetes técnicos de una gran sofisticación, por las relaciones múltiples y simultáneas que proponen para mirar la diversidad organizada. Algo así como la teoría del caos.

En ciencias sociales tenemos el primero y segundo nivel con cierta claridad, el tercero es aún un nivel posible. Esto en tanto aparece un observador que se exterioriza a lo que acontece y observa a la distancia y concluye en hipótesis. En cuanto el observador se mueve en el interior de la trama social la situación es otra y aparece el tercer nivel en toda su magnitud. Es la trama misma de las observaciones interactuando la que configura la ecología social. Una ecología baja en observación reflexiva o en interacción perceptual se configura mejor desde afuera, el caso contrario impide su manipulación desde el exterior y sólo el movimiento social mismo produce el proceso de tercer nivel. Como podrá observarse aquí parecen dos puntos de vista no del todo reconciliables, uno apunta hacia la división social entre los que reflexionan y los que no, el otro promueve la reflexión particular y general de todos los integrantes de la ecología social.

Cultura de investigación. ¿Por qué y para qué investigar?

De la sociedad de información a la sociedad de comunicación.

La sociedad cerrada que se abre

La investigación depende del tipo de sociedad donde se realiza, de la cultura y de la ecología específicas. Decir algo así no tiene gran problema, el punto es desplegar la argumentación sobre esas dimensiones en relación a la actividad reflexiva de segundo nivel. No será lo mismo vivir la experiencia reflexiva en un medio que no promueve las preguntas, que incluso las inhibe o reprime, a vivir en una ecología donde la vivencia estética es el corazón de las relaciones humanas, donde abrir la percepción y enriquecer la conciencia son actividades centrales e indispensables. Lo esencial puede expresarse en el orden de prioridades, la coartada es que nun-

ca habrá tiempo para lo secundario, y que lo primario siempre será más importante. También puede expresarse en relación a determinismos fatales ante los cuales nada se puede ni debe intentar. El asunto es que todo queda en el ámbito de lo relativo cuando se mira la diversidad humana en lo social, lo cultural y lo ecológico, y aparece la imagen de lo posible como horizonte de lo realizable. No hay forma humana definitiva, todo puede cambiar, cualquier forma puede ser modificada a voluntad.

De acuerdo a la tipología social en cuatro tipos, comunidad de información, sociedad de información, sociedad de comunicación y comunidad de comunicación, la sociedad mexicana como parte de un movimiento global transita del segundo al tercero en forma compleja y particular. Es importante caracterizar en breve este proceso para ubicar a la investigación en su devenir. Así pues, un marco sociohistórico viene bien.

La sociedad de información es la forma cerrada de lo social. Se configura en el agrupamiento de conglomerados humanos en lugares acotados y ordenados, las ciudades, jerarquizados y controlados por lo más alto de la jerarquía. Sociedades con centro que gobierna la periferia, donde los pocos toman decisiones por los muchos. Forma social donde las mayorías delegan autoridad en una minoría. Forma social que cruza casi la totalidad de la historia humana en teocracias, monarquías, imperios, dictaduras, y casi todas las macroformas de organización colectiva, incluyendo las democracias occidentales actuales. La sociedad de información tiene una característica clave, unos manejan la información sobre el todo, son los únicos que pueden actuar sobre el todo, y por tanto necesitan medios que les aseguren el control, el principal es la información sobre una masa previsible y manipulable.

Las sociedades contemporáneas son de este tipo, durante el siglo xx se han desarrollado tecnologías y saberes diversos que promueven el que esa parte que decide tenga la mayor y mejor información para decidir y actuar según sus intereses. En tanto, las mayorías consumen y obedecen, cumplen el guión dictado por sus élites conductoras. Las formas de este tipo social no están agotadas, el futuro tiene un lugar aún para la sociedad de información.

Lo importante es que la investigación social emerge en el momento en que mayor necesidad hay de un control rápido y eficaz de una masa que siendo homogénea desde cierto punto de vista, en general, tiene diversos grados de heterogeneidad. El mercado es el gran nicho ecológico de este nuevo escenario. La investigación es indispensable para que unos sepan de todos, y esos todos puedan ser dirigidos en sus comportamientos por esos unos. Ese es el lugar general de la investigación en nuestro medio.

Frente a este escenario aparece el proyecto de la modernidad occidental, la sociedad de comunicación, la sociedad abierta, la compuesta por ciudadanos libres y participativos, la de individuos críticos y reflexivos. La paradoja es que entonces nace la ciencia social, pero no para promover la sociedad de comunicación

solamente, sino para ser instrumentada para la sociedad de información. La democracia es la cualidad central de este tipo social, para su movimiento requiere del diálogo de los iguales, del acuerdo entre los distintos pero tolerantes para un gobierno más horizontal. El asunto lleva poco más de dos siglos y aún batalla por abrirse paso en la forma dominante.

La sociedad de comunicación supone un manejo de información distinto, pone énfasis en la relación más que en el contenido. No importa tanto lo que sabemos de los otros y lo que podemos hacer sobre los demás con ese saber, lo importante es el acuerdo de diálogo y concertación con los demás sobre lo que a todos compete, poniendo en juego todo el saber posible para un mejor diálogo y una mayor decisión concertada y ejecutada. La investigación en este tipo social se separa del anterior sustantivamente, ya no es el control y la dominación lo importante, lo básico es el autocontrol y la comunicación. Una investigación social donde las formas del diálogo, del escuchar, sean el centro del trabajo reflexivo aún es escasa, casi inexistente, pero existen iniciativas emergentes en ese sentido.

La situación actual en México es consistente con los dos escenarios presentados, la sociedad de información es la dominante, la de comunicación es emergente. La investigación toma algún lugar en esta bifurcación, por una parte mantener la situación ecológica general, por otra parte promover la democracia y la reflexividad como formas elementales de convivencia.

De la cultura de información a la cultura de comunicación. Dominación frente a comunicación y democracia

Los grandes tipos sociales configuran los marcos ecológicos generales de la vida social y cotidiana. Para llegar a la observación de las situaciones particulares que la conforman se requiere de otra mediación perceptual que permita identificar las estructuras que ponen en forma las energías que llevan a constituir una u otra forma general ecológica. Esa mediación se configura en las imágenes de cultura de información y cultura de comunicación.

La cultura de información representa las tendencias en el comportamiento a la búsqueda, manejo y distinción de la información pertinente para la acción. Todo actor social requiere de cierta cultura de información para relacionarse con los demás, para vivir en sociedad. Esta cultura será la adecuada cuando el ajuste situacional entre la acción y el objeto de la acción se verifiquen con consistencia. Para comer se requiere cierta información, si la gente cumple con el objetivo es que tuvo la información pertinente a tiempo. En general nuestra cultura de información es la suficiente para sobrevivir y sobrellevar el curso de vida en el cual nos movemos. El punto es que dicho curso puede ser reconocido si aparece cierta informa-

ción, o si el sujeto la busca. Las situaciones cambian si la información sobre ellas se configura en forma distinta.

Por otra parte, la cultura de información adquiere su verdadera profundidad cuando se la observa en sectores sociales distintos y en forma comparativa. Ahí se perciben diferencias que están lejos de ser casuales o circunstanciales, más bien son consistentes con la ecología social general. Así, unos sectores manejan sólo cierto tipo de información sobre ciertos ámbitos de la organización social, y otros tienen un espectro más amplio. Esto permite que unos puedan percibir y decidir más allá de lo evidente, y otros no. Su cultura de información es distinta, su configuración y complejidad es diferente, el efecto de sus saberes sobre la vida social es desproporcionado.

La investigación social forma parte de las formas de la cultura de información. El saber sobre lo social se distribuye entre la población de manera disimétrica, a algunos les llega más información y a otros menos, por otra parte, unos buscan y necesitan más información que otros. La situación no es objeto de acción de la investigación tradicional, pero la investigación reflexiva de segundo orden intenta promover en los actores sociales una mayor cultura de información, una más intensa, superior y, sobre todo, una mejor distribución de los saberes estratégicos. El derecho a la información es parte de la investigación reflexiva, no es asunto pertinente en la investigación tradicional.

El asunto se pone más interesante con la cultura de comunicación. La sociedad de información tiene una muy baja cultura de comunicación, le interesa más el flujo de datos en ciertas direcciones, que constituir formas sociales de encuentro y diálogo. La razón es simple, una organización con trazos verticales no incluye a los horizontales mas que en un orden secundario y subordinado, como en el caso de las democracias actuales.

En la sociedad de comunicación se invierte el orden de subordinación prioritaria y primaria, la información depende de la comunicación. La información sigue teniendo una importancia clave, pero es estructuralmente más relevante lo que hacen con ella en interacción dialógica los actores. Es decir, el flujo de información no se mueve en una dirección predominante, se reconstituye en cada nodo interactivo. Esto supone una organización más compleja, así como un gasto de energía más alto en la interacción. De la misma manera, conlleva ciertos elementos de transformación en toda la vida social, en sus valores y objetos prioritarios.

La investigación social se modifica sustancialmente con el aumento en la cultura de comunicación, y en tránsito a una sociedad de comunicación. Indagar sigue siendo un oficio relevante, pero su ubicación en la organización no está sólo cerca del centro de control, se distribuye en el todo social que ahora tiende a multiplicar los centros de autocontrol, y a necesitar más los nodos interactivos para llegar a acuerdos o sólo para participar colectivamente de la reflexividad de cada

lugar que se auto-organiza. La configuración de la observación reflexiva se torna horizontal y con ello las reglas del juego cambian drásticamente.

En México la cultura de información es baja, sólo se tiene la necesaria para ir la llevando, el patrón de búsqueda o de exploración es casi inexistente como norma general. En contraste, ciertos sectores minoritarios aprenden con cierto desorden a manejar más y mejor información. Sin embargo, incluso estos sectores desperdician energía e información y toman decisiones sin información suficiente y adecuada. El panorama en la investigación es patético, sólo unos pocos lugares la emplean para potenciarse como tomadores de decisiones. La imagen general es la de una sociedad dependiente e inmovilizada.

En el caso de la cultura de información el asunto es peor. Hay un dictador en potencia en cada individuo, un pequeño tirano en cada lugar de autoridad. El diálogo es una ausencia constante. Y aún así, hay un movimiento que busca redes horizontales de relación y organización. Este movimiento inicia la cultura de investigación como una actividad colectiva, reflexiva y dialógica.

De la cultura de investigación

Este es un nuevo punto de partida, la pregunta por la situación de la cultura de investigación contemporánea, en particular en el medio mexicano. La escena es casi de risa. Por una parte, las vidas particulares de millones de personas navegan sin ningún vínculo con la formalidad académica y con una creciente dependencia de los resultados de la investigación fuera de nuestras fronteras; lo que se observa sobre todo en la tecnología de uso doméstico. Este es el grado cero de la situación, pero hay más.

Una segunda escena la componen los investigadores académicos del país, aislados casi por completo del movimiento social en cualquiera de sus manifestaciones, produciendo para una institucionalidad que les cambia sus esfuerzos por premios, algunos reconocimientos y una mínima garantía de carrera profesional dentro de unas reglas del juego que tienden a formalizarse para hacer más seria la cosa en un sentido oficial.

Una tercera escena puede ser la compuesta por los ejecutivos de investigación de las compañías especializadas, en el sector privado, y sus pares dentro de la carrera burocrático-política, en el llamado sector público. Ahí se hace lo que mandan los jefes, para obtener alguna ganancia económica o política, el investigador cumple con sólo acatar las órdenes que se le dan.

Existen otras escenas y más cuadros que cruzan actores, guionistas y directores dramáticos. Un comentario posible del resultado sería lo contenido en calificativos como dependencia, control, indiferencia. La cultura de investigación es una forma social subordinada a otros órdenes de configuración social, una forma

secundaria que no parece útil, que no se entiende, que no tiene valor pragmático evidente y que se manipula a cada paso, se ignora o simplemente permanece aislada.

El aspecto más interior de la cultura de investigación también es preocupante. Si, por una parte, la investigación conforma una pequeña aldea en el universo social, con relaciones de alta subordinación, discriminación y falta de comprensión, por otra, lo que sucede dentro de la aldea está sobredeterminado por las condiciones exteriores. Siempre con excepciones y con una apariencia de trabajo ordenado y sistemático, lo que sucede en la investigación académica, y en la otra, se parece mucho más a un proceso sólo administrativo y de movilidad burocrática que a un ejercicio creativo de exploración y sensibilidad estética.

El punto culminante de esta pequeña reseña toma tierra con la parte medular del proceso investigativo, la metodología y la tecnología. Para empezar, todo lo que sucede en el país se mueve bajo los parámetros y normas de control de calidad de una legislación exterior. Las miradas siempre se dirigen hacia afuera cuando de método y técnica se tiene que hablar. Por otra parte, el ensayo de estos medios es siempre sólo una mínima parte de lo que podría ser. La exploración metodológica y tecnológica es mínima, se tiende a operar con lo mismo, de la misma manera, y cuando más abierto aparenta ser el ejercicio forma parte de una moda, de algo que no se profundizará y será abandonado al poco tiempo.

En esta dimensión, la cultura de investigación muestra su rostro más descarnado, si en lo propio, su mundo interior, su espacio cultural específico, no hay apertura ni exploración, qué se puede esperar de la investigación en su convocatoria hacia el resto de la sociedad en el sentido de la creatividad y la promoción de una cultura de información y de comunicación más intensas. La investigación se separa de la política, la investigación social con mayor razón, pero la política interna, la elemental, la de lo cotidiano del investigador, esa es la menos reflexionada, la menos observada, la más ciega, la más inmovilizadora y corrupta.

Hay una dimensión más, de singular configuración, la que se refiere a la perspectiva general sobre el investigar y sus fundamentos y proyecciones. El siglo XIX fue muy rico en propuestas y polémicas en este sentido, siempre europeo, participando de la hegemonía mundial del sistema de vida social europeo. El don más reconocido de ese siglo, hacia finales del presente, es la acuñación positiva de la noción de ciencia. Esa forma del conocimiento que sustituiría al dogma religioso, según declara la leyenda, y que permitiría dar su lugar al proyecto de humanidad que promovió el espíritu romántico hasta el desastre de las dos guerras mundiales del siglo XX, según continúa diciendo la leyenda. Pero este sentido de ciencia, basado en la verdad, en la objetividad, en la pureza de la intención indagadora, se evaporó en el tránsito del antiguo régimen al mundo de la postguerra. Desde la filosofía, el pensamiento reflexivo de la física y las matemáticas, todo cambió, el observador constituye la calidad de lo observado, le confiere existencia, sólo la

crítica y la imaginación pueden hacer consistente el oficio de la ciencia. La verdad no existe, sólo existe la percepción y sus juegos cognitivos; la objetividad es sólo un momento de la reflexividad que es un movimiento constante de la observación, en el mundo social los objetos de observación no son mudos, también reflexionan y hablan, el dominio de la ciencia positiva se transforma en la interacción democrática de la comunicación. El tiempo pasó, el mundo cambió, la ciencia es otra.

El panorama actual es de una riqueza impresionante, la libertad vuelve a tener un lugar, pero no es una obligación social, es una condición del juego, de la crítica, de la creación, de la imaginación, del diálogo, del crecimiento de la mente ecológica en cada uno de sus componentes holonómicos. La cultura de investigación es de todos o es una fachada sospechosa.

Tecnología, metodología y epistemología

Los niveles de relación entre sujeto de conocimiento e interacción

Aquí la pregunta es casi pedagógica, ¿quién conoce y qué sucede entonces? También es un punto de partida cartesiano, el que se da cuenta que conoce se interroga sobre los sucedidos. Pero no es sólo eso, el proceso de conocimiento va más allá, desborda pronto el continente del individuo e interconecta redes, formas plasmáticas de percepción y entendimiento. No hay interior sin exterior, no hay exterior sin interior, las preguntas necesitan ser más audaces, el conocimiento no es sólo la relación de interiorización de un objeto por un sujeto, las categorías ahora son otras, la resonancia mórfica, la comunidad hofográfica, el nuevo pensamiento pretende ver más allá de lo evidente con imaginación y valentía.

La historia de la ciencia y la filosofía de la ciencia están en una revolución desde la emergencia de la perspectiva cognitiva y las formas críticas de la cibernética de segundo orden. Nada volverá a ser lo mismo, todo el pensamiento está en revisión, lo que se veía como lugar común hace una o dos décadas tiene nuevas facetas y percepciones posibles. El futuro vuelve a existir como una forma de la creación, el pasado también tienen el mismo *status*. El sentido puede ser configurado de diversas formas sin el ahogo de la incertidumbre o la confusión. No más verdad, sólo posibilidades, desde la teoría y el mundo práctico.

Un movimiento tal afecta la reflexión metodológica y tecnológica de la investigación en forma intensa y devastadora, aunque no al mismo tiempo y en impactos semejantes a todas las áreas. En el caso de la investigación social el fenómeno podría decirse que recién inicia, aunque está configurado desde el inicio del pensamiento sociológico, en aquellos tiempos decimonónicos del enfrentamiento de la sociología y el socialismo. Hay formas que tardan mucho tiempo en adquirir su

evidente profundidad, quizás un siglo sea aún poco tiempo para lo que viene, y lo paradójico es que el programa estaba bosquejado casi desde el principio.

En el principio fue el positivismo, con su certidumbre de que el mundo puede ser conocido, que es único como único es el conocimiento sobre él. Las ciencias naturales avanzaron con velocidad en la demostración pragmática de que la configuración de la naturaleza poseía regularidades que era pertinente explicitar. Ese fue el modelo, las ciencias sociales también se dedicaron a ensayar proposiciones de leyes generales sobre el comportamiento social. El esquema suponía una separación tajante entre el observador y lo observado, requería de esa distancia para afirmar que un sujeto estaba conociendo a un objeto, y que el único que tenía cualidades de reconfiguración reflexivas era el sujeto. En el caso de las ciencias físicas y naturales la imagen tuvo cierta utilidad, la superioridad de lo humano estaba fuera de toda duda, pero en el caso de lo social la propuesta estuvo envuelta en la polémica desde el principio.

Si bien a cierta mentalidad le venía justo el observar a lo social como un mundo-objeto que se diseccionaba con pulcritud y sistema, a otras mentalidades les pareció inconveniente, los actores sociales también significan, hablan, son reflexivos y no caben en la misma categoría del afán clasificador sobre rocas y pájaros. En este conflicto de visiones se configura lo que hoy conocemos como el pensamiento hermenéutico, el que interpreta, el que se mueve en significados no en datos, el abierto en forma permanente frente al cerrado positivo. Pero había más.

Este aparentemente irconciliable enfrentamiento tenía en común un elemento, al sujeto de conocimiento. Se partía de la experiencia directa y obvia de la vivencia individual del conocimiento. Tanto los positivistas como los humanistas hermenéuticos ubicaban la sede del conocimiento en la mente individual, la diferencia era que unos concedían mayor subjetividad a los actores sociales que los otros. De cualquier manera, el camino al reconocimiento de la reflexividad como cualidad humana y social generalizada y promovible estaba abierto.

Sería hasta la emergencia del pensamiento cibernético cuando parte de la discusión se aclara en un tercer escenario, o segundo, según se vea. Los actores sociales pueden ser observados como objetos-cosas, y de esa observación obtener cierto resultado que permita actuar sobre ellos considerándolos objetos-cosas manipulables, y tener éxito en la empresa. También pueden ser observados como subjetividades que toman decisiones y tienen capacidad de reflexionar sobre su situación, lo que los configura como seres más libres y autónomos de la voluntad de dominación. Esto convierte al pensamiento social en un juego de ajedrez donde se trata de prever el movimiento reflexivo del otro. Pero sucede que tanto el aparente observador solitario como los observados reflexivos, pueden observarse observando, y comunicarse en dicha situación reflexiva de segundo orden. Esto mueve a las ciencias sociales hacia otra parte, integrando las fases anteriores

como de menor complejidad, y proyectándose hacia una configuración de redes de interacción que se hacen y rehacen según los efectos en la acción de los contactos comunicativos reflexivos.

En este plano no hay dominación absoluta posible, lo que se abre es el horizonte de la autopoiesis social, una sociedad reflexiva que se auto-organiza sin centro ni control absoluto.

La teoría, la metodología y la acción técnica concreta en investigación. La búsqueda del sentido reflexivo organizado

La configuración sobre la investigación social que abre la cibernética de segundo orden compromete a un ejercicio más atento del proceso y de las operaciones concretas, al tiempo que exige un esfuerzo extra de visión más allá del proceso y de las acciones particulares. En este nuevo nicho, las relaciones entre teoría, metodología y operación técnica se reconfiguran con mayor claridad. Esto sucede porque el ejercicio reflexivo es más intenso, y se requiere de la auto-observación constante. Lo que pasa con todo esto es que cada proceso concreto es único, alimentado por las experiencias previas, y alimentador de las futuras. Es decir, estamos ante una forma de trabajo que se auto-organiza en la interacción entre la práctica y la previsión teórico-metodológica, además de la relación con los otros. La investigación toma así la forma de un proceso en movimiento constante, en cierto sentido hay un principio, pero después sólo aparece el movimiento y sus resultados parciales. El investigador es un viajero de la creación y la exploración, cambia a cada paso, aprende, se modifica a sí mismo por la auto-observación reflexiva, no repite una operación de la misma manera, es parte de la propia trayectoria auto-organizadora de lo social. La interacción de él con él mismo y de él con los demás, lo agrega al movimiento social, la investigación social pasa a formar parte del movimiento social mismo en su proceso de auto-organización.

Aún así, todo tiene un principio, un punto de partida. Para este caso es necesaria una imagen de un esquema de propaganda de los elementos básicos de la configuración dinámica mostrada. Estos elementos son las imágenes parciales de la tecnología, la metodología y la epistemología de la investigación social.

La tecnología configura los paquetes técnicos con los cuales el investigador opera en la relación de su mente con el mundo explorado y construido. Estas operaciones pueden ser de diversos tipos, su oferta es variada y aún espera un ejercicio de orden y catalogación analítico-reflexiva. Este catálogo sería sólo un punto de inicio, pues de inmediato aparece la labor concreta de investigación que combina y ajusta la oferta a las necesidades de indagación. Sea el caso que sea, por tecnología estamos entendiendo lo más concreto de la operación de investigación, el cómo de la acción indagadora.

En la tecnología se incluye todo tipo de operación, desde la más elemental hasta la más sofisticada. Para darnos una idea de lo que se trata imaginemos la descripción detallada de todo acontecimiento relacionado con la tarea de investigar. En algo así como una etnografía de la investigación aparecería una enorme cantidad de acciones concretas de distintos investigadores. El segundo paso sería ordenar en paquetes lo que en cada caso se operó. El procedimiento seguiría en esa línea hasta obtener formas estables construidas *a posteriori* sobre cómo trabajan la investigación un grupo o comunidad de investigadores. En el *a priori* posterior, un aprendiz de investigación tendría ante sí un menú de esas formas para escoger y configurar la propia y adecuada al proyecto que quisiera iniciar.

La tecnología de investigación está asociada entonces a la acción y, por lo tanto, es la investigación en sí, nada de especulaciones o deseos, nada de supuestos y propuestas, sólo la acción concreta. Como podrán imaginarse los lectores, en ocasiones hay una gran distancia entre lo que se supone pretende el investigador y lo que concretamente hace. De ahí que las decisiones sobre la acción sean las básicas, ahí se cocina todo, ahí se verifica, no hay coartada ni disculpa, sólo acción.

El asunto de la metodología es de un orden lógico diferente, responde a la pregunta de por qué se hacen las cosas así y no de otra manera. Se entiende que aquí se configura la guía de operaciones, muchas de las cuales pueden intercambiarse sin modificar sustantivamente el resultado, tal vez sólo un poco. El punto es que aquí es donde opera el estratega, el visionario de la investigación a corto plazo, por el tiempo que dura un proyecto particular.

La metodología decide el camino general donde las operaciones concretas representan los pasos particulares. Cuando los pasos se van agrupando en sendas parciales hay un punto donde se integran con la estrategia general; la decisión sobre los paquetes técnicos es la base de la acción estratégica, existen decisiones por encima de ella, y hacia el interior de los paquetes. Todo este juego de decisiones puede denominarse como metodológico pues está en referencia a la totalidad del camino por recorrer, aunque sólo se refiera a una parte de él.

La epistemología se ordena en una dimensión superior a la metodología, pero íntimamente relacionada con ella y con la tecnología. Ahí se define el para qué y para quién se hace lo que se hace. En este punto la investigación se mueve de su foco lógico interior y desplaza su atención hacia el mundo social donde actúa. Para muchos este punto es irrelevante, asunto de la política, fuera del oficio elemental de investigación. Como sea que fuere está ahí y afecta, asumiéndolo o no.

Es importante percibir en forma integrada las tres dimensiones, la operación concreta no está separada de la lógica estratégica general, ni de la interacción que promueve con el mundo social. Es más, la nueva investigación social supone la relación de los tres configuradores cuando define a la interacción social reflexiva como el escenario y el objeto de la acción de investigación.

Los paquetes técnicos. Operación concreta y mundo asociado y suturado perceptualmente por ella

Todo parte de la acción y se ve afectado por ella. La decisión sobre la selección de los paquetes técnicos pertinentes para una investigación determina la configuración de todo el proceso. Y siendo esta una operación del orden metodológico, supone un conocimiento de los caminos posibles a seguir ante un asunto por explorar. Este conocimiento implica una mínima experiencia con el transitar por distintos caminos, una vivencia de lo que sucede cuando se camina por una senda o por otra. Esta es una condición que muy pocas veces se cumple, la tendencia es más bien la de avanzar una y otra vez por el mismo camino ignorando opciones, e incluso renegando de ellas a pesar de la ignorancia. La tesis aquí es que lo básico es conocer rutas ya exploradas, ensayar su vivencia y actuar en todo momento en forma reflexiva y autoreflexiva. Un investigador también es un explorador de su oficio, además de un explorador de los mundos sociales.

En este marco el conocimiento de los paquetes técnicos es elemental. Iniciar una exploración vivencial en un vecindario sin conocer el paquete técnico etnográfico es un error mayor; argumentar eficiencia en tales situaciones una estupidez rotunda. La cultura de investigación específica del investigador se configura por el conocimiento, imágenes y conceptos del espacio-tiempo social que conforma su referente de trabajo, su mundo de trabajo, pero también, y sobre todo, por el árbol de búsqueda con el que explora ese espacio-tiempo. Ese árbol de búsqueda es su conocimiento y experiencia en los paquetes técnicos con los cuales puede moverse dentro de su espacio-tiempo social. Entre más pobre sea su árbol de búsqueda más pobre será su exploración o, cuando menos, será más pobre su espacio de posibilidades de exploración.

El orden de configuración de los paquetes técnicos depende del criterio con el cual se les asocia a los niveles de relación entre el sujeto de conocimiento y la interacción social. Este primer bosquejo se complementa con la combinación de niveles y paquetes que se establezca para una investigación concreta. Las rutas concretas de trabajo así configuradas son múltiples, las combinaciones y las modificaciones están en proporción de la experiencia, el conocimiento y la capacidad creativa. El proceso de investigación es un balance entre competencias administrativas, lógicas y creativas.

En principio, la decisión sobre el paquete técnico pertinente depende del objeto cognitivo correspondiente y el curso metodológico-epistemológico de todo el movimiento. Si interesa moverse sólo a nivel descriptivo y de superficie de comportamiento, el paquete estadístico y su combinatoria en el más sofisticado paquete encuesta puede ser adecuado, sobre todo si la reducción requerida de la superficie social se propone en números, cuantitativa. Si lo que se busca es superficie exhaustiva, una representación de situaciones en imágenes más icónicas, más pa-

recidas a lo que un observador en el lugar puede registrar, será conveniente el paquete etnográfico. Ambos, el estadístico-encuesta y el etnográfico parten del observador que da cuenta del acontecer social como dato.

La lógica previa se configura en la distribución, es decir, en la presencia-ausencia de cualidades que pueden o no ser contabilizadas. El asunto se complica cuando el nivel pasa de superficie a estructura, entonces se requieren paquetes que puedan manejar significados, construir nuevas categorías, organizar grupos de categorías en esquemas y constructos. Todas las técnicas de análisis de discurso son pertinentes aquí, aquellas que trabajan con el lenguaje, las que se aproximan al eje de la interpretación hermenéutica, de la significación de lo observado.

Los objetos cognitivos de la comprensión y la explicación se van organizando en paquetes de lógica positiva o hermenéutica según el caso, y la sofisticación de los paquetes aumenta. Cuando la reflexividad del actor social se toma en cuenta, estamos en una dimensión que se adentra en la interacción entre participantes del proceso de investigación. Así, los paquetes de historia de vida, historia oral, etnometodología y grupos de discusión, están de lleno en la operación reflexiva que afecta tanto al investigador que la induce como al actor social que la realiza. La relación con la vida social en vivo y espontánea es intensamente mayor.

Los umbrales de la interacción se traspasan por completo cuando el *status* del investigador pasa a segundo término y es la interacción misma entre los actores sociales y su reflexividad lo que importa. El grupo de discusión está cerca, la historia de vida también, pero el socioanálisis y la investigación-acción están dentro por completo. Ahí el objeto cognitivo es la interacción misma, la comunicación. Lo que sucede en ciertas psicoterapias se configura socialmente, los actores toman su proceso reflexivo y se mueven con autonomía en su propia exploración creativa y activa.

Como puede observarse, los paquetes técnicos se configuran con cierta afinidad a ciertos objetos cognitivos, los que definen el programa de investigación, pero también pueden ser combinados según la iniciativa y audacia del investigador. Lo primero es conocerlos. Así, las formas de análisis psicológico, historiográfico, heurístico y matemático, se agregan al grupo de posibilidades que aumentan posibilidades que aumentan posibilidades.

BIBLIOGRAFÍA

- ACEVEDO Ibañez, Alejandro y Alba Florencia A., López M., *El proceso de la entrevista. Conceptos y modelos*, Acevedo y asociados, México, 1981.
- ACEVES, Jorge, *Escribir la oralidad*, Conafe, México, 1992.
- ACOSTA Valverde, Miguel y Luz Paula Parra Rosales, *Guía para el análisis de contenido electoral de medios de comunicación*, Academia Mexicana de Derechos Humanos, México, 1994.
- ADORNO, Theodor W., *Teoría estética*, Editorial Taurus, Madrid, 1992.

Lectura 3

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 tampoco 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 descriptivas y causales válidas. Por lo tanto, se sitúa en una 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-

Tomada de: Gary King et al., "La ciencia en las ciencias sociales", en *El diseño de la investigación social. La inferencia científica en los estudios cualitativos*, Madrid, Alianza, 2007, pp. 13-43.

cho, a veces se consideran 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 cualitativistas como los cuantitativistas 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 detallados de materiales históricos; utiliza un método discursivo e intenta estudiar de forma global o exhaustiva un acontecimiento o unidad. Los cualitativistas, 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, institución, lugar, problema o ley. Al igual que en la investigación cuantitativa, el objeto de estudio puede ser importante en sí mismo: es un cambio crucial en una nación, unas elecciones, una decisión trascendental o una crisis mundial. ¿Por qué se derrumbó tan súbitamente el régimen germano 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 violenta-

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 (Skocpol, 1984, pp. 374-86).

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 diferencie fundamentalmente de la «cuantitativa», excepto en el estilo.

La mayoría de las investigaciones no corresponde 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 comparar, 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).

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 ambiguas: 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*

(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, así como a 115 líderes comunitarios en 1976 y a 118 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 interocular», «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 deberían 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.

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 describirlas 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 quizá haya datos —recogidos con los métodos científicos más útiles— que se queden obsoletos antes de organizarlos. Si una persona enloquecida corre hacia nosotros blandiendo un hacha, puede que entregarle 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 pretenden extraer del mundo inferencias descriptivas y causales. Los que no comparten 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 sofocarlo. 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 éstas 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-

gación, y para ello utilizamos la palabra «científica» como calificativo¹. Esta palabra tiene muchas connotaciones injustificables, inapropiadas o, para muchos investigadores cualitativos, completamente incendiarias. De ahí que nosotros le demos una definición explícita. Tiene que quedar claro que no consideramos más científica la investigación cuantitativa que la cualitativa. El estilo de una buena investigación —o sea, la que es científica— puede ser cuantitativo o cualitativo. Sin embargo, en lo tocante al diseño, la investigación tiene estas cuatro características:

1. El objetivo es la inferencia

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. Éstos pueden ser reunidos (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 entrañar 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.

2. Los procedimientos son públicos

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, sí 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-

servaciones y hacen inferencias sobre causas y consecuencias. Si el método y la lógica de las observaciones e inferencias de un investigador no se hacen explícitos, la comunidad académica no tiene forma de juzgar la validez de lo que se ha hecho. No podrá evaluar los principios de selección utilizados para registrar las observaciones, ni cómo se han procesado éstas, ni la lógica de las conclusiones. No podremos aprender de sus métodos ni reproducir sus resultados. Esta investigación no será un hecho *público*. Aunque sea una buena lectura, no será una aportación a las ciencias sociales.

Todos los métodos —sean o no explícitos— tienen sus limitaciones. La ventaja de lo explícito es que pueden captarse tales limitaciones y, si es posible, remediarlas. Además, los métodos se pueden enseñar y compartir. Este proceso permite que los resultados de la investigación se comparen con los de otros investigadores, que se reproduzcan los estudios de proyectos ajenos y que los académicos aprendan.

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 libro². 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.

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 contemple 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 confinado 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

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 compararlas 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 contrafáctico, 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 de la geología y la biología evolutiva, ciencias naturales con orientación histórica. Stephen J. Gould ha indicado que una de las formas de distinguir entre características evolutivas sistemáticas y acontecimientos estocásticos, aleatorios, puede ser imaginar cómo sería el mundo si se fijaran todas las condiciones, hasta un determinado punto, y se volviera a dar cuerda 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 Álvarez y sus colaboradores de Berkeley a finales de los años setenta (W. Álvarez y Asaro, 1990), postula que hubo una colisión cósmica: un meteorito chocó con la Tierra a unos 72.000 kilómetros 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 iridio (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 en las capas que se había pronosticado se ha considerado una prueba que confirma parcialmente 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 cráter del meteorito en algún lugar de la Tierra (y ya se han encontrado algunos candidatos)³.

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 Álvarez 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. Matizar 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 utilizarlos mejor hará posible que extraigamos más consecuencias de ellos. El hecho de estudiar acontecimientos muy complejos no hace irrelevante la

realización de un cuidadoso diseño de investigación. Si recogemos datos sobre la mayor cantidad posible de consecuencias observables de nuestra teoría, el estudio mejorará, independientemente de que estemos barajando muchos o pocos fenómenos —o incluso uno.

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 indicaba su proyecto, a menudo se encontrará con que el engarce entre las principales preguntas de la investigación, la teoría y los datos disponibles es imperfecto. En este momento, suele cundir 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 retiran los andamios 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.

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 ilimitados. Evidentemente, en cualquier investigación real siempre hay que hacer concesiones. Creemos que si el investigador entiende 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 qué hacer 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 “intuición 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-

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 lleven 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 probabilidades 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 comiencen a estudiarlo: las mujeres han solido llevar 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 peculiares 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.

Aunque no existan normas precisas para elegir un tema, si hay maneras de determinar —aparte de las preferencias personales— la posible trans-

endencia que tiene para la comunidad científica una empresa investigadora. 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 beneficiosos (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 otras el objetivo ni siquiera es hacer tal inferencia, sino únicamente observar de cerca ciertos acontecimientos o 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 principalmente social. La segunda condición nos sitúa en la bibliografía académica de las ciencias sociales, es decir, en las perplejidades 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 cuatrocientos 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-

rrollo 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 descriptivas 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 comprende el «estado de la cuestión» y reduce 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 carecía 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 continuación 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 alejado.

Si nos preocupamos demasiado por hacer una aportación a la bibliografía académica sin prestar cierta atención a asuntos importantes para el mundo real, correremos 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 comenzarán 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 decisión 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 estando más cerca de uno que de otro⁴.

Independientemente del punto de partida, el proceso de diseñar una investigación para dar respuesta a una pregunta debe cumplir nuestras dos condiciones, y es evidente que la dirección de este movimiento dependerá de dónde comience. Si lo que nos motiva es una perplejidad científica, tendremos que preguntarnos cómo hacer que el objeto de la investigación sea más relevante para asuntos importantes del mundo real: por ejemplo, cómo podría un experimento de laboratorio dilucidar mejor las elecciones estratégicas de los decisores del mundo real o las consecuencias que, para el comportamiento, puede tener esta teoría. Si partimos de un problema real, ten-

dreemos que preguntarnos cómo puede estudiarse con métodos científicos modernos para hacer una aportación al fondo de explicaciones de las ciencias sociales. Bien pudiera ser que apartarse demasiado de uno u otro criterio no fuera la mejor opción. Los que hacen experimentos de laboratorio pueden señalar que es prematuro buscar referencias externas y que se avanzará más si se perfecciona la teoría y el método en un medio controlado; quizá, en un estudio de larga duración, esto sea cierto. Por el contrario, quien esté preocupado por un problema del mundo real puede indicar que antes de explicar es necesario describir con precisión, y puede que también tenga razón, ya que la descripción exacta es un paso importante de los estudios explicativos.

En cualquier caso, si es posible, todo proyecto de investigación tiene que intentar cumplir nuestras dos condiciones: ocuparse de un asunto relevante del mundo real y estar concebido para realizar una aportación, directa o indirectamente, a una determinada área bibliográfica académica. En este libro, al ser nuestra principal preocupación que la investigación cualitativa sea más científica, nos ocuparemos más del investigador que parte del «mundo real», pero nuestro análisis es aplicable a ambas tendencias.

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 convertirlo en un proyecto de investigación concreto, que permita 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» (Lieberson, 1992, p. 4; véase también Woods y Walton, 1982).

El desarrollo de la teoría suele presentarse como el primer paso de la investigación. A menudo, en la práctica, es lo primero, pero no tiene por qué serlo. De hecho, no podemos elaborar una teoría sin conocer antes los trabajos realizados sobre el tema y sin haber reunido algunos datos, ya que en este estadio ni siquiera se sabe cuál es la pregunta de la investigación. No obstante, independientemente de la cantidad de datos que se hayan reunido ya, hay algunas formas generales de evaluar y aumentar la utilidad de una teoría. Ahora sólo vamos a esbozarlas, pero las analizaremos en detalle en capítulos posteriores.

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 principio⁵. Es necesario poder responder directamente a la pregunta: ¿qué datos nos convencerían de que estamos equivocados?⁶. 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 «conciación». Por desgracia, este término se ha utilizado tantas veces en conversaciones informales y en escritos académicos que el principio ha quedado eclipsado (véase Sober [1988] para un análisis completo). Jeffreys (1961, p. 47) dio la definición más clara de conciación: «Las teorías más simples tienen en principio más probabilidades»⁷. Por lo tanto, la conciación 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 conciación, pero los biólogos suelen creer que es absurda. En las ciencias sociales, los hay que la defienden con vehemencia para sus subáreas (por ejemplo, Zellner, 1984), pero nosotros creemos que sólo es útil en ciertas ocasiones. Dada la precisa definición de conciación como supuesto 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 conciación como un bien en sí mismo, porque, a menos que ya se-

pamos 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. Cualquier 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 disciplina⁸.

Todavía nos queda el problema de qué hacer cuando hayamos terminado tanto la recogida de datos como nuestro análisis y queramos 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 inadecuada restringiéndola 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

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 supongamos 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 periodos de prosperidad excepto allí donde el ejército es débil, los líderes políticos son represivos, la economía se basa en un reducido número de productos y el clima es cálido. Tal formulación es una forma caprichosa (y engañosa) de afirmar que «mi teoría 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 espuria. Si no nos esforzamos por recabar nuevos datos, no dispondremos 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 estarlo. En ciertos momentos, aceptar que estamos equivocados es lo mejor y, de hecho, algunos resultados negativos pueden ser bastante útiles para un área académica. ¿Quién no habría de preferir una sólida conclusión negativa antes que varias de carácter insustancial que se basen en teorías elaboradas para salir del paso?

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 más 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 diseños de investigación o proyectos de recogida de datos que podrían utilizarse

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 particularmente costosos. La recogida de datos preliminar puede llevarnos a alterar 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 cualitativo o cuantitativo. 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 Coombs, 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-

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 enunciar 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 cualquiera que se utilice, así como la información necesaria para que otros lo apliquen⁹.

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, podemos (a) recoger más observaciones sobre la misma variable dependiente o (b) registrar más variables dependientes. Podemos, por ejemplo, desagregar los datos en períodos de tiempo más cortos o áreas geográficas más pequeñas. También se puede recabar información sobre variables dependientes de un interés menos directo: si los resultados son como los que 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 creíbles. 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 improductividad.

La tercera directriz es: *maximizar 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*. Fiabilidad 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 entretanto 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 renombre— 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 imposibilita 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 sí 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.

Cómo 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 anterioridad.

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 prever los problemas antes de comenzar a reunir los datos. Sin embargo, aquí mencionamos brevemente estas 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 debe contemplar consecuencias relacionadas con las observaciones que esperamos 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 la 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 esté sujeta 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, qué nos permiten 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 persigue uno de los objetivos más importantes de las ciencias sociales: *explicar 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 princi-

pio parece un complicado efecto, el *control* que tendremos sobre un problema será muy elevado. Al revés, si podemos explicar muchos efectos a partir de una variable o de unas pocas, también dispondremos de un control considerable. En general, el control es escaso en las ciencias sociales y lo es aún más en determinadas áreas temáticas. Esto puede deberse a que los investigadores no saben aún cómo aumentarlo, a que la naturaleza no está organizada adecuadamente o a ambas razones. Las áreas que se estudian habitualmente de forma cualitativa suelen ser aquellas en las que hay poco control. Para explicar cualquier cosa se suele necesitar un conjunto de variables explicativas: utilizamos muchos elementos para explicar poco. En estos casos tendríamos que proponernos diseñar investigaciones con más control.

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 concisión, 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 *recomendamos fervientemente 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 «falacia 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 agregación¹². 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 hacernos 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 respuestas de los revolucionarios a entrevistas 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. Sólo 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 habría 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 limitados 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 habría que calcular de forma razonable el *grado de incertidumbre* que existe en cada una de ellas. Neustadt 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 pondrías en ello?». Esta pregunta tiene sentido siempre que también nos planteemos: «¿Con qué posibilidades?».

3.4 Pensar como un científico social: escepticismo e hipótesis contrapuestas

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

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.

Lectura 4

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 ésta, 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

Tomada de: Gary King *et al.*, "La inferencia descriptiva", en *El diseño de la investigación social. La inferencia científica en los estudios cualitativos*, Madrid, Alianza, 2007, pp. 45-85.

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 calidad. 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: descubrir conocimientos 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 pormenores 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

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 impensables —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 imaginémonos el análisis que hizo Robert Caro (1983) de las elecciones tejanas de 1948 al Senado sin Lyndon Johnson y Coke Stevenson¹. Acontecimientos concretos como la Revolución francesa o las primarias demócratas de 1948 para el Senado en Tejas 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).

1.1 «Interpretación» e inferencia

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 pormenores históricos 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 quitarle al comportamiento humano, antes de comenzar a examinarlo, precisamente las propiedades que nos atraen».

Los investigadores que hacen hincapié en la «interpretación» pretenden arrojar luz sobre los aspectos deliberados del comportamiento humano utilizando el concepto de *Verstehen* («empatía: comprender 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 relacionándolas 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. Duneier (1993), por ejemplo, estudió el comportamiento colectivo de los hombres de clase trabajadora negros y blancos en un bar autoservicio 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 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 menoscabar 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 Fenno del Congreso estadounidense (Fenno, 1978) a través de lo que él 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». Cualquier definición de ciencia que no deje lugar a las ideas relativas a la elaboración de hipótesis será tan frívola 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 conspirativo para un amigo. Los dos movimientos son, como tales movimientos, idénticos. Si se realizara una observación «fenomenística» 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 guiñar, 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 conspirativo, hacer este movimiento *es* guiñar.

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 «empaparse y hurgar» 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 qué 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 pormenorizada 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 definimos 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 interpretamos 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 voluntaria, fracasarían sistemáticamente: no podríamos generalizar, y lo sabríamos².

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 interpretemos el significado de una acción. En política exterior, los que toman 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 guiño 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 que:

cualquier 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 distinguirlas. 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 presentimiento, 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-

cuencias observables diferentes. Por otra parte, si las hay, los métodos que describimos en este libro proporcionan formas de distinguirlas.

En la práctica, los etnógrafos (y todos los buenos científicos sociales) *sí* que buscan comportamientos observables con el fin de establecer diferencias entre las teorías. Puede que se sumerjan en la cultura, pero todos se basan en diversas formas de *observación*, y cualquier nueva «apreciación» del contexto cultural procede directamente de ellas o de otros medios comparables. No siempre resulta fácil identificar observaciones relevantes. Por el contrario, encontrar las apropiadas quizá sea la parte más difícil de un proyecto de investigación, especialmente (y necesariamente) en aquellas áreas de estudio tradicionalmente dominadas por la investigación cualitativa.

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 «únicos», 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 intentaran 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 únicos. 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-

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 —cualquiera que sea su densidad— y ninguna explicación —independientemente del número de factores explicativos que contemple— pueden llegar a captar, siquiera levemente, toda la «florecente 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 podría 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 aislarnos del tiempo sin perder el vínculo con él, y la insistencia en uno de estos objetivos puede variar de una investiga-

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. Confía 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 critica 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 cualitativistas 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 sondan 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

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 pasarse 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 sistemática³.

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-

tandariza las necesidades de los datos de los estudios de caso [...] formulando 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 traído: «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 mágnam 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 inalterable durante el proceso. La teoría y los datos interaccionan. 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 utilicemos los datos para compro-

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 partió⁴.

En tercer lugar, insistir en que se recojan datos que sean consecuencias observables de una hipótesis deja mucho más claro el territorio común que comparten 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, ¿qué 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: cuándo se tomó, 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 elementos 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»

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 escaños 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.

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. Ésta 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; preguntar si cada aspecto del modelo es verdadero es como plantearse si la modelo que posó para la *Gioconda* de Leonardo da Vinci tenía realmente una sonrisa tan seductora. Aunque la tuviera, no esperamos 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 esperamos 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.

¿Podemos evaluar un modelo sin conocer qué características del objeto queremos evaluar? Está claro que no. Por ejemplo, podríamos pensar que un modelo que se ocupara 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 (ahorrándose 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 ocupe 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 juguetes aeronáuticos 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 cuantitativa 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

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 cómo 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 ayudarnos, 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 entrevistas no estructuradas podría utilizar los valores «participativo», «cooperador» o «intransigente», 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. Éstas, 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 cuál 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 número 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 mucho más general. En realidad, nuestra definición de «observación» coincide exactamente con lo que Harry Eckstein (1975, p. 85) denomina «caso». 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 $n = 120.000.000$.

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 disponemos. 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 historiadores 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 víveres que encontraba a su paso; pero puede que sea irrelevante conocer el color del pelo de Napoleón o si prefería los huevos fritos a los cocidos. Los buenos textos históricos suelen incluir, aunque no se limiten a ello, un sucinto resumen verbal de gran cantidad 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 objeti-

vo es exponer características apropiadas de éstos mediante un formato útil⁵. Por ejemplo, la *media muestral*, o promedio, es un estadístico:

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

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 representan, 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 si 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 examinaríamos valores medios (de presupuestos, cometidos, número de miembros, etc.), pero para comprender el abanico de actividades tendríamos 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

ningún tipo de resumen ⁶. La síntesis que utilicemos 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 suficiente.

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 cómo 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 escaños 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 aleatorio y lo impredecible, al igual que el conjunto de la vida social y toda investigación científica ⁷. Supongamos que por un súbito descuido (o como deferencia hacia las ciencias sociales) el Parlamento Británico hubiera decidido celebrar elecciones cada semana en 1979 y supongamos también (de forma contrafáctica) que éstas 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, sí podrían influir en los resultados semanales. De este modo, numerosos acontecimientos transitorios podrían producir resultados electorales ligeramente diferentes. Después de todo, nuestra observación de cualquiera 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-

comunidades 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 sí se reduciría 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 referentes al apoyo tradicional a cada partido. En hipotéticas 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

de observación en una pequeña comunidad es sólo una de las muchas semanas 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 prevalecer sobre el sistemático 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 atiende 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 Y 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 Y_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 Y .

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-Bireh⁸. La variable aleatoria Y_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 observación⁹.

Uno de los objetivos de la inferencia es conocer las *características sistemáticas* de las variables aleatorias Y_1, \dots, Y_n (hay que señalar lo contradictorio de esta terminología estándar: aunque, en general, queremos distinguir el componente sistemático del no sistemático en 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 se espera que obtenga el Partido Laborista en la circunscripción 5 (el promedio de votos Y_5 para este partido en un gran número de elecciones hipotéticas en esa circunscripción). Al ser ésta 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 aleatorios¹⁰.

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

$$E(Y_5) = \mu_5$$

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 característica sistemática para una variable aleatoria Y_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 *media*:

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

Un estimador de μ podría ser la media de los niveles de conflictividad observados en todas las comunidades que se estudian, \bar{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-

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

Todavía hay otra característica de estas variables aleatorias que podría ser de interés: la variación del nivel de conflictividad dentro de una comunidad aun cuando las características sistemáticas no cambien, es decir, en qué medida las observaciones de diferentes semanas (diversas aplicaciones hipotéticas 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):

$$V(Y_i) = \sigma_i^2 \quad (2.3)$$

en donde σ^2 (la letra griega sigma) indica lo que se obtiene al aplicar el operador de varianza a la variable aleatoria Y_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 registraran una elevada varianza y, por tanto, fueran muy impredecibles, 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 vista fundamentales sobre la variación aleatoria¹¹, que constituyen los dos extremos de una gradación. Aunque hay un número significativo de investigadores que se siente 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 variación aleatoria existe tanto en la naturaleza como en los mundos social y político, y nunca podrá eliminarse. Aunque midiéramos todas las variables sin cometer errores, registráramos todos los datos (en vez de sólo una muestra) e incluyéramos toda variable explicativa concebible, nuestros análisis no podrían hacer nunca predicciones perfectas. Un investigador puede dividir el mundo en componentes aparentemente sistemáticos 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 variación no sistemática que existe en diversas partes del mundo empírico.

Punto de vista 2: un mundo determinista

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

de qué variables explicativas se encuentren disponibles y de cuáles 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 variación aleatoria. Al carecerse de cualquier consecuencia observable 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 sistemático 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 encontrar otro conjunto de datos (o incluso un solo caso más) con el que comprobar la persistencia de un efecto o pauta, será muy difícil hacer una apreciación 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 acontecimiento es estocástico o sistemático es imposible o irrelevante. Una versión más realista de esta perspectiva acepta que el punto de vista 1 puede considerarse de forma correcta o incorrecta que una pauta es aleatoria o sistemática, pero deja cierta libertad para decidir cuál será objeto de examen en cada estudio concreto y cuál 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. Decidir si un acontecimiento o proceso no explicado es un hecho realmente aleatorio o si procede simplemente de variables explicativas aún 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 utilidad para las ciencias sociales cuando también es explícitamente inferencial. Para conceptualizar las variables aleatorias de las que proceden las observaciones e intentar hacer una estimación de sus características sistemáticas —en vez de limitarse a resumir los pormenores históricos— no es preciso recoger datos a gran escala. De hecho, uno de los rasgos que definen 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 sistemática¹³. 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 perfectos) 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 mandatarios, 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 laborista 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).

En pocas palabras, resumir los pormenores históricos es un importante paso intermedio en el proceso de utilización de los datos, pero también te-

nemos que hacer inferencias descriptivas que distingan 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 tienen 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 hipotéticas 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 (si, por ejemplo, los conservadores fueran más reacios a votar en sábado por razones religiosas). Las estimaciones reproducidas también podrían haberse basado en informes de contabilizadores de votos corruptos que favorecieran 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 favoreciera 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 si el sistema electoral tuviera un método de recuento que favoreciera más a 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 esté 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 que hay que ser consciente: las personas que nos proporcionan los datos brutos 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 cuidados experimentos, sólo lo aplicamos una vez. En este caso, la falta de sesgo es importante, pero quisiéramos 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 constituirán un apretado conglomerado cercano al valor equivocado). Como describimos 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 estuviéramos 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 quisiéramos hacer una estimación de μ en la ecuación (2.2) y que decidiéramos 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, $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$, 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, \bar{Y} producirá resultados electorales que se acerquen a μ y en otras ocasiones éstos se alejarán de ella. La cuestión es si \bar{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:

$$\begin{aligned} E(\bar{Y}) &= E\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) && (2.4) \\ &= \frac{1}{n} \sum_{i=1}^n E(Y_i) \\ &= \frac{1}{n} n\mu \\ &= \mu \end{aligned}$$

Así, \bar{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).

tra 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

un n grande ($n = 25$) y que representa el nivel medio de conflictividad 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 $V(Y_{\text{típica}}) = \sigma^2$. Esto quiere decir que habríamos elegido una circunscripción que consideramos «típica», la cual, no obstante, se vería afectada por variables aleatorias. La varianza del gran estimador n es $V(\bar{Y}) = \sigma^2/25$, 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 variabilidad (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 variabilidad se acerca a cero y la estimación es igual al parámetro que estamos intentando calcular¹⁴.

Sin embargo, es frecuente que las condiciones no sean las mismas. Supongamos, por ejemplo, que cualquier medida individual de un fenómeno que estemos 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 sólo 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-

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 conflictividad observado para hacer una estimación del nivel medio, μ , que se registra en Cisjordania. Ésta 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. Mediante 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 observaran de cerca un país, podríamos corregir las estimaciones de su gobierno para acercarlas 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 la Comisión y del Consejo de Ministros. Podríamos recabar información en 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

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 enfrentaríamos de nuevo a la necesidad de alcanzar un equilibrio entre sesgo y eficiencia, y quizá tuviera 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 indicador de 100 casos y usar después una versión corregida de este último como 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 contemporáneas.

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 trabaja de cerca con los materiales y comprende 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:

$$\begin{aligned} V(\bar{Y}) &= V\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n V(Y_i) \end{aligned}$$

Además, si presuponemos 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:

$$\begin{aligned} V(\bar{Y}) &= \frac{1}{n^2} \sum_{i=1}^n V(Y_i) & (2.5) \\ &= \frac{1}{n^2} \sum_{i=1}^n \sigma^2 \\ &= \frac{1}{n^2} n \sigma^2 \\ &= \sigma^2/n \end{aligned}$$

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.

go y eficiencia. 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, supongamos que quisiéramos hacer una estimación de μ y que el estudio de n grande produce un estimador d :

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

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 circunscripciones 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 \bar{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:

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

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

$$\begin{aligned}
E(c) &= E\left(\frac{Y_1 + Y_2}{2}\right) \\
&= \frac{E(Y_1) + E(Y_2)}{2} \\
&= \frac{\mu + \mu}{2} \\
&= \mu
\end{aligned}$$

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:

$$\begin{aligned}
V(d) &= V\left(\frac{1}{n} \sum_{i=1}^n Y_i - 0,01\right) \\
&= V\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) - V(0,01) \\
&= \sigma^2/n \\
&= \sigma^2/650
\end{aligned}$$

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:

$$\begin{aligned}
V(c) &= V\left(\frac{Y_1 + Y_2}{2}\right) \\
&= \frac{1}{4} [V(Y_1) + V(Y_2)] \\
&= \frac{1}{4} 2\sigma^2 \\
&= \sigma^2/2
\end{aligned}$$

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:

$$\begin{aligned} \text{EMC}(g) &= V(g) + E(g - \gamma)^2 && (2.6) \\ &= \text{varianza} + \text{sesgo al cuadrado} \end{aligned}$$

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 preguntamos 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 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 congresistas 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:

$$\begin{aligned} \text{EMC}(d) &= \frac{\sigma^2}{650} + (0,01)^2 & (2.7) \\ &= \frac{\sigma^2}{650} + 0,0001 \end{aligned}$$

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 sesgaría ligeramente 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 apliquemos 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.

Lectura 5

3. Causalidad e inferencia causal

Hemos analizado dos estadios de la investigación social: resumen de los pormenores históricos y extracción de inferencias descriptivas mediante la división del mundo en componentes sistemáticos y no sistemáticos (en los apartados 5 y 6, respectivamente, del capítulo anterior). Muchos estudiosos de los fenómenos sociales y políticos se detendrían en este punto, evitando hacer afirmaciones causales y pidiendo a unos hechos seleccionados y bien ordenados que «hablen por sí mismos».

Al igual que los historiadores, los científicos sociales necesitan resumir los pormenores históricos y hacer inferencias descriptivas. Sin embargo, cuando las ciencias sociales pretenden alcanzar ciertos objetivos, si no hay inferencia causal, el análisis es insuficiente. Esto significa que, al igual que una inferencia causal es imposible sin una buena inferencia descriptiva, ésta, por sí misma, suele resultar insatisfactoria e incompleta. No obstante, con esto no queremos decir que todos los científicos sociales tengan que intentar, en cualquiera de sus trabajos, dar explicaciones causales a los fenómenos que estudian. A veces la inferencia causal es demasiado difícil y en otras muchas situaciones la descriptiva constituye el fin último de la empresa investigadora.

Es evidente que siempre tenemos que dejar claro si el objetivo de nuestro proyecto de investigación es describir o explicar. Muchos científicos sociales no se sienten cómodos con la inferencia causal. Les preocupa tanto la advertencia de que «correlación no es causalidad» que no plantean hipótesis o inferencias de tipo causal y señalan que su trabajo «estudia asociaciones y no causas». Otros caen con facilidad en afirmaciones aparentemente

Tomada de: Gary King *et al.*, "Causalidad e inferencia causal", en *El diseño de la investigación social. La inferencia científica en los estudios cualitativos*, Madrid, Alianza, 2007, pp. 87-124.

causales, calificando de «explicaciones» hipótesis o especulaciones no evaluadas, que se basan en diseños de investigación imprecisos¹. Creemos que ambas posturas esquivan el problema de la inferencia causal.

Evitar el lenguaje causal cuando la causalidad es el auténtico objeto de la investigación o bien hace irrelevante el estudio o bien le permite no respetar las reglas de la inferencia científica. Nuestra incertidumbre acerca de las inferencias causales nunca desaparecerá, pero esto no debe significar que evitemos extraerlas. Por el contrario, habría que hacerlas en cualquier lugar que parezcan apropiadas, pero aportando al lector, al mismo tiempo, la mejor y más sincera estimación de su incertidumbre. Hay que ser valiente al hacer inferencias causales, siempre que detallemos cuidadosamente su incertidumbre. Además, es importante que las hipótesis causales se disciplinen, acercándose lo más posible a las reglas de la inferencia causal. Entre los capítulos 4 y 6 casi toda nuestra atención se dedica a explicar las circunstancias en las que dicha inferencia es apropiada y a facilitar que las investigaciones de los cualitativistas tengan más posibilidades de aportar pruebas fiables de sus hipótesis causales.

En el apartado 1 damos una definición rigurosa de causalidad, que sirve igualmente para la investigación cualitativa que para la cuantitativa; después, en el apartado 2, clarificamos algunos conceptos de causalidad alternativos que figuran en la bibliografía sobre el tema y demostramos que no entran en conflicto con nuestra definición más fundamental. En el apartado 3 analizamos los presupuestos precisos sobre el mundo y las hipótesis que se necesitan para hacer inferencias causales fiables. Posteriormente, en el apartado 4, nos centramos en cómo aplicar a la inferencia causal los criterios que desarrollamos para juzgar la descriptiva. En el apartado 5 terminamos el capítulo con consejos más generales sobre cómo elaborar explicaciones, teorías e hipótesis de tipo causal.

1. Definición de causalidad

En este apartado, definimos la causalidad como un concepto *teórico* que es independiente de los datos utilizados para conocerlo. A continuación, entramos a considerar la *inferencia* causal a partir de nuestros datos (para un análisis de los problemas específicos de este tipo de inferencia, véanse los capítulos 4 al 6). En el apartado 1.1 planteamos con todo detalle nuestra definición de causalidad, junto a un simple ejemplo cuantitativo, mientras que en el 1.2 retomamos esta definición con un caso cuantitativo más elaborado.

1.1 Definición y un ejemplo cuantitativo

Nuestra definición teórica de causalidad se aplica de la forma más simple y clara a una sola unidad². Una unidad, tal como se definió en el apartado 4 del capítulo 2, es uno de los muchos elementos que hay que observar en un

estudio, como una persona, un país, un año o una organización política. Para mayor precisión y claridad, hemos elegido un ejemplo corriente de la investigación cuantitativa: el efecto causal que tiene sobre la proporción de votos que recibe un candidato demócrata a la Cámara de Representantes estadounidense el hecho de que ya esté en el poder (utilizar sólo un candidato demócrata simplifica el ejemplo). La variable dependiente será la proporción de votos que obtienen los demócratas en el sistema bipartidista de la Cámara. Por tanto, la variable explicativa clave es dicotómica: el demócrata está en el poder o no lo está (para lograr una mayor sencillez, en este apartado sólo tenemos en cuenta las circunscripciones en las que el candidato republicano fue derrotado en los últimos comicios).

El lenguaje causal puede ser confuso, y nuestra opción aquí no es muy original. A veces la «variable dependiente» se denomina «variable de resultados» y con frecuencia las «variables explicativas» son «variables independientes». Dividimos las explicativas entre «variable causal principal» (también denominada «causa» o «variable de tratamiento») y «variables de control». Para terminar, la variable causal principal siempre adopta dos o más valores, que suelen conocerse como «grupo de tratamiento» y «grupo de control».

Pensemos ahora únicamente en la cuarta circunscripción electoral de Nueva York para las elecciones al Congreso e imaginemos que en 1998 hubieran tenido lugar unos comicios en los que se hubieran enfrentado un congresista demócrata y un aspirante republicano (no en el poder). Supongamos que el candidato demócrata hubiera recibido en estas elecciones una fracción del voto expresada como y_4^1 (el subíndice 4 indica la cuarta circunscripción de Nueva York y el superíndice 1 el hecho de que el demócrata sea ya congresista). Por tanto, y_4^1 es un valor de la variable dependiente. Para *definir* el efecto causal (una cantidad *teórica*), imaginemos que damos marcha atrás en el tiempo hasta el comienzo de la campaña electoral y que todo sigue igual, a excepción de que el congresista demócrata decide no presentarse a la reelección y que su partido nombra a otro candidato (probablemente al ganador de las elecciones primarias). Expresamos la fracción de voto que el candidato demócrata (no congresista) recibirá mediante y_4^0 (N se refiere a un candidato demócrata que no está en el poder)³.

Esta condición *contrafáctica* es la esencia de la definición de causalidad, y la diferencia, en esta situación hipotética, entre voto real (y_4^1) y voto probable (y_4^0) es el efecto causal, concepto que definimos con mayor precisión más adelante. Hay que tener mucho cuidado al definir contrafácticos porque, aunque se oponen evidentemente a los hechos, tienen que ser razonables: debería ser posible que el acontecimiento contrafáctico hubiera ocurrido en unas determinadas circunstancias. Al definir la condición contrafáctica adecuada resulta esencial precisar qué se mantiene constante cuando alteramos el valor de la variable de tratamiento. En nuestro ejemplo, la variable causal principal (o de tratamiento) es estar en el poder, y

cambia desde «ser congresista» a «no ser congresista». Durante esta transformación hipotética se mantienen todos los elementos constantes (la fuerza relativa de los demócratas o los republicanos en esa circunscripción en elecciones anteriores, el carácter del proceso de nominación del candidato, las características de la circunscripción, el contexto económico y político del momento, etc.) hasta que el Partido Demócrata decide su candidato. *No* sometemos a ningún control las cualidades de los candidatos, como en qué medida se reconoce su nombre, su visibilidad o si se sabe cómo funciona el Congreso, ni tampoco cualquier otra consecuencia que tenga la nominación por parte del partido. Esto se debe a que, en parte, estos elementos son *consecuencia* de nuestra variable de tratamiento: estar en el poder, lo cual quiere decir que, entre las ventajas de esta situación, se incluye el reconocimiento del nombre, la visibilidad y todo lo demás. Si realmente mantuviéramos constantes estos elementos, estaríamos controlando algunos de los efectos más importantes que tiene estar en el poder y, por tanto, haciendo caso omiso de ellos. En consecuencia, malinterpretaríamos su influencia global en el conjunto del voto. De hecho, controlar un número suficiente de las consecuencias que tiene estar en el poder podría llevarnos a la incorrecta conclusión de que dicha situación no tiene efecto alguno⁴.

De manera más formal, el efecto causal que tiene estar en el poder en la cuarta circunscripción de Nueva York —la proporción de votos recibida por el candidato demócrata que puede atribuirse al hecho de que ya sea congresista— sería la diferencia entre estas dos fracciones del voto: $(y_4^D - y_4^R)$. Por razones que pronto quedarán claras, esta diferencia la denominamos *efecto causal realizado*, y, para expresarla con una notación más general, utilizamos la unidad i en vez de sólo la circunscripción 4⁵:

$$(\text{Efecto causal realizado en la unidad } i) = y_i^D - y_i^R \quad (3.1)$$

Es obvio que este efecto se define únicamente en teoría, ya que, en cualquier elección real, sólo podríamos observar y_4^D o y_4^R , o ninguno de los dos, pero nunca ambos a la vez. De este modo, esta sencilla definición de causalidad demuestra que nunca podemos conocer con certeza un efecto causal. Para Holland (1986), éste es *el problema fundamental de la inferencia causal*, y en verdad lo es porque, independientemente de lo perfecto que sea el diseño de la investigación, de la cantidad de datos que recojamos, de lo perspicaces que sean los observadores, de lo diligentes que sean los ayudantes y del grado de control experimental que tengamos, nunca conoceremos a ciencia cierta la inferencia causal. De hecho, gran parte de las cuestiones empíricas de los diseños de investigación que se analizan en este libro tienen que ver con este problema fundamental y casi todas las indicaciones que damos son intentos parciales de evitarlo.

Nuestra definición básica de causalidad se aparta de la de Holland, ya que en el apartado 6 del capítulo 2 hemos señalado que las ciencias sociales

siempre necesitan dividir el mundo en componentes sistemáticos y no sistemáticos, mientras que la definición de Holland no establece claramente esta distinción⁶. Para ver la importancia de esta división, pensemos en qué ocurriría si pudiéramos repetir la campaña electoral de 1998 en la cuarta circunscripción de Nueva York, con el congresista demócrata y el aspirante republicano al cargo. La suma total de votos tendría un resultado ligeramente diferente, a causa de ciertas características no sistemáticas de las campañas electorales: aspectos políticos que no se mantienen de unas campañas a otras, incluso si éstas parten de situaciones idénticas. Entre estas características no sistemáticas se pueden incluir las meteduras de pata verbales, un discurso o pronunciamiento sobre un problema que concita una atención sorprendente, una actuación inesperadamente desafortunada en un debate, mal tiempo durante un mitin del candidato o el día de las elecciones o también los resultados de alguna investigación periodística. Por lo tanto, podemos imaginar una variable que representara los valores del voto demócrata en diversas reproducciones hipotéticas de las mismas elecciones.

Como hemos señalado anteriormente (véase el apartado 6 del capítulo 2), esta variable se denomina «variable aleatoria», ya que tiene características no sistemáticas: le afectan las variables explicativas que no contempla nuestro análisis teórico o contiene una variabilidad fundamentalmente inexplicable⁷. Definimos la variable aleatoria que representa la proporción de votos recibida por el candidato demócrata que ya es congresista como Y_4^I (obsérvese la Y mayúscula), y la proporción de los que habría recibido un demócrata no congresista en reproducciones hipotéticas, como Y_4^N .

Ahora definimos el *efecto causal aleatorio* en la circunscripción 4 como la diferencia entre estas dos variables aleatorias. Como queremos mantener cierto carácter general, de nuevo utilizamos una notación que se refiere a la unidad i en vez de a la circunscripción 4:

$$(\text{Efecto causal aleatorio en la unidad } i) = (Y_i^I - Y_i^N) \quad (3.2)$$

(Al igual que en la definición de variable aleatoria, un efecto causal del mismo tipo es el que cambia en cada reproducción hipotética del mismo experimento, pero también representa muchas interesantes características sistemáticas de las elecciones.) Si pudiéramos observar a la vez dos proporciones de voto diferentes en la circunscripción 4 —la primera referida a unas elecciones a las que se presentara un congresista demócrata y la segunda sobre otras sin él—, podríamos ver directamente el efecto causal realizado de la ecuación (3.1). Es evidente que, a causa del problema fundamental de inferencia causal, no podemos observar el efecto causal realizado; en consecuencia, en la ecuación 3.1, dicho efecto es una única aplicación *no observada* del efecto causal aleatorio de la ecuación 3.2. Dicho de otro modo, si en la circunscripción 4 se hacen diferentes reproducciones hipotéticas de las mismas elecciones, unas con un candidato congresista de-

mócrata y otras tantas con uno no congresista, el efecto causal realizado (no observado) se convertirá en un efecto causal aleatorio.

Describir la causalidad como una de las características sistemáticas de las variables aleatorias puede parecer excesivamente complicado, pero tiene dos ventajas. La primera es que hace que nuestra definición de causalidad sea directamente análoga a las características sistemáticas de un fenómeno (como son la media o la varianza) que sirven de objeto a la inferencia descriptiva: las medias y las varianzas también son características sistemáticas de las variables aleatorias (como en el apartado 2 del capítulo anterior). La segunda es que posibilita la división de un problema inferencial causal en sus componentes sistemáticos y no sistemáticos. Aunque muchas de las características sistemáticas de una variable aleatoria puedan ser de interés, *la más relevante para nuestro simple ejemplo es el efecto causal medio* de la unidad i . Para explicar qué significa esto retomamos nuestro ejemplo de las elecciones en Nueva York.

Recuérdese que la variable aleatoria representa la fracción del voto recibida por el candidato demócrata (ya sea congresista o no) en un número elevado de reproducciones hipotéticas de los mismos comicios. Definimos de la siguiente manera el valor esperado de esta variable aleatoria —la media de fracciones del voto en dichas reproducciones— para el no congresista:

$$E(Y_i^N) = \mu_i^N$$

y para el congresista como:

$$E(Y_i^C) = \mu_i^C$$

En consecuencia, el efecto causal medio de estar en el poder para la unidad i es una característica sistemática del efecto causal aleatorio, y se define como la diferencia entre estos dos valores esperados (utilizamos de nuevo, para generalizar, la unidad i en vez de la circunscripción 4):

$$\begin{aligned} \text{Efecto causal} \\ \text{medio en la unidad } i &\equiv \beta \end{aligned} \tag{3.3}$$

$$= E(\text{Efecto causal aleatorio en la unidad } i)$$

$$= E(Y_i^C - Y_i^N)$$

$$= E(Y_i^C) - E(Y_i^N)$$

$$= \mu_i^C - \mu_i^N$$

de modo que en la primera línea de esta ecuación, β (beta) expresa este efecto causal medio. En la segunda indicamos que en la unidad i este indicador es sólo la media (valor esperado) del efecto causal aleatorio, mientras que en las líneas tercera y cuarta se muestra cómo calcular la media. La última línea es otra manera de expresar la diferencia entre las medias de los dos conjuntos de elecciones hipotéticas (la media de la diferencia de dos variables aleatorias es igual a la diferencia de sus medias). Para resumirlo con palabras: *el efecto causal es la diferencia que existe entre el componente sistemático de las observaciones que se hacen cuando la variable explicativa tiene un valor y el componente sistemático de observaciones comparables cuando la variable explicativa tiene otro valor.*

La última línea de la ecuación 3.3 es similar a la ecuación 3.1, y, al ser así, el problema fundamental de la inferencia causal existe todavía en esta fórmula. De hecho, expresado de esta manera, el problema es aún más grande, porque, aunque pudiéramos esquivar el problema fundamental en un efecto causal realizado, todavía seguiríamos teniendo todos los habituales en la inferencia, incluido el de separar los componentes sistemáticos y no sistemáticos del efecto causal aleatorio. A partir de ahora, utilizaremos la expresión de Holland, «problema fundamental de la inferencia causal», para hacer referencia a la dificultad que él identificó y también a los problemas inferenciales habituales, que hemos añadido a la formulación de este autor. En el recuadro de la página 108 mostramos una notación más general de los efectos causales, que resultará útil en el resto del libro.

Hay otras muchas características sistemáticas de estos efectos causales aleatorios que podrían ser de interés en diversas circunstancias. Por ejemplo, quizá queramos saber la varianza de los posibles efectos causales (realizados) que tiene el hecho de estar en el poder sobre el voto demócrata en la unidad i , al igual que quisimos conocer la del propio voto mediante la ecuación 2.3 del apartado 6 del capítulo anterior. Para calcular este indicador del efecto causal, aplicamos la operación de la varianza:

$$(\text{varianza del efecto causal en la unidad } i) = V(Y_i^T - Y_i^C)$$

en la que evitamos introducir un nuevo símbolo que represente el resultado del cálculo de la varianza, $V(Y_i^T - Y_i^C)$. Con seguridad, los nuevos congresistas querrán saber la variación del efecto causal que tiene el hecho de estar en el poder para poder evaluar en qué medida su experiencia se parecerá a la de los congresistas anteriores y hasta qué punto pueden fiarse de la estimación del efecto causal medio que tuvo el hecho de estar en el poder en elecciones anteriores. Resulta especialmente importante comprender que esta varianza del efecto causal es una parte fundamental del mundo y que no representa una incertidumbre causada por la estimación.

1.2 Un ejemplo cualitativo

En el apartado 1 hemos desarrollado nuestra precisa definición de causalidad y, al ser delicados y bastante complejos algunos de los conceptos que se barajaban, ilustramos la explicación con un ejemplo simple y corriente de la investigación cuantitativa, que nos ayudó a transmitir los conceptos que queríamos recalcar sin tener que ocuparnos de los detalles contextuales y de la sensibilidad cultural que caracterizan la investigación cualitativa de calidad. En este apartado retomamos nuestra definición de causalidad, pero esta vez utilizamos un ejemplo cualitativo.

Los politólogos aprenderían mucho si pudieran volver a poner en marcha la historia manteniendo todos sus elementos constantes, a excepción de la variable explicativa controlada por el investigador. Por ejemplo, una de las cuestiones más importantes a las que se enfrentan los que participan de la política y el gobierno es la relacionada con las consecuencias de una determinada ley o normativa. El Congreso de los Estados Unidos aprueba una ley impositiva para que tenga ciertas consecuencias: que conduzca a unas inversiones concretas, incremente los ingresos en cierta cantidad y cambie las pautas de consumo. ¿Acaso tiene tal efecto? Podemos observar lo que ocurre después de que se aprueba el impuesto, para comprobar si surgen las consecuencias deseadas, pero, aunque sea así, nunca estaremos seguros de que *provengan* de la ley. El cambio en la política de inversiones podría haberse producido en cualquier caso. Si fuera posible volver a poner en marcha la historia, con y sin la nueva normativa, tendríamos mucho más control a la hora de hacer una estimación del efecto causal de la ley. Evidentemente, no podemos hacerlo, pero la lógica nos ayudará a diseñar una investigación que dé una respuesta aproximada a nuestra pregunta.

Pensemos ahora en un ejemplo general de la política comparada. Después del derrumbe del sistema soviético numerosos gobiernos de las repúblicas ex soviéticas de Europa del Este instauraron nuevas formas de gobierno. Ahora se hallan inmersos —y son conscientes de ello— en un importante experimento político: están aprobando nuevas constituciones con la intención de que éstas creen sistemas democráticos estables. Uno de los dilemas constitucionales que se les presentan es elegir entre sistemas de gobierno parlamentarios o presidenciales. Determinar cuál de ellos tiene más posibilidades de conducir a una democracia estable ha suscitado un considerable debate entre los académicos del área (Linz, 1993; Horowitz, 1993; Lijphart, 1993). La polémica es compleja, principalmente porque hay muchos tipos de sistemas parlamentarios y presidenciales y también por la variedad de las demás disposiciones constitucionales (como la naturaleza del régimen electoral) que podrían acompañar a este dilema e interactuar con él. Nuestro objetivo no es analizar en profundidad estas opciones, sino dar una versión muy simplificada del dilema que plantean, con el fin de definir un efecto causal dentro de este ejemplo cualitativo. De este modo, su-

brayamos la diferencia que hay entre las características sistemáticas y no sistemáticas del efecto causal.

El debate sobre sistemas presidenciales y parlamentarios tiene que ver con diversas características de ambos. Nos centraremos en dos de ellas: en qué medida representa cada ordenamiento los diversos intereses de la ciudadanía y si favorece la existencia de un liderazgo fuerte y con capacidad de decisión. Se suele decir que los sistemas parlamentarios permiten una mejor representación de todos los grupos e intereses sociales en el gobierno, ya que hay muchos escaños que llenar y los pueden ocupar representantes elegidos por diversos sectores. Por el contrario, el carácter drástico (todo o nada) de los sistemas presidenciales conlleva que algunos grupos se sientan apartados del gobierno, le sean desafectos y generen una mayor inestabilidad. Por otra parte, es posible que los regímenes parlamentarios —especialmente los que representan de manera adecuada a todo el arco de grupos e intereses sociales— se bloqueen y que no cuenten con un gobierno decidido. También estas características pueden conducir a la desafección y la inestabilidad⁸.

El objetivo principal de este apartado es formular una definición precisa de efecto causal. Para ello, imaginemos que podemos instaurar un sistema parlamentario y que, periódicamente, durante unos diez años, calibramos su grado de estabilidad democrática (quizá la propia supervivencia o desaparición del sistema, así como los intentos de golpe de Estado que sufre u otros indicadores de inestabilidad); a la vez, en el mismo país, fundamos un sistema presidencial y también medimos su estabilidad en el mismo período y con los mismos elementos. El *efecto causal realizado* sería la diferencia entre los grados de estabilidad observados en el sistema presidencial y en el parlamentario. La imposibilidad de calibrar este efecto causal directamente es otro ejemplo del problema fundamental que plantea la inferencia causal.

Como parte de esta definición, también necesitamos distinguir los efectos sistemáticos y no sistemáticos de la forma de gobierno. Para ello, imaginemos que llevamos a cabo este experimento hipotético muchas veces. Definimos el *efecto causal medio* como la media de los efectos causales realizados en las reproducciones de estos experimentos. Utilizar así la media hace que se anulen las características no sistemáticas del problema y permite que nuestro indicador incluya sólo las sistemáticas. Entre éstas se encuentran la indecisión en un sistema parlamentario o la desafección de las minorías en uno presidencial, mientras que entre las no sistemáticas se puede incluir una repentina enfermedad del presidente que precipita el gobierno en el caos. Este último acontecimiento no sería un rasgo persistente en un sistema presidencial; aparecería en una de las repeticiones del experimento, pero no en otras⁹.

Otro rasgo notable de este ejemplo es la varianza del efecto causal. A cualquier país que se esté planteando elegir entre estos sistemas políticos le interesará conocer el efecto causal medio que tiene uno u otro sobre la esta-

bilidad democrática; sin embargo, para cada país sólo hay una oportunidad: sólo una reproducción de este experimento. En tal situación, puede que a los líderes políticos les interese algo más que el efecto causal medio. Quizá quieran comprender cuáles podrían ser los efectos causales máximo y mínimo o, al menos, la *varianza* de los efectos causales. Por ejemplo, quizá el sistema presidencial reduzca la estabilidad democrática media, pero la variabilidad de este efecto es enorme: a veces aumenta mucho la estabilidad y otras la disminuye de forma considerable. Esta varianza se convierte en un riesgo para el sistema político. En tales circunstancias, puede que los ciudadanos y los líderes políticos prefieran optar por un ordenamiento que genere una estabilidad media sólo un poco menor, pero con una varianza también inferior en su efecto causal, de manera que se minimicen las posibilidades de llegar a un resultado desastroso.

2. Clarificación de las definiciones alternativas de causalidad

En el apartado 1 hemos definido la causalidad en función de un efecto causal: el efecto causal medio es la diferencia que hay entre los componentes sistemáticos de una variable dependiente cuando la variable causal tiene dos valores diferentes. En este apartado, utilizamos nuestra definición de causalidad para clarificar diversas propuestas alternativas e ideas aparentemente complejas. Demostramos que las importantes afirmaciones ---contempladas en este mismo capítulo--- que han hecho otros autores acerca de los «mecanismos causales» (apartado 2.1), la causalidad «múltiple» (apartado 2.2) y la causalidad «simétrica» frente a la «asimétrica» (apartado 2.3) no entran en conflicto con nuestra más básica definición de causalidad.

2.1 «Mecanismos causales»

Algunos investigadores señalan que la idea principal de la causalidad es el conjunto de «mecanismos causales» que se dice que existen entre causa y efecto (véase Little, 1991, p. 15). De forma intuitiva, este punto de vista tiene sentido: cualquier explicación coherente que se dé a la causalidad tiene que precisar cómo se producen los efectos. Por ejemplo, supongamos que a un investigador le interesen las consecuencias que puede tener un nuevo acuerdo fiscal entre los Estados Unidos y Japón sobre el déficit de la balanza de pagos del primer país respecto al segundo. Según nuestra definición de causalidad, aquí el efecto causal se basaría en comparar la reducción del déficit de la balanza de pagos esperado, cuando se aplique el nuevo acuerdo fiscal, y la misma situación (en igual momento y con los mismos países) sin que haya entrado en vigor el tratado. A su vez, el meca-

nismo causal que opera en este proceso conllevaría —hasta que se produjera el efecto final sobre la balanza de pagos estadounidense respecto a Japón— la firma y ratificación del acuerdo, la información que dieran los periódicos, las reuniones de los actores afectados en las compañías multinacionales y las acciones compensatorias que realizaran para reducir la carga fiscal de sus empresas (como el cambio en las reglas de fijación de precios de transferencia o el traslado de las fábricas de un país a otro), así como las acciones de otras compañías y trabajadores que quisieran aprovecharse de los movimientos de capital y de fuerza productiva entre los países, etc.

Desde el punto de vista de los procesos con los que opera la causalidad, hacer hincapié en los mecanismos causales tiene sentido intuitivamente: toda explicación coherente de la causalidad tiene que explicar de qué manera se producen sus efectos. Identificar mecanismos causales es una forma habitual de hacer análisis empíricos a la que se ha denominado con expresiones ligeramente diferentes: «rastreo de procesos» (que analizaremos en el apartado 3.3 del capítulo 6), «análisis histórico» y «estudios de caso detallados». Muchos de los pormenores de los estudios de caso bien hechos conllevan la identificación de mecanismos causales.

Sin embargo, para precisar cuáles son tales mecanismos es necesario utilizar la inferencia causal, empleando los métodos que se analizan posteriormente. Esto quiere decir que el investigador, para demostrar el carácter causal de cada posible vínculo en un mecanismo como éste, tendrá que definir el efecto causal subyacente y luego hacer una estimación del mismo. Si se quiere representar un mecanismo causal con coherencia interna es necesario utilizar, para cada uno de los eslabones de la cadena de acontecimientos, la definición de causalidad fundamental que hemos dado en el apartado 1 de este capítulo.

De ahí que nuestra definición de causalidad sea, lógicamente, previa a la identificación de los mecanismos causales. Además, en las ciencias sociales siempre existen infinidad de pasos causales entre cualquier par de eslabones de esa cadena de mecanismos. Si postulamos que una variable explicativa produce otra dependiente, para utilizar un enfoque a base de «mecanismos causales» tendremos que listar los vínculos causales que hay entre ambas variables. En esta definición también necesitaremos determinar una serie de enlaces causales, definir la causalidad en cada par de variables consecutivas de la secuencia y precisar qué unión existe entre cada dos variables, así como la conexión entre parejas diferentes. Este enfoque nos conducirá rápidamente a un regreso al infinito y en ningún momento ofrece por sí solo una definición precisa de causalidad, relativa a una causa y un efecto.

En el ejemplo referido al efecto que tiene un sistema presidencial o parlamentario en la estabilidad democrática (apartado 1.2) los mecanismos causales de la hipótesis incluyen una mayor desafección de las minorías en

el régimen presidencial y una menor decisión del gobierno en el parlamentario. Estos efectos intervinientes —que genera el ordenamiento constitucional y que, a su vez, influyen en la estabilidad política— pueden observarse directamente. Podríamos hacer un seguimiento de las actitudes o comportamientos de las minorías para ver de qué modo difieren en dos situaciones experimentales o estudiar lo decididos que son los gobiernos en cada uno de los sistemas. Sin embargo, aunque los efectos causales de los sistemas presidenciales y de los parlamentarios operaran de forma diferente, nuestra definición de efecto causal seguiría siendo válida. Se puede definir dicho concepto sin comprender todos los mecanismos causales que conlleva, pero no determinar éstos sin definir el efecto causal.

Para nosotros, identificar los mecanismos que hacen que una causa tenga un efecto suele servir de apoyo a una teoría y es un procedimiento operativo muy útil. A veces, determinar mecanismos causales puede darnos más control sobre una teoría, al convertir las observaciones de otros niveles de análisis en consecuencias de dicha teoría. El concepto también puede generar nuevas hipótesis causales que investigar. Sin embargo, no debemos confundir la definición de causalidad con el procedimiento operativo de identificar mecanismos causales, que no es definitorio pero sí suele ser útil.

2.2 «Causalidad múltiple»

En un trabajo reciente, Charles Ragin (1987, pp. 34-52) aboga por una metodología que utilice muchas variables explicativas y pocas observaciones, con el fin de que se pueda tener en cuenta lo que él denomina «causalidad múltiple». Esto quiere decir que «El fenómeno que se investiga tiene determinantes alternativos; lo que Mill (1843) llamaba el problema de la “pluralidad de causas”». Para la teoría general de sistemas (George, 1982, p. 11), este problema se denomina «equifinalidad». En situaciones con causas múltiples, estos autores señalan que un mismo resultado pueden producirlo combinaciones de diferentes variables independientes¹⁰.

Para Ragin, cuando diversas variables explicativas pueden explicar el mismo resultado en una variable dependiente, hay métodos estadísticos que rechazan equivocadamente la hipótesis de que tales variables tengan un carácter causal. Ragin tiene razón al afirmar que algunos modelos estadísticos (o relevantes diseños de investigación cualitativa) quizá no alerten al investigador de la existencia de una «causalidad múltiple», pero sí los hay adecuados (algunos de los que Ragin analiza) para manejar con facilidad este tipo de situaciones.

Además, los rasgos fundamentales de la «causalidad múltiple» pueden compatibilizarse con nuestra definición de causalidad y son los mismos en la investigación cuantitativa y en la cualitativa. Esta idea no contiene rasgos o requisitos teóricos nuevos. Consideremos, por ejemplo, la hipótesis de

que el nivel de renta de una persona depende *tanto* de que culmine con éxito estudios superiores *como* del hecho de que sus padres también lo hicieran. Resulta insuficiente cumplir una condición y no la otra. En este caso, necesitamos comparar diversas categorías de nuestra variable causal: los encuestados que **cumplen** ambas condiciones, los dos grupos en los que sólo se da una u otra y aquel que no cuenta con ninguna. De este modo, el concepto de «causalidad múltiple» exige más de nuestros datos, ya que ahora nuestras variables causales se dividen en cuatro categorías, aunque esto no haga necesario un cambio en nuestra definición de causalidad. Para ésta, tendríamos que medir la renta esperada en la misma persona, al mismo tiempo y en cada una de las cuatro situaciones.

Sin embargo, ¿qué ocurre si diferentes explicaciones causales producen los mismos valores en la variable dependiente? Supongamos, por ejemplo, que en una población de trabajadores fabriles nuestra variable causal (dicotómica) fuera el hecho de tener o no tener una licenciatura universitaria. En esta situación sería bastante razonable que licenciados y no licenciados tuvieran la misma renta (nuestra variable dependiente). Una de las razones puede ser que esta variable explicativa (los estudios universitarios) no tiene un efecto causal en la renta de los trabajadores fabriles, quizá porque la educación superior no ayuda a un mejor rendimiento. Por el contrario, hay diversas razones que pueden explicar que ambos grupos tengan el mismo nivel de renta. Quizá el de los licenciados tenga que ver con su título y el de los que no fueron a la universidad con sus cuatro años de antigüedad en el puesto. En esta situación, ¿no podríamos llegar a la conclusión de que la «educación universitaria» no tiene un efecto causal en los niveles de renta de los que van a convertirse en trabajadores fabriles?

Por fortuna, nuestra definición de causalidad requiere una mayor precisión en la condición contrafáctica. En el presente ejemplo los valores de la variable causal principal que hay que cambiar son (1) tener un título universitario frente a (2) carecer de él pero llevar cuatro años en un puesto. La variable dependiente es la renta anual inicial. En consecuencia, nuestro efecto causal se define de la siguiente manera: registramos la renta de una persona que termina su carrera y va a trabajar a una fábrica. Después, retrocedemos cuatro años y ponemos al mismo individuo a trabajar en la misma fábrica, en vez de mandarle a la universidad, y al cabo de cuatro años medimos «de nuevo» sus ingresos. La diferencia esperada entre estos dos niveles de renta en el mismo individuo es nuestra definición de efecto causal medio. En la situación actual, hemos imaginado que este efecto causal es igual a cero, pero esto no significa que «los estudios universitarios no tengan un efecto sobre la renta», sólo que la diferencia media entre los grupos de tratamiento (1) y (2) es cero. De hecho, no hay una única definición lógica del «efecto causal de los estudios universitarios», ya que no se puede definir un efecto causal sin que haya, al menos, dos condiciones. No tienen por qué ser las que hemos mencionado, pero es preciso identificarlas claramente.

Otro posible par de condiciones causales es el que surge de comparar a una persona con título universitario con otra que carece de él pero que tiene la misma antigüedad en el puesto que la primera. En cierto sentido, esto es poco realista, ya que el no universitario habrá hecho algo en los cuatro años en los que no ha estudiado, pero quizá estemos dispuestos a imaginar que en ese tiempo ha tenido otro empleo de poca importancia. Dicho de otro modo, este contrafáctico alternativo se basa en comparar el efecto de tener estudios universitarios con el de no tenerlos, siempre que la antigüedad en la empresa sea la misma. Si este elemento no se mantiene constante en las dos condiciones causales, las estimaciones de cualquier diseño de investigación darán los resultados de nuestro primer contrafáctico, y no los de esta versión revisada. Si el objetivo es el segundo pero no introducimos controles, nuestro análisis empírico será defectuoso por el «sesgo de la variable omitida» (que abordamos en el apartado 2 del capítulo 5).

De este modo, las cuestiones que se agrupan bajo la etiqueta de «causalidad múltiple» no confunden nuestra definición de causalidad, aunque exigen más de los análisis posteriores. El hecho de que algunas variables dependientes, y quizá todas las interesantes de este tipo que hay en las ciencias sociales, se vean influidas por muchos factores causales no hace problemática nuestra definición de causalidad. La clave para entender estas situaciones tan habituales es definir la condición contrafáctica mediante una precisa concepción de cada efecto causal. En el capítulo 5 demostramos que, para estimar el único efecto causal que tiene interés, no es necesario identificar «todos» los de una variable dependiente (aunque esto fuera posible). El investigador puede centrarse en el único efecto que le interesa, llegar a conclusiones consistentes y pasar después a otros efectos que puedan ser relevantes (véanse los apartados 2 y 3 del capítulo 5)¹¹.

2.3 Causalidad «simétrica» y «asimétrica»

Stanley Lieberman (1985, pp. 63-64) distingue entre lo que clasifica como formas de causalidad «simétricas» y «asimétricas» y se centra en las diferencias que se producen en los efectos causales cuando se aumenta una variable explicativa o cuando se reduce. Según sus propias palabras:

Al examinar, por ejemplo, la influencia causal de X_1 [una variable explicativa] sobre Y [una dependiente], también hay que considerar si los cambios producidos desde cualquier dirección en un valor dado de X_1 tienen las mismas consecuencias para Y [...] Si la relación causal entre X_1 [la variable explicativa] e Y [la dependiente] es simétrica o realmente reversible, el efecto que tendrá sobre Y un aumento de X_1 desaparecerá si X_1 vuelve a su nivel anterior (siempre que el resto de las condiciones sea constante).

Como ejemplo de la afirmación de Lieberson, imaginemos que en las elecciones al Congreso estadounidense de 1998 no se hubiera presentado ningún congresista en la cuarta circunscripción de Nueva York y que el candidato demócrata recibiera el 55 por cien de los votos. Lieberson definiría el efecto causal de estar en el poder como el aumento de votos que se produciría si el demócrata ganador en 1998 se presentara, siendo ya congresista, a las siguientes elecciones del año 2000. Este efecto sería «simétrico» si la ausencia de un candidato congresista en las elecciones posteriores (las del año 2002) hiciera que los votos volvieran a ser del 55 por cien. El efecto podría ser «asimétrico» si, por ejemplo, el congresista demócrata reuniera dinero y mejorara la organización de la campaña de su partido: en consecuencia, si ningún congresista se presentara a las elecciones del 2002, el candidato demócrata podría recibir más del 55 por cien de los votos.

El argumento de Lieberson es inteligente y de gran importancia. Sin embargo, para nosotros no constituye una definición de *causalidad* sino que sólo es aplicable a algunas *inferencias* causales: al proceso de aprendizaje sobre un efecto causal a partir de las observaciones existentes. En el apartado 1 de este capítulo hemos definido la causalidad en una sola unidad. En el ejemplo actual, un efecto causal puede definirse desde el punto de vista teórico en función de los acontecimientos hipotéticos que han tenido lugar únicamente en las elecciones de 1998 en la cuarta circunscripción de Nueva York. Nuestra definición se basa en la diferencia que se da en el componente sistemático del voto en esa circunscripción cuando se presenta un congresista y cuando no se presenta, si las elecciones, la fecha y la circunscripción son las mismas.

Por el contrario, el ejemplo de Lieberson no incluye cantidades hipotéticas y, por tanto, no puede ser una definición causal. Sólo se ocupa de lo que ocurriría en dos elecciones reales si la variable explicativa pasara de ser no congresista a congresista y de lo que sucedería si se diera el proceso contrario —de congresista a no congresista— en otros dos comicios. Cualquier análisis empírico de este ejemplo plantearía numerosos problemas inferenciales. Analizamos muchos de los problemas relativos a la inferencia causal entre los capítulos 4 y 6. En el presente ejemplo podríamos preguntarnos si el efecto estimado sólo nos parecía mayor por no haber explicado el hecho de que se hubieran registrado recientemente una gran cantidad de ciudadanos en la cuarta circunscripción, o si la oleada de apoyo para el congresista demócrata en las elecciones se nos antojaba más pequeña de lo que era porque, necesariamente, descartábamos aquellas circunscripciones en las que los demócratas habían perdido en las primeras elecciones.

De este modo, es importante considerar los conceptos de causalidad «simétrica» y «asimétrica» de Lieberson en el contexto de la inferencia causal. Sin embargo, no deben confundirse con la definición teórica de causalidad que ya hemos dado en el apartado 1 de este capítulo.

3. Supuestos necesarios para la estimación de efectos causales

¿Cómo evitar el problema fundamental de la inferencia causal y también el de separar los componentes sistemáticos de los no sistemáticos? Los capítulos que van del 4 al 6 se dedicarán íntegramente a dar una respuesta completa a esta pregunta, pero aquí proporcionamos una visión general de los dos supuestos que pueden evitarnos este problema fundamental: la *homogeneidad de las unidades* (que analizaremos a continuación en el apartado 3.1) y la *independencia condicional* (apartado 3.2). Al igual que cualquier intento de esquivar el problema fundamental de la inferencia causal, tales supuestos siempre conllevan otros que no son comprobables. Cada investigador tiene la responsabilidad de mostrar a los lectores con toda claridad las consecuencias esenciales de este punto débil de su diseño de investigación. Las inferencias causales no deben parecer mágicas y los supuestos pueden y deben justificarse mediante cualquier información colateral o investigación anterior que se pueda aportar, siempre que se reconozca explícitamente.

3.1 Homogeneidad de las unidades

Si no podemos volver a poner en marcha la historia en el mismo momento y lugar, dando diferentes valores a nuestra variable explicativa en cada ocasión — tal como requeriría una auténtica solución para el problema fundamental de la inferencia causal—, si se puede partir de un segundo supuesto: poner otra vez en marcha nuestro experimento en dos unidades diferentes que sean «homogéneas». *Dos unidades son homogéneas si los valores esperados de las variables dependientes de cada una de ellas son los mismos cuando nuestra variable explicativa adopta un valor determinado* (o sea, $\mu_1^X = \mu_2^X$ y $\mu_1^0 = \mu_2^0$). Por ejemplo, si observamos que $X = 1$ (un congresista) en la circunscripción 1 y $X = 0$ (no congresista) en la 2, la premisa de homogeneidad de las unidades supone que podemos utilizar las proporciones observadas del voto en dos circunscripciones separadas para extraer una inferencia sobre el efecto causal β , que presuponemos es el mismo en ambas. En un conjunto de datos con n observaciones, la homogeneidad de las unidades consiste en presuponer que todas las que tengan igual valor en sus variables explicativas tendrán también el mismo valor esperado en la dependiente. Es evidente que esto es sólo un supuesto y que puede ser erróneo: quizá ambas circunscripciones sean diferentes por alguna razón desconocida que sesgue nuestra inferencia causal. De hecho, cualquier par de circunscripciones reales *diferirá* en ciertos sentidos, y para aplicar este supuesto ambas tendrán que ser iguales en cuanto a la media de muchas reproducciones hipotéticas de la campaña electoral. Por ejemplo, la pluviosi-

dad (que puede reducir la asistencia a las urnas en algunas áreas) no variará, como promedio, de una circunscripción a otra, a menos que haya diferencias climáticas sistemáticas entre las dos.

En la cita siguiente, Holland (1986, p. 947) da un ejemplo claro del supuesto de homogeneidad de las unidades (definido a partir de su idea de efecto causal realizado y no en función del efecto causal medio). Al no existir apenas aleatoriedad en el experimento del siguiente ejemplo, su definición y la nuestra se parecen (de hecho, como planteamos en el apartado 2 del capítulo 4, en un pequeño número de unidades, el supuesto de la homogeneidad de las unidades es más útil cuando el grado de aleatoriedad es bastante bajo).

Si [la unidad] es la habitación de una casa, t [de «tratamiento»] significa que pulso el interruptor de la luz en esa habitación, c [de control] significa que no lo hago y [la variable dependiente] indica si la luz está encendida o no durante un corto periodo después de que se aplique t o c , de manera que podría *creer* que se pueden *conocer* los valores de [la variable dependiente, tanto en t como en c] al pulsar el interruptor. Sin embargo, está claro que mi *creencia* sólo la puede compartir alguien por la plausibilidad de ciertos supuestos relativos a la situación. Si, por ejemplo, la luz se ha estado apagando y encendiendo sin razón aparente mientras estoy pensando en realizar mi experimento, podría dudar de que se puedan conocer los valores de [la variable dependiente tanto en t como en c] después de pulsar el interruptor. ¿al menos hasta que fuera lo suficientemente ingenioso como para concebir un nuevo experimento!

En este ejemplo, el supuesto de homogeneidad de las unidades consiste en que si hubiéramos pulsado el interruptor (aplicado t , según la notación de Holland) en ambos periodos, el valor esperado (si la luz iba a estar encendida o no) habría sido el mismo. La homogeneidad de las unidades también presupone que si no hubiéramos apretado el interruptor (aplicado c) en los dos periodos, el valor esperado habría sido el mismo, aunque no necesariamente igual a cuando se aplicara t . Hay que señalar que, para asegurarnos de esto, tendríamos que haber puesto el interruptor en la posición de apagado después del primer experimento, pero también tendríamos que haber partido de un supuesto improbable: que pulsar el encendido del interruptor en el primer periodo no tiene como consecuencia los dos valores hipotéticos esperados en el siguiente (como sería el caso si se fundieran los plomos después de la primera pulsación). En general, el supuesto de homogeneidad de las unidades no puede comprobarse en una única unidad (aunque, en este caso, abriendo la pared e inspeccionando la instalación eléctrica podríamos generar algunas nuevas hipótesis en relación con el mecanismo causal).

El supuesto del *efecto constante* es una versión más débil, pero completamente aceptable, de la homogeneidad de las unidades. En lugar de presuponer que el valor esperado de la variable dependiente es el mismo en

diferentes unidades si también lo es el de la variable explicativa, sólo tendremos que presuponer que el efecto causal es constante. Ésta es una versión más endeble del supuesto de homogeneidad de las unidades porque el efecto causal no es más que la diferencia entre los dos valores esperados. Si los de las unidades cuya variable explicativa vale lo mismo cambian de igual manera, se vulnerará el supuesto de homogeneidad de las unidades, pero seguirá siendo válido el del efecto constante. Por ejemplo, en unas elecciones al Congreso estadounidense dos circunscripciones pueden variar en cuanto a la proporción de voto esperada para candidatos demócratas no congresistas (digamos, un 45% frente a un 65%), pero puede que el hecho de tener ese cargo siga aumentando en un 10% el voto a los candidatos demócratas en cualquier circunscripción.

El concepto de homogeneidad de las unidades (o el supuesto de los efectos causales constantes, que es menos exigente) está en la raíz de la investigación científica. Por ejemplo, es la premisa que subyace en los estudios de caso comparados. Contrastamos diversas unidades cuyas variables explicativas tienen diferentes valores y observamos los de las dependientes. Creemos que los resultados que observamos en estos últimos valores proceden de las diferencias que hay entre los de las variables explicativas que se relacionan con las observaciones. Lo que hemos planteado aquí es que, en este caso, nuestra «creencia» se apoya necesariamente en el supuesto de homogeneidad de las unidades o de los efectos constantes.

Hay que señalar que puede que busquemos unidades homogéneas a lo largo del tiempo o del espacio. Podemos comparar los votos que recibe un candidato demócrata, cuando es congresista y cuando no lo es, en la misma circunscripción y en diferentes ocasiones, así como en varias circunscripciones al mismo tiempo (o combinar ambas posibilidades). Como un efecto causal sólo puede estimarse y no conocerse, no resulta sorprendente que, en general, el supuesto de homogeneidad de las unidades no sea comprobable; sin embargo, es importante que se haga explícita su naturaleza. ¿En qué gama de unidades esperamos que tenga sentido nuestro supuesto de que estar en el poder tenga un efecto uniforme?, ¿en cualquiera de las elecciones para el Congreso?, ¿para el Congreso sí pero no para el Senado?, ¿sólo en las del norte?, ¿sólo en las de los últimos veinte años?

Véase de qué manera el supuesto de homogeneidad de las unidades se relaciona con el análisis que hicimos en el apartado 1.3 del capítulo 1 de la complejidad y de la «singularidad». Señalamos entonces que las generalizaciones de las ciencias sociales dependen de nuestra capacidad para simplificar la realidad coherentemente. En su forma más extrema, generalizar de ese modo con el fin de extraer inferencias causales precisa del respeto a las normas de homogeneidad de las unidades; para realizar el análisis, las observaciones que se examinan se hacen idénticas en aspectos importantes. Con frecuencia, resulta imposible lograr la homogeneidad de las unidades; las elecciones al Congreso, por no hablar de las revoluciones, apenas pue-

den compararse con interruptores de la luz. Sin embargo, entender el grado de homogeneidad de nuestras unidades de análisis nos ayudará a hacer una estimación del nivel de incertidumbre o de otros sesgos similares que puedan atribuirse a nuestras inferencias.

3.2 Independencia condicional

Según el supuesto de *independencia condicional*, los valores que se adjudican a las variables explicativas son independientes de los de las dependientes (la misma expresión se utiliza a veces en estadística, pero su definición no es la misma que suele tener en la teoría de la probabilidad). Esto quiere decir que, después de tener en cuenta las variables explicativas (o de controlarlas), el proceso de asignar valores a la variable explicativa es independiente de las (en general, dos o más) variables dependientes, Y_i y y_i' . Utilizamos la expresión «asignar valores» a las variables explicativas para describir el proceso mediante el cual éstas consiguen unos valores determinados. En el trabajo experimental el investigador *asigna* realmente valores a estas variables: unos objetos se colocan en el grupo de tratamiento y otros en el de control, mientras que en el trabajo no experimental los pueden «asignar» la naturaleza o el medio. En estos casos, lo esencial es que los valores de las variables explicativas no sean consecuencia de las dependientes. En el apartado 4 del capítulo 5 se describe el problema de la «endogeneidad», que aparece cuando las variables explicativas son consecuencia, al menos parcialmente, de las dependientes.

Los análisis de un n grande en los que hay que utilizar un sistema aleatorio, tanto de selección como de asignación de valores, son la forma más fiable de garantizar la independencia condicional y no necesitan del supuesto de homogeneidad de las unidades. Este sistema aleatorio de selección y asignación nos ayuda a extraer inferencias causales porque cumple *automáticamente* tres supuestos que subyacen en el concepto de independencia condicional: (1) que el proceso de asignación de valores a las variables explicativas es independiente de las dependientes (o sea, que no existe problema de endogeneidad), (2) que no hay sesgo de selección —problema que analizamos en el apartado 3 del capítulo 4— (3) y tampoco sesgo de la variable omitida (apartado 2 del capítulo 5). De este modo, si podemos cumplir de alguna manera estas condiciones, ya sea mediante el sistema aleatorio de selección y asignación (tal como se analiza en el apartado 2 del capítulo 4) o utilizando algún otro procedimiento, podremos evitar el problema fundamental de la inferencia causal.

Por fortuna, el sistema aleatorio de selección y asignación *no* es necesario para que se cumpla el supuesto de independencia condicional y, si el proceso de «asignación» de valores a las variables explicativas no es independiente de las dependientes, aún podremos cumplir dicho supuesto cono-

ciendo el proceso e incorporando una medida del mismo a nuestras variables de control. Por ejemplo, supongamos que nos interesa hacer una estimación de las consecuencias que tiene el nivel de segregación residencial en la intensidad del conflicto entre israelíes y palestinos en la Cisjordania ocupada por los primeros. Estaríamos transgrediendo considerablemente nuestro supuesto de independencia condicional si, para hallar el efecto causal, sólo nos ocupáramos de la relación entre estas dos variables. La razón es que puede que los israelíes y palestinos que deciden vivir en zonas segregadas lo hagan en virtud de una convicción ideológica que determina quién tiene realmente el derecho a estar en Cisjordania. Por lo tanto, el extremismo ideológico (en ambos bandos) puede llevar al conflicto. Una medida que creemos que representa la segregación residencial quizá sea, en realidad, un sustituto de la ideología. La diferencia entre ambas explicaciones puede ser bastante importante, ya que una nueva política de vivienda podría ayudar a remediar el conflicto si la segregación residencial fuera su auténtica causa, pero sería ineficaz, e incluso contraproducente, si su fuerza motriz es la ideología. Podríamos intentar solucionar este problema midiendo también de forma explícita la ideología de los residentes y controlando este factor. Por ejemplo, podríamos informarnos de qué apoyo tienen los partidos extremistas entre los israelíes y de la afiliación a la OLP entre los palestinos. Posteriormente, se podrían controlar las posibles consecuencias engañosas de la ideología comparando comunidades en las que haya un mismo nivel de extremismo ideológico pero un grado de segregación residencial distinto.

Cuando no es factible ni seleccionar ni asignar valores de manera aleatoria y tampoco se puede controlar el proceso de asignación y de selección, tenemos que recurrir a alguna versión del supuesto de homogeneidad de las unidades para hacer inferencias causales válidas. Como en las investigaciones sociales nunca se cumplirá perfectamente tal supuesto, habrá que poner un cuidado especial en especificar el grado de incertidumbre de las inferencias causales. Este supuesto se pondrá especialmente de manifiesto cuando analicemos, en el apartado 6 del capítulo 5, los procedimientos que se utilizan para analizar observaciones «equiparables».

4. Criterios para evaluar las inferencias causales

Recordemos que al definir la causalidad en función de variables aleatorias logramos establecer una estricta analogía entre ese fenómeno y otras características sistemáticas, como son la media o la varianza, en las que nos centramos al extraer inferencias descriptivas. Esta analogía nos permite evaluar las inferencias causales precisamente con los mismos criterios que utilizamos en el apartado 7 del capítulo 2 para juzgar las descriptivas: la falta de *sesgo* y la *eficiencia*. De ahí que gran parte de lo que dijimos sobre este

asunto en el capítulo 2 se aplique igualmente a los problemas de la inferencia causal que abordamos aquí. En este apartado, formalizamos brevemente las diferencias relativamente escasas que hay entre estas dos situaciones.

En el apartado 7 del capítulo 2 el objetivo de nuestra inferencia era una media (el valor esperado de una variable aleatoria) que denominamos μ y que conceptualizamos como un número fijo pero desconocido. Se dice que un estimador de μ no está sesgado si es igual a la media de μ en muchas reproducciones hipotéticas del mismo experimento.

Al igual que entonces, seguimos conceptualizando el valor esperado de un efecto causal aleatorio, denominado β , como un número fijo pero desconocido. Por tanto, la falta de sesgo se define de manera análoga: un estimador de β carece de sesgo si es igual a la media de β en muchas reproducciones hipotéticas del mismo experimento. La eficiencia también se define de forma parecida a la variación que se registra de una reproducción hipotética a otra. Éstos son conceptos muy importantes que, entre los capítulos 4 y 6, nos servirán de base para estudiar muchos de los problemas de la inferencia causal. En los dos recuadros siguientes se presentan definiciones formales.

5. Reglas para elaborar teorías causales

Muchos consejos sensatos para mejorar la investigación cualitativa son precisos, específicos y detallados, y se refieren a un aspecto manejable y, por tanto, delimitado de ese tipo de investigación. Sin embargo, aun cuando estemos enfrascados en la resolución de multitud de problemas concretos, hay que tener en mente una clara visión general: toda solución específica debe ayudarnos a solventar cualquiera que sea el problema general de la inferencia causal del que nos estemos ocupando. Hasta ahora, en este capítulo hemos proporcionado una precisa definición teórica de efecto causal y hemos analizado algunos de los problemas que conlleva la extracción de inferencias causales. Ahora vamos a dar un paso atrás para ofrecer una panorámica más amplia de ciertas normas relativas a la elaboración de teorías. Para nosotros (tal como se ha señalado en el apartado 2 del capítulo 1) la mejora de una teoría no acaba cuando comienza la recogida de datos.

Las *teorías causales* tienen como objetivo mostrar las causas de un fenómeno o conjunto de fenómenos. Cualquier teoría, independientemente de que en principio haya sido deductiva o inductiva, conlleva un conjunto interrelacionado de hipótesis causales. Cada una de ellas postula la existencia de una relación entre variables que genera consecuencias observables: si unas determinadas variables explicativas tienen ciertos valores, se predice que las dependientes tendrán otros valores específicos. Para comprobar o evaluar cualquier hipótesis causal se necesitan inferencias del mismo tipo. La teoría general, que se compone de hipótesis, ha de tener *coherencia interna* porque si no se plantearán hipótesis contradictorias.

Notación para un modelo formal de efecto causal

Con el fin de hacer más comprensibles los apartados siguientes, a continuación generalizamos nuestra notación. Casi siempre tendremos n realizaciones de la variable Y_i . En nuestro ejemplo cuantitativo corriente, n es el número de circunscripciones en las elecciones al Congreso (435), mientras que la realización y_i de la variable aleatoria Y_i es la proporción de votos demócratas observada (como 0,56%) en la circunscripción i , dentro del reparto entre ambos partidos. La proporción de votos esperada para el demócrata no congresista en ese reparto (la media de todas las reproducciones hipotéticas) en la circunscripción i es μ_i^0 . Definimos la variable explicativa como X_i y, en el presente ejemplo, la codificamos como cero cuando en la circunscripción i no se presenta un congresista demócrata y como uno cuando sí se presenta. De este modo, podemos expresar el efecto causal medio en la unidad i como:

$$\beta = E(Y_i | X_i = 1) - E(Y_i | X_i = 0) = \mu_i^1 - \mu_i^0 \quad (3.4)$$

e incorporarla al siguiente modelo formal simple:

$$\begin{aligned} E(Y_i) &= \mu_i^0 + X_i(\mu_i^1 - \mu_i^0) \\ &= \mu_i^0 + X_i\beta \end{aligned} \quad (3.5)$$

Así, cuando en la circunscripción i no se presenta un congresista e $X_i = 0$, el valor esperado se ve determinado por la sustitución de X_i por 0 en la ecuación, y la respuesta es la siguiente:

$$\begin{aligned} E(Y_i | X = 0) &= \mu_i^0 + (0)\beta \\ &= \mu_i^0 \end{aligned}$$

Las teorías e hipótesis que se ajustan a esta descripción abarcan una amplísima gama. En este apartado presentamos cinco reglas que ayudan a elaborar buenas teorías y analizaremos cada una de ellas mediante ejemplos.

5.1 Regla 1: elaborar teorías falsables

Con esta primera regla no sólo queremos decir que una «teoría» que no puede estar equivocada no es tal teoría, sino que las teorías hay que dise-

De la misma manera, cuando un congresista demócrata se presenta en la circunscripción i , el valor esperado es μ_i^D :

$$\begin{aligned} E(Y_i | X = 1) &= \mu_i^D + (1)\beta \\ &= \mu_i^D + \beta \\ &= \mu_i^D + (\mu_i^R - \mu_i^D) \\ &= \mu_i^R \end{aligned}$$

De este modo, la ecuación (3.5) proporciona un modelo útil de inferencia causal, siendo β —la diferencia entre las dos proporciones teóricas— nuestro efecto causal. Para terminar, simplificamos por última vez esta ecuación para poder remitirnos a ella en el futuro. Si presuponemos que la media de Y_i es cero (o que se representa como una desviación de esta media que no limita en modo alguno la aplicabilidad del modelo), podremos prescindir del término constante en esta ecuación y escribirla de forma más simple como:

$$E(Y_i) = X_i\beta \tag{3.6}$$

El parámetro β sigue siendo el valor teórico del efecto causal medio, una característica sistemática de las variables aleatorias y uno de nuestros objetivos en la inferencia causal. Este modelo constituye un caso especial del «análisis de regresión» que es habitual en la investigación cuantitativa, pero los coeficientes de regresión sólo coinciden a veces con las estimaciones de los efectos causales.

ñarlas para que se pueda demostrar de la forma más rápida y fácil posible que están equivocadas. Es evidente que no tenemos que intentar equivocarnos, pero incluso una teoría incorrecta es mejor que una afirmación que no es ni falsa ni cierta. Subrayar la necesidad de que las teorías sean falsables nos obliga a mantener una perspectiva adecuada sobre la incertidumbre y garantiza que vamos a considerar tales teorías como algo provisional, sin dejarlas convertirse en dogmas. Siempre hay que estar dispuesto a rechazar las teorías si existen suficientes pruebas científicas en su contra. Una de las preguntas que hay que plantearse en relación a una teoría (o a cualquier hipótesis que se derive de ella) es simplemente: ¿qué datos la falsarían? Esta

Un análisis formal de la falta de sesgo en las estimaciones causales

En este recuadro demostramos la falta de sesgo del estimador que representa el parámetro del efecto causal que aparecía en el apartado 1 de este capítulo. La notación y la lógica de estas ideas son bastante paralelas a las de la definición formal de falta de sesgo que se dio en el apartado 7 del capítulo 2, dentro del contexto de la inferencia descriptiva. El simple modelo lineal que tiene una variable explicativa y otra dependiente es como sigue¹²:

$$E(Y_i) = \beta X_i$$

Nuestra estimación de β es simplemente la de regresión por mínimos cuadrados:

$$b = \frac{\sum_{i=1}^n Y_i X_i}{\sum_{i=1}^n X_i^2} \quad (3.7)$$

Para determinar si b es un estimador no sesgado de β es preciso tomar el valor esperado, haciendo su media en las reproducciones hipotéticas:

$$\begin{aligned} E(b) &= E\left(\frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}\right) & (3.8) \\ &= \frac{\sum_{i=1}^n X_i E(Y_i)}{\sum_{i=1}^n X_i^2} \\ &= \frac{\sum_{i=1}^n X_i^2 \beta}{\sum_{i=1}^n X_i^2} \\ &= \beta \end{aligned}$$

lo cual demuestra que b es un estimador no sesgado de β .

pregunta debe aplicarse a cualquier teoría o hipótesis, pero, sobre todo, tiene que hacerlo el investigador que primero formula una teoría.

Karl Popper es el autor que más se relaciona con la idea de falsabilidad (Popper, 1968), y para él existe una asimetría fundamental entre confirmar

Un análisis formal de la eficiencia

A continuación, calibramos la eficiencia del estimador estándar que representa el parámetro del efecto causal β que aparecía en el apartado 1 de este capítulo. En la ecuación (3.8) hemos demostrado que este estimador no está sesgado, y ahora calculamos su varianza:

$$\begin{aligned} V(b) &= V\left(\frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}\right) & (3.9) \\ &= \frac{1}{\left(\sum_{i=1}^n X_i^2\right)^2} \sum_{i=1}^n X_i^2 V(Y_i) \\ &= \frac{V(Y_i)}{\sum_{i=1}^n X_i^2} \\ &= \frac{\sigma^2}{\sum_{i=1}^n X_i^2} \end{aligned}$$

En consecuencia, la varianza de este estimador es una función de dos componentes. En primer lugar, cuanto más aleatoria sea *cada* unidad en nuestros datos (la mayor es σ^2), más variable será nuestro estimador b , y esto no debería sorprendernos. Además, cuanto mayor sea la varianza observada en la variable explicativa $\left(\sum_{i=1}^n X_i^2\right)$, menos variable será nuestra estimación de b . En el caso extremo de que no haya variabilidad en X , nada podrá ayudarnos a estimar el efecto que tienen los cambios producidos en la variable explicativa sobre la dependiente, y, en este ejemplo, la fórmula pronosticará una varianza infinita (incertidumbre completa). De manera más general, este componente indica que la eficiencia es mayor cuando, en la variable explicativa, tenemos datos que proceden de una gama de valores más amplia. Por tanto, lo más habitual es que sea mejor evaluar nuestras hipótesis causales en tantas situaciones diferentes como sea posible. Se puede pensar en este último asunto como si consistiera en trazar una línea utilizando una regla, dos puntos sobre una página y poco pulso. Si los dos puntos están muy juntos (X tiene una pequeña varianza), los errores que se registren al colocar la regla serán mucho mayores que si estuvieran más separados (situación en la que X tiene una varianza grande).

una teoría (verificación) y desmentirla (falsación). Lo primero es prácticamente irrelevante, mientras que lo segundo es la clave de la ciencia. Popper cree que una teoría, una vez que se enuncia, pasa a convertirse de inmediato en parte integrante del cuerpo de conocimientos científicos aceptados. Al ser generales las teorías y específicas las hipótesis, las primeras implican, técnicamente, la existencia de un número infinito de las segundas. Sin embargo, sólo es posible someter a pruebas empíricas a un número limitado de hipótesis. En este sentido, «las teorías no son verificables» porque nunca podemos comprobar todas sus consecuencias observables (Popper, 1968, p. 252). Al comprobar cada hipótesis se puede demostrar que es coherente con la teoría, pero un determinado número de resultados empíricos coherentes no cambiarán nuestra opinión, ya que la teoría sigue constituyendo un conocimiento científico aceptado. Por otra parte, si se demuestra que una sola hipótesis es incorrecta y, por tanto, que no se ajusta a la teoría, ésta será falsada y dejará de pertenecer al bagaje del conocimiento humano. «Por lo tanto, la sucesión de comprobaciones no altera en absoluto la posición de ninguna hipótesis, pero el fracaso de una sola de esas pruebas sí puede tener muchas consecuencias» (Miller, 1988, p. 22). Popper no pretendía que la falsación fuera un concepto determinista y reconocía que toda inferencia empírica es, en cierto modo, incierta (Popper, 1982). En su análisis del desmentido escribió que «aunque se admita la asimetría [entre falsación y verificación], sigue siendo imposible, por varias razones, que cualquier sistema teórico sea falsado de manera concluyente» (Popper, 1969, p. 42).

Las ideas de Popper nos parecen fundamentales para *formular* teorías. Siempre hay que diseñarlas de manera que puedan ser falsadas y también debemos aprender de la insistencia de Popper en el carácter provisional de toda ellas. Sin embargo, para *evaluar* las teorías existentes en las ciencias sociales no resulta tan significativa la asimetría entre verificación y falsación. Ninguna de ellas contribuye a nuestro conocimiento científico. En cierto sentido general, la cuestión no es tanto si una teoría es o no falsa —casi todas las teorías sociales interesantes cuentan al menos con una consecuencia observable que parece equivocada—, sino *en qué medida nos ayuda dicha teoría a explicar el mundo*. Si siguiéramos la regla de Popper, las teorías que se basan en el supuesto de la elección racional hubieran sido rechazadas hace tiempo, ya que han sido falsadas en muchos ejemplos específicos. Sin embargo, los científicos sociales optan con frecuencia por mantener este supuesto, convenientemente modificado, porque les concede un poder considerable en muchas clases de problemas de investigación (véase Cook y Levi, 1990). Lo mismo puede decirse de casi cualquier otra teoría social de interés. De hecho, en las ciencias sociales, intentar falsar teorías es buscar los límites de su aplicabilidad. Aprendemos cuando alguna consecuencia observable indica que la teoría no se cumple y también cuando funciona correctamente.

Para los científicos (y especialmente para los sociales) que evalúan teorías formuladas con propiedad, la asimetría fundamental de Popper resulta prácticamente irrelevante. O'Hear (1989, p. 43) hizo una afirmación similar acerca de la aplicación de las ideas de Popper a las ciencias físicas:

Popper siempre suele hablar teniendo en mente *explicaciones de teorías universales*. Pero hay que insistir una vez más en que proponer y comprobar teorías universales es sólo uno de los objetivos de la ciencia. Puede que, debido a condiciones que han variado considerablemente a lo largo del tiempo y el espacio, no haya teorías universales verdaderas, y ésta es una posibilidad que no podemos subestimar. Sin embargo, aunque así fuera, la ciencia aún podría cumplir muchos de sus objetivos, al darnos conocimientos y hacer verdaderas predicciones acerca de las condiciones que constituyen y rodean nuestro restringido ámbito espacio-temporal.

Seguramente, esta misma idea puede aplicarse aún con más razón a las ciencias sociales.

Además, la evaluación que hace Popper de las teorías no establece una distinción fundamental entre las que acaban de formularse y las que ya han resistido numerosas pruebas empíricas. Cuando estamos comprobando la distinción determinista entre verdad o ficción en una teoría universal (de la cual no hay ejemplos interesantes), la idea de Popper es apropiada, pero no resulta útil para lo que nosotros pretendemos, que es buscar los límites de aplicación de una teoría. Como hemos señalado muchas veces en este libro, propugnamos que todas las inferencias relativas a hipótesis específicas se hagan planteando el mejor cálculo (una estimación) y midiendo su incertidumbre. La conclusión a la que lleguemos respecto a si la inferencia concuerda o no concuerda con nuestra teoría, tendrá el mismo efecto sobre la fe que tenemos en ésta. Tanto la coherencia como la incoherencia aportan información sobre la verdad de la teoría y tienen que influir en la certidumbre de nuestras creencias¹³.

Pensemos en la hipótesis de que las campañas electorales de los demócratas y de los republicanos durante las elecciones presidenciales estadounidenses tienen poco impacto global sobre su resultado. Esta hipótesis implica otras muchas de carácter más específico, como las que propugnan que los anuncios televisivos y radiofónicos, así como los debates, tampoco tienen mucha influencia en los votantes. En realidad, cualquier comprobación de la teoría debe contrastar una de estas hipótesis. Hay una comprobación de esta teoría que ha demostrado que se pueden hacer pronósticos más precisos del resultado utilizando las variables de que se dispone únicamente al hacerse las convenciones y, por lo tanto, antes de las campañas (Gelman y King, 1993). Esta comprobación concuerda con la teoría (si podemos predecir las elecciones antes de la campaña, no puede decirse que ésta tenga mucha importancia), pero no la verifica completamente. Puede que haya algún aspecto de la campaña que tenga alguna pequeña influencia que expli-

que ciertos errores de pronóstico (y pocos investigadores dudan de ello). Además, la predicción puede haber sido fruto de la suerte o quizá, en los años en que se han recogido los datos, las campañas no hayan incorporado ninguna táctica innovadora (y, por tanto, impredecible).

Podríamos llevar a cabo otras muchas comprobaciones, introduciendo en el modelo de pronóstico variables que midieran aspectos de la campaña, como la cantidad relativa de espacios televisivos o radiofónicos, la elocuencia de los candidatos y la evaluación de los resultados de los debates. Si todas estas hipótesis no tuvieran ningún efecto, Popper diría que nuestra opinión no cambia de forma reseñable, de manera que la teoría de que las campañas presidenciales no tienen impacto se mantendría. De hecho, si hiciéramos mil comprobaciones similares y todas concordaran con la teoría, ésta todavía podría estar equivocada, ya que no hemos puesto a prueba todas las infinitas variables posibles que miden una campaña. De manera que, aunque dispusiéramos de muchos resultados coherentes con la teoría, aún *podría* ser cierto que las campañas de elecciones presidenciales influyen en el comportamiento del votante.

Sin embargo, si se demostrara que un único acontecimiento de la campaña —como una importante acusación de comportamiento inmoral— tiene alguna influencia en los votantes, la teoría quedaría falsada. Según Popper, aunque esta teoría no se falsara de forma concluyente (lo cual reconoció que era imposible), aprenderíamos más de esta manera que con mil comprobaciones que concordaran con ella.

Para nosotros, no es así como se hacen las ciencias sociales o como deberían hacerse. Después de hacer mil comprobaciones a favor o en contra, no prescindiríamos de la teoría que afirma que las campañas no tienen impacto, sino que la modificaríamos para decir, quizá, que las campañas normales carecen de impacto, a no ser que haya pruebas concluyentes de que uno de los candidatos se ha comportado de forma inmoral. Sin embargo, como esta modificación haría más restrictiva nuestra teoría, necesitaríamos evaluarla con un nuevo conjunto de datos antes de confiar en su validez. La teoría seguiría siendo convincente, y, de alguna manera, con cada nueva evaluación empírica sabríamos más sobre sus límites. Cada prueba que pasa una teoría influye tanto en la estimación de su validez como en la incertidumbre de ese cálculo, y el impacto puede alcanzar nuestro grado de interés en que la teoría funcione.

En el análisis anterior hemos planteado una forma importante de abordar las teorías y también hicimos una advertencia: nuestro enfoque subraya el carácter contingente de teorías e hipótesis. Más adelante, propugnamos que se intente dar a ambas una aplicación amplia. Ésta es una estrategia útil para investigar, pero siempre debemos recordar que no es probable que las teorías sociales tengan una aplicación universal. Las teorías que se dicen aplicables a cualquier circunstancia, en cualquier lugar —algunas variantes del marxismo y de la teoría de la elección racional son ejemplos de esta

pretensión—, o bien se presentan de forma tautológica (en cuyo caso ni son verdaderas ni falsas) o de manera que es posible desmentirlas empíricamente (en cuyo caso descubriremos que hacen predicciones incorrectas). Las teorías sociales más útiles son válidas en determinadas condiciones (en campañas electorales en las que no hay pruebas concluyentes de comportamiento inmoral por parte de ningún candidato) o en ciertos contextos (en países industrializados, pero no en los menos desarrollados; en las campañas para elegir la Cámara de Representantes, pero no en las del Senado). Siempre hay que intentar concretar los límites de aplicación de la teoría o hipótesis. El paso siguiente es preguntarse: ¿por qué existen esos límites?, ¿qué tienen las elecciones al Senado que invalidan las generalizaciones aplicables a las de la Cámara de Representantes?, ¿qué hace que la industrialización transforme los efectos causales?, ¿qué variable falta en nuestro análisis que podría generar una teoría de aplicación más general? Al hacernos estas preguntas traspasamos los límites de nuestra teoría o hipótesis para mostrar qué factores hay que considerar con el fin de ampliar su campo de acción.

Sin embargo, hay que añadir una advertencia. Hemos señalado que el proceso de evaluación de teorías e hipótesis es flexible: ciertas comprobaciones empíricas no las confirman ni desmienten definitivamente. Cuando una prueba empírica no concuerda con nuestras expectativas teóricas, no tiramos la teoría inmediatamente por la borda. Podemos hacer varias cosas: llegar a la conclusión de que quizá los datos sólo son pobres por casualidad o ajustar lo que consideramos que es el ámbito de aplicación de una teoría o hipótesis aunque no se sostenga en un caso determinado y, en virtud de ese ajuste, seguir aceptando esa elaboración teórica. La ciencia avanza mediante tales ajustes, pero éstos pueden ser peligrosos. Si los llevamos demasiado lejos haremos que nuestras teorías sean invulnerables al desmentido. En consecuencia, hay que tener mucho cuidado al adaptar teorías para que concuerden con nuevos datos y no hay que forzarlas más allá de lo plausible, añadiendo numerosas excepciones y casos especiales.

Si nuestro estudio desmiente algún aspecto de la teoría podemos optar por mantenerla, pero incorporando una excepción. Este procedimiento resulta aceptable siempre y cuando se reconozca el hecho de que estamos reduciendo las pretensiones de la teoría. De este modo, ésta será menos valiosa porque explica menos: según nuestra terminología, tendremos menos *control* sobre el problema que queremos comprender¹⁴. Además, este enfoque puede producir una «teoría» que no sea más que una mezcla de diversas excepciones y conclusiones. En algún momento hay que estar dispuesto a prescindir por completo de ciertas teorías e hipótesis. Si hay demasiadas excepciones, la teoría debe rechazarse. De manera que, en sí misma, *la concisión, que es la preferencia normativa por teorías que tienen pocos componentes, no siempre es aplicable*. Lo único que necesitamos es nuestro concepto general de maximización del control, del que se deriva

completamente la idea de concisión cuando es útil. Pensar que, en gran medida, la ciencia consiste en explicar muchos fenómenos mediante unos pocos pone de manifiesto que las teorías de pocos componentes no son mejores ni peores. Para maximizar el control hay que intentar formular teorías que expliquen lo más posible con el menor número de elementos. A veces esta formulación se logra mediante la concisión, pero otras no. Podemos pensar en ejemplos en los que una teoría ligeramente más complicada explica una parte del mundo mucho más amplia. En tal situación, seguramente utilizaríamos una teoría poco concisa, ya que ésta maximiza mejor el control que la más concisa¹⁵.

5.2 Regla 2: construir teorías que tengan coherencia interna

Una teoría sin coherencia interna no sólo es falsable: es falsa. De hecho, ésta es la única situación en la que se conoce la veracidad de una teoría sin necesidad de datos empíricos: si tiene dos o más componentes que generan hipótesis contradictorias, no hay datos del mundo empírico que puedan sostenerla. No tendría que ser polémico garantizar la coherencia interna de las teorías, pero esa cualidad suele ser difícil de alcanzar. Se puede lograr mediante modelos formales, matemáticos. Es en la economía donde más desarrollada está la *creación de modelos formales*, aunque es cada vez más habitual en sociología, psicología, ciencia política, antropología y en otras disciplinas (véase Ordeshook, 1986). Los politólogos han desarrollado numerosas teorías esenciales a partir de modelos matemáticos en áreas como la elección racional, la social, los modelos electorales espaciales, la economía pública o la teoría de juegos. Con estas investigaciones se han logrado importantes resultados y una gran cantidad de hipótesis plausibles. Una de las aportaciones más valiosas de los modelos formales es su capacidad para poner de manifiesto la incoherencia de teorías planteadas mediante la expresión verbal.

Sin embargo, al igual que ocurre con otras hipótesis, los modelos formales no son explicaciones verificadas si no evalúan empíricamente sus predicciones. En la investigación social, la formalización nos ayuda a razonar con más claridad y garantiza realmente que nuestras ideas tengan coherencia interna, pero no resuelve los problemas de evaluación empírica de las teorías. En las ciencias sociales, los supuestos de los modelos formales suelen ser sencillas expresiones matemáticas o maneras de garantizar que se puede llegar a un equilibrio. Hay pocos autores que creen que el mundo político es matemático en la misma medida en que lo creen algunos físicos del mundo que ellos analizan. Por lo tanto, estos modelos formales no son más que modelos: abstracciones que hay que diferenciar del mundo que estudiamos. De hecho, algunas teorías formales hacen predicciones a partir de supuestos enormemente simplificadores y a veces no tienen gran valor

empírico. Sólo desde un punto de vista abstracto son más precisas que las teorías sociales no formales: sus predicciones acerca del mundo real no son más específicas porque sus condiciones no se corresponden, ni siquiera de forma aproximada, con las reales.

Simplificar es esencial para los modelos formales, al igual que lo es en toda investigación, pero hay que tener cuidado con las inferencias que se extraen de la realidad a partir de esos modelos. Por ejemplo, presuponer que todas las variables omitidas no tienen consecuencias sobre los resultados puede ser muy útil al construir modelos, y, precisamente, esto es lo que hacemos en muchos de los modelos formales cualitativos de este libro. Estos supuestos no suelen justificarse por ser rasgos del mundo: sólo se dice que son instrumentos útiles para construir modelos a partir de él. En consecuencia, sus resultados se refieren concretamente a situaciones en la que tales variables omitidas son irrelevantes y pueden parecerse o no a los del mundo real. No tenemos que comprobar el supuesto para elaborar el modelo y sus consecuencias, pero sí es *esencial* que lo pongamos a prueba durante la evaluación empírica. Para que el modelo formal sea útil, el supuesto no tiene que ser correcto, pero no podemos emplear en la elaboración de diseños de investigación empírica supuestos teóricos no comprobados o justificados. Por el contrario, para que las teorías formales sean útiles en los estudios empíricos, suele ser preciso complementarlas con otros elementos.

Un buen modelo formal debe ser abstracto, de manera que los rasgos clave del problema se pongan de manifiesto y pueda aplicarse fácilmente el razonamiento matemático. Así, consideremos un modelo formal sobre las consecuencias de la representación proporcional en los sistemas de partidos que implique que la primera fragmenta los segundos. La variable causal clave es el ordenamiento electoral, o sea, si es un sistema de representación proporcional en el que los partidos obtienen escaños en función del porcentaje de votos que logran o si es un sistema mayoritario en el que sólo hay un ganador en cada circunscripción. La variable dependiente es el número de partidos políticos, que suele denominarse «grado de fragmentación del sistema de partidos». La hipótesis principal es que los ordenamientos electorales que se basan en la representación proporcional generan más partidos políticos que los de tipo mayoritario. Para una mayor sencillez, este modelo podría incluir únicamente variables que midieran algunos rasgos esenciales del sistema electoral y también el grado de fragmentación del de partidos. El modelo sólo generaría una *hipótesis*, no una conclusión, acerca de la relación que existe entre la representación proporcional y la fragmentación del sistema de partidos en el mundo real, y esa hipótesis habría que comprobarla utilizando métodos cualitativos o cuantitativos empíricos.

Sin embargo, aunque una de las consecuencias de este modelo sea que la representación proporcional fragmenta los partidos políticos y aunque no se utilizaran otras variables en el modelo, en un análisis empírico sería absurdo

utilizar solamente dos. Un estudio que indique que los países que utilizan la representación proporcional tienen sistemas de partidos más fragmentados estará dejando de lado el problema de la endogeneidad (apartado 4 del capítulo 5), puesto que quizá los países que instauran sistemas electorales de tipo proporcional lo hagan precisamente porque ya tienen un sistema de partidos fragmentado. Del mismo modo, el sesgo de la variable omitida sería un problema porque también es probable que los países en los que existen profundas divisiones raciales, étnicas o religiosas tengan un sistema de partidos fragmentado y que su sistema electoral sea de tipo proporcional.

De este modo, parecen cumplirse las dos condiciones para que exista sesgo de la variable omitida (apartado 2 del capítulo 5): dicha variable se correlaciona tanto con la explicativa como con la dependiente, y, por tanto, todo análisis que prescindiera de la variable relativa a la división social producirá inferencias sesgadas.

Tendría que quedar claro que los modelos formales son extremadamente útiles para aclarar nuestras ideas y para desarrollar teorías con coherencia interna. Quizá sólo un modelo formal pueda poner de manifiesto y corregir las incoherencias internas de muchas teorías especialmente complejas que se expresan con palabras. Al mismo tiempo, no es probable que esas elaboraciones teóricas proporcionen modelos correctos para la comprobación empírica, y podemos asegurar que no sirven para evitar ninguno de los problemas empíricos de la inferencia científica.

5.3 Regla 3: seleccionar cuidadosamente las variables dependientes

Es evidente que al investigar todo debe hacerse con cuidado, pero elegir las variables, principalmente las dependientes, es una decisión especialmente importante. A continuación damos tres indicaciones (basadas en errores que se cometen con demasiada frecuencia en escritos cuantitativos y cualitativos):

En primer lugar, *las variables dependientes deben ser realmente dependientes*. Elegir una variable dependiente que en realidad produce cambios en las explicativas es un error muy habitual. Analizamos las consecuencias específicas de la endogeneidad y ciertas maneras de esquivarla en el apartado 4 del capítulo 5, pero hay que subrayar aquí este problema porque la forma más fácil de evitarlo es elegir variables explicativas claramente exógenas y variables dependientes manifiestamente endógenas.

En segundo lugar, *no hay que seleccionar observaciones que estén basadas en la variable dependiente para que ésta se mantenga constante*. Esto también puede parecer bastante obvio, pero hay investigadores que suelen elegir observaciones en las que la variable dependiente no cambia en absoluto (como en el ejemplo analizado en el apartado 3.1 del capítulo 4). Aunque el diseño de investigación no se haga con el propósito de que esta va-

riable se mantenga constante, puede resultar que lo sea. Sin embargo, esto no supondrá un problema a menos que hayamos predeterminado tal situación mediante la selección de criterios. Por ejemplo, supongamos que se seleccionan observaciones en dos categorías de una variable explicativa y que resulta que la dependiente es constante en los dos grupos. Éste no será más que un caso en el que el efecto causal estimado es cero.

Para terminar, tenemos que *elegir una variable dependiente que represente la variación que queremos explicar*. Aunque esta indicación parezca evidente, el asunto es bastante delicado, como pone de manifiesto Stanley Lieberson (1985, p. 100):

Un simple experimento gravitatorio, presente en el Ontario Science Centre de Toronto, inspira un ejemplo heurístico. En este montaje se deja caer una moneda y una pluma dentro de un tubo en el que se ha hecho el vacío, y ambas llegan al fondo casi al mismo tiempo. Como el vacío no es total, es de suponer que la moneda llegará al fondo un poquito antes que la pluma. En cualquier caso, imaginémosnos un estudio en el que se dejen caer varios objetos, sin que se tenga el control tan férreo que otorga el vacío (que es lo que ocurriría en una investigación social no experimental). Si los científicos sociales perciben que varía el tiempo que tarda cada objeto en llegar al suelo, lo normal es que quieran saber qué características determinan esa diferencia. Es probable que rasgos como la densidad y la forma de los objetos influyan en la velocidad de la caída si no existe el vacío. Si el investigador social tiene suerte, la conjunción de esos factores explicará completamente las diferencias de velocidad que presenta la caída de cada uno de los objetos y, si es así, se pondrá muy contento porque se habrán explicado todas las variaciones que hay entre un objeto y otro. Mediante el razonamiento habitual en las ciencias sociales, el investigador llegará a la conclusión de que se entiende completamente el fenómeno *porque se han explicado todas las diferencias que existen entre los objetos que se estudian*. Seguramente, debe de haber algo que falla en nuestros procedimientos si podemos abordar este problema sin siquiera referirnos a la gravedad.

En este ejemplo, los procedimientos del investigador sólo serían erróneos si la variable más interesante fuera la gravedad. Si ésta fuera la variable explicativa que nos preocupara, el experimento no la cambiaría (ya que sólo ocurre en un lugar) y, por tanto, no nos daría información sobre ella. Sin embargo, el experimento que describe Lieberson tendría un mayor interés si quisiéramos comprender las variaciones que se producen en el tiempo que tardan diferentes objetos en tocar el suelo, cuando se dejan caer desde una misma altura a diferentes niveles de presión del aire. De hecho, aunque lo supiéramos todo de la gravedad, este experimento nos seguiría dando una información valiosa, pero si, como presupone el autor, lo que nos interesa realmente es hacer una inferencia sobre el efecto causal de la gravedad, necesitaríamos una variable dependiente que cambiara nuestras observaciones cada vez que variara el grado de atracción de la gravedad. Del mismo modo, en las ciencias sociales hay que asegurarse de que realmente nos in-

teresa más nuestra variable dependiente que los factores contextuales que el diseño de investigación mantiene constantes.

En consecuencia, para obtener una estimación no sesgada del impacto de las variables explicativas es preciso que todo el abanico de posibles variaciones de la variable dependiente pueda ser resultado del experimento. Poner límites artificiales al abanico de posibilidades de la variable dependiente o a sus valores genera lo que definimos (en el apartado 3 del capítulo 4) como sesgo de selección. Por ejemplo, si nos interesan las condiciones que propician el estallido de conflictos armados, no podemos elegir sólo observaciones en las que ése sea precisamente el resultado. Tal estudio podría darnos mucha información sobre cómo varían las observaciones de conflictos armados (al igual que el experimento de la gravedad nos informaba de las que se registran en la velocidad con que caen diversos objetos), pero no nos permitiría examinar las causas de esos conflictos. Si esto es lo que queremos entender, será mejor utilizar un diseño en el que seleccionemos observaciones a partir de nuestras variables explicativas, dejándole a la dependiente la *posibilidad* de abarcar todas las situaciones posibles: en las que hay poco o ningún conflicto, en las que amenaza con producirse o en las que realmente estalla.

5.4 Regla 4: maximizar lo concreto

La cuarta regla, que surge de nuestra insistencia en la falsabilidad, la coherencia y la variación de la variable dependiente, es maximizar lo concreto. Siempre que sea posible, hay que elegir conceptos observables en vez de no observables. En las teorías de las ciencias sociales se utilizan con frecuencia conceptos abstractos, no observados, como los de utilidad, cultura, intenciones, motivaciones, identificación, inteligencia o interés nacional. Pueden ser muy útiles para *formular* teorías, pero también una carga para la *evaluación* empírica de teorías e hipótesis, a menos que se puedan definir de manera que sea posible observarlos y medirlos o, al menos, hacer ambas cosas con sus consecuencias. Plantear explicaciones mediante ideas como cultura, interés nacional o utilidad resultará sospechoso, a menos que tales conceptos se puedan calibrar al margen de la variable dependiente que estamos explicando. Cuando se utilizan tales términos en las explicaciones, es fácil caer en la tautología o no presentar consecuencias observables y diferenciadoras. La acción de un individuo o de una nación puede explicarse a partir de un deseo de maximizar la utilidad, materializar unas intenciones o trabajar en beneficio del interés nacional; pero la prueba de que esa acción ha logrado tales fines residirá en el hecho de que el actor o la nación hayan participado en ella. El investigador es responsable de formular una teoría que especifique con claridad y precisión qué consecuencias observables van a indicar su veracidad y a distinguirla de otras alternativas lógicas.

Con esta regla no queremos decir en modo alguno que conceptos como los de intencionalidad o motivación carezcan de importancia. Sólo pretendemos señalar que, en una ciencia *empírica* como la nuestra, la principal norma explicativa tiene que ser la verificación o falsación *empírica*. Siempre será más difícil y menos provechoso buscar pruebas empíricas para conceptos abstractos, inconmensurables e inobservables que para muchas ideas específicas y concretas concebidas de manera imperfecta. Cuanto más abstractos sean nuestros conceptos, más difusas serán sus consecuencias observables y menos posibilidades habrá de falsar la teoría.

Los investigadores suelen utilizar la siguiente estrategia. Parten de algún concepto abstracto como los mencionados anteriormente, aceptan que no pueden medirlo de forma directa y, en consecuencia, señalan indicadores específicos de dicho concepto que sí pueden medirse y los utilizan en sus explicaciones. La elección de ese indicador concreto se basa en que es observable. A veces, es *lo único* observable (por ejemplo, sólo hay datos disponibles de este fenómeno o es el único tipo de acontecimiento histórico del que se guarda información). Este aspecto de la investigación empírica es perfectamente respetable y, de hecho, suele ser necesario.

Sin embargo, tiene una parte mala, pues con frecuencia el indicador específico está lejos del concepto original y sólo se relaciona con él de forma indirecta e incierta. Quizá no sea en absoluto un indicador válido del concepto abstracto, pero el investigador, después de pedir perdón rápidamente por el desfase que hay entre uno y otro, adjudica al primero el nombre del segundo y sigue adelante como si estuviera midiendo el concepto directamente. Por desgracia, esta objetualización es frecuente en las ciencias sociales, quizá más en la investigación cuantitativa que en la cualitativa, pero demasiado habitual en ambas. Por ejemplo, el investigador utiliza las cifras que tiene sobre correos, comercio, turismo e intercambio de estudiantes para elaborar un índice de «integración social» en Europa: realiza una encuesta para averiguar si a los entrevistados les preocupa más el medio ambiente que ganar dinero, clasificándolos después como «materialistas» o «postmaterialistas», u observa que las instituciones federales estadounidenses se diferencian entre sí por el período medio de permanencia de los trabajadores en sus empleos y con este dato mide la «institucionalización» de dichos organismos.

Aquí tenemos que dejar claro lo que queremos decir. El desfase entre concepto e indicador es inevitable en gran parte de los estudios sociales, pero hay buenas razones para utilizar conceptos generales en vez de específicos: nos permiten ampliar nuestro marco de referencia y la aplicabilidad de las teorías. De este modo, podemos referirnos al poder legislativo en vez de utilizar categorías más estrechas, como parlamento, o instituciones concretas, como el Bundestag alemán. Podemos hablar de «órganos decisorios» en lugar de legislativo cuando queremos que nuestra teoría sea aplicable a una gama de instituciones aún más amplia (de hecho, en el siguiente apar-

tado lo recomendamos). La ciencia se basa en estas clasificaciones abstractas, si no estaríamos volviendo al resumen de los pormenores históricos. Sin embargo, nuestros términos abstractos y generales deben relacionarse en algún momento con conceptos específicos y mensurables que hagan posible la comprobación empírica. Esta relación —y la distancia que se recorre para establecerla— siempre hay que tenerla en mente y hacerla explícita. Además, el problema teórico que se planteé debe justificar el determinado nivel de abstracción elegido, que, a su vez, tiene que ayudar a relacionar una investigación concreta —en la que un determinado indicador es el actor clave— con un problema general. Para nosotros será una responsabilidad ver que se llevan a cabo nuevas investigaciones, que utilizan otros indicadores específicos, con el fin de reforzar el supuesto de que los nuestros están realmente relacionados con un concepto más amplio. Los términos abstractos anteriores —«integración social», «postmaterialismo» e «institucionalización»— pueden medirse de forma razonable con los indicadores específicos indicados. No negamos que haya que dar un salto desde el indicador específico hasta el concepto abstracto general —en las ciencias sociales esto es preciso para continuar la investigación—, pero hay que darlo con cuidado, justificándolo y «recordando» siempre de dónde se parte.

Por lo tanto, no nos oponemos a las abstracciones, pero pedimos que el lenguaje de la investigación social sea lo más concreto y preciso posible. Si, como suele ocurrir en las ciencias sociales, no hay más remedio que utilizar elaboraciones teóricas inobservables, al menos tendremos que *elegir ideas con consecuencias observables*. Por ejemplo, la «inteligencia» nunca se ha observado directamente, pero no deja de ser por ello un concepto muy útil. Existen numerosos tests y otras formas de evaluar sus implicaciones. Por otra parte, si podemos elegir entre «institucionalización de la presidencia» y «número de empleados de la Casa Blanca», suele ser mejor quedarse con lo segundo. Podemos señalar que este concepto se relaciona con el de institucionalización de la presidencia —que es más general—, pero no hay que concretar el concepto limitado hasta el punto de considerar que es idéntico al más amplio. Además, si número de empleados significara institucionalización, tendríamos que poder encontrar otras medidas de este segundo concepto que respondieran a las mismas variables explicativas que el primero. Más adelante analizaremos cómo se «maximiza el control» ampliando las variables dependientes.

Nuestra defensa de la concreción se extiende en general a las palabras que utilizamos para describir nuestra teoría. Si el lector tiene que pasar mucho tiempo averiguando los significados precisos de una teoría, ésta tendrá menos utilidad. Al describir teorías, hay que reducir al mínimo la polémica sobre nuestras intenciones. Para lograr esta mayor especificidad, aunque no estemos llevando a cabo investigación empírica, hay que dedicar cierto tiempo a considerar, explícitamente, las consecuencias observables de la teoría e incluso los posibles proyectos de investigación que podríamos realizar.

Cuanto más impreciso sea nuestro lenguaje, menos posibilidades habrá de que nos equivoquemos, pero también de que nuestro trabajo sea útil. Es mejor equivocarse que ser impreciso.

En nuestra opinión, al presentar la lógica de un proyecto de investigación, defender su importancia y describir los acontecimientos en detalle, hay que cultivar una redacción elegante (y disfrutarla), práctica que constituye un bien escaso en las ciencias sociales. El aburrimiento nunca ha hecho avanzar ninguna ciencia. Sin embargo, cuando el objeto de estudio se transforma en una inferencia causal o descriptiva, en la que nos interesan observaciones y generalizaciones que esperamos se mantengan, es preciso utilizar una terminología y una forma de pensar concretas y específicas¹⁶.

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

Aceptando los condicionantes que implica asegurarse de que una teoría es falsable y de que maximizamos la concreción, hay que formular tal teoría de manera que explique una parte del mundo lo más grande posible. Nos damos cuenta de que hay cierta tensión entre esta quinta regla y el hecho de que antes instáramos a la concreción. Sólo podemos decir que ambos objetivos son importantes, aunque en muchos casos pueden contraponerse y hay que tenerlos presentes para llegar a un equilibrio.

Por ejemplo, no debemos presentar nuestra teoría como si sólo fuera aplicable al Bundestag alemán, cuando hay razones para creer que pudiera serlo en todo poder legislativo independiente. Para formularla no es preciso dar pruebas de todas sus consecuencias, siempre que se haga una estimación razonable de la incertidumbre que conlleva. Puede que hayamos dado pruebas consistentes de que la teoría es aplicable al Bundestag alemán. Aunque no las tengamos de que funciona en otros ámbitos, tampoco las tenemos de que no sea así. Un marco de referencia más amplio resulta útil si no perdemos de vista el hecho de que es necesario evaluar su aplicabilidad. De hecho, expresarla en relación a un marco de referencia hipotéticamente más amplio podría obligarnos a pensar en qué rasgos estructurales de la teoría la harían aplicable o no a otros poderes legislativos independientes. ¿Sería adecuada, por ejemplo, para el Senado estadounidense, cuyos integrantes no se eligen de una sola vez sino de manera escalonada, para la asamblea de New Hampshire, que es mucho más numerosa en relación al número de electores, o para la Cámara de los Comunes británica, en la que es mucho más fuerte el voto de partido? Presentar qué rasgos sistemáticos de la teoría creemos que la hacen aplicable en diferentes áreas es un importante ejercicio. Quizá así nos demos cuenta de que estamos equivocados, pero esto es bastante mejor que no formular la teoría con suficiente precisión desde el principio.

Puede que parezca que esta regla entra en conflicto con la preferencia que mostró Robert Merton ([1949], 1968) por las «teorías de alcance medio», pero incluso una lectura rápida de este autor nos indicará que no es así. Merton estaba reaccionando contra una tradición sociológica en la que había «teorías», como la de «la acción» de Parsons, que se formulaban de manera tan amplia que no podían falsarse. En la ciencia política, la «teoría de sistemas» de Easton (1965) se halla en la misma tradición (véase Eckstein, 1975, p. 90). Éste es un ejemplo del tipo de críticas que a Merton ([1949], 1968, p. 43) le gustaba hacer: «Por lo que se ve, la teoría de la fijación de los roles no concuerda con orientaciones teóricas tan generales como el marxismo, el análisis funcionalista, el conductismo social, la sociología integral de Sorokin o la teoría de la acción de Parsons». Merton no critica la teoría de la fijación de los roles, que consideraba de alcance medio, sino que ataca las «amplias orientaciones teóricas» con las que concuerda casi cualquier teoría más específica u observación empírica. Merton está a favor de las teorías de «alcance medio», pero creemos que estaría de acuerdo en que hay que formular teorías de la manera más general posible, siempre que sigan siendo falsables y concretas. Para retomar un concepto que hemos mencionado anteriormente, enunciar teorías de este modo es maximizar el control. Si la teoría puede comprobarse —y es evidente que el peligro de las más generales es que quizá estén expresadas de manera que no se pueda—, cuanto más general sea, mejor. O sea, a mayor amplitud, más control.

Lectura 6

Capítulo 2

METODOLOGIA DE LA INVESTIGACION POLITICA

Stefano Bartolini

1. Introducción

La expresión «metodología» se usa en las ciencias del hombre con diversos significados. Por una parte, con frecuencia se incluye en su esfera de competencia el estudio del fundamento filosófico del conocimiento de tipo científico, la teoría o filosofía del método científico, o sea su interpretación, evaluación y justificación en referencia a otros métodos y otras consideraciones. Por otra, aún con mayor frecuencia se entiende por metodología el estudio de técnicas específicas de investigación y de su lógica, de los instrumentos y las operaciones necesarios para crear tales técnicas y para interpretar sus resultados. En este caso la reflexión metodológica tiende inevitablemente a centrarse y a privilegiar los problemas de aquellos sectores y campos de la investigación social que recurren con más frecuencia a las técnicas estadísticas y matemáticas de tratamiento de datos, dedicando poca atención a ese amplio sector de investigación que no utiliza —o no lo hace sistemáticamente— tales técnicas. Está, por fin, un tercer nivel o acepción del término metodología, que, por su ubicación intermedia entre la filosofía del método y la lógica de las técnicas, es el más difícil de definir en términos positivos, tanto porque en el intento de indicar su contenido específico nos vemos remitidos continuamente hacia arriba, hacia el nivel y la acepción epistemológica, o hacia abajo, hacia las aplicaciones instrumentales. Esta tercera acepción de metodología de la investigación incluye aquellos procedimientos lógicos que se refieren a la formulación de los problemas de investigación, la formación y el tratamiento de los conceptos, la elección de los casos y las variables, los procedimientos de control de los resultados. En este sentido, la metodología de la investigación trata los métodos en términos de sus objetivos putativos y de los intereses de los usuarios. Muchos de los problemas citados, caen dentro también de la esfera de competencia de la metodología entendida como reflexión sobre las técnicas de investigación. Pero, además, en tal caso estos problemas se tratan haciendo referencia concreta

Tomada de: Stefano Bartolini, "Metodología de la investigación política", en *Manual de ciencia política*, Madrid, Alianza, 1996, pp. 39-79.

a las técnicas que se estudian, y no como problemas generales de método inherentes a la investigación política y social en su conjunto.

En este capítulo trataremos, pues, de una metodología entendida estrictamente como lógica del método de investigación, evitando discutir tanto los problemas epistemológicos que planean sobre toda empresa cognoscitiva de carácter científico, como las técnicas concretas usadas en las investigaciones cuantitativas. Nos centraremos en el nivel intermedio al que nos hemos referido antes, tratando sus procedimientos generales. En la mayor parte de los casos estos procedimientos metodológicos suponen una racionalización de los procedimientos dictados por el sentido común que cotidianamente permiten efectuar descubrimientos, producir afirmaciones y hacer opciones. Esta racionalización de los procedimientos del sentido común implica un «conocimiento» y una «potenciación» de ellos que permita formular proposiciones sobre la base de un saber cuyo carácter empírico y lógicamente no contradictorio y cuyos procedimientos públicos y repetibles —que puedan ser retomados por otros y repetidos para controlar la validez de los resultados— constituyen la base de la posibilidad de decidir y juzgar su valor de verdad. De estos dos elementos deriva el carácter de «objetividad» científica, entendida en el sentido de intersubjetividad de ese conocimiento¹.

Esta caracterización del conocimiento científico plantea *condiciones mínimas* a respetar en toda investigación y deliberadamente no recurre a definiciones del método científico que provienen de disciplinas más desarrolladas y que implican la necesaria referencia a «leyes» y «explicaciones» en sentido fuerte. En la situación actual, la investigación social y política en su conjunto no está en condiciones de satisfacer los requisitos de una estricta definición de método y explicación científica. Apegarse a tal estándar proponiendo comparaciones con las ciencias naturales más desarrolladas determina de hecho o bien la limitación de la aplicación de una rígida metodología sólo a aquellos campos de la investigación político-social en que el recurso a técnicas estadístico-matemáticas permite una aproximación mayor al modelo científico-físico, o bien una distancia incolmable entre las aspiraciones metodológicas y la realidad de la investigación diaria. El poner el acento sobre estándares metodológicos inalcanzables por las actuales ciencias sociales puede producir daños aún más graves: un rechazo o abandono de toda disciplina o precaución metodológica bajo el pretexto de la imposibilidad de alcanzar el tipo ideal de investigación y de explicación científica.

En una disciplina como la ciencia política, la posibilidad de evitar los riesgos inherentes a estas posiciones extremas está ligada al desarrollo y a la potenciación de ese nivel intermedio de conocimiento metodológico bosquejado antes. Por este motivo centraremos en él la atención, tratando de ofrecer una guía de los procedimientos de método que un investigador enzarzado con un problema de investigación debe seguir para obtener resultados lo más *resolubles lógicamente y empíricamente, transparentes y repetibles posible*.

2. Selección y formulación de los problemas

La investigación científica empieza con la selección y la correcta formulación de los problemas de investigación. Este primer paso influye sobre todos los demás

¹ BRUSCHI, A., *Scienza, tecnica e senso comune*, en A. BRUSCHI, U. GORI y F. ARRINÀ, *Relazioni internazionali. Metodi e tecniche di analisi*, Milán, Etas Compass, 1973, pp. 3-53

procesos. En la elección de un problema el investigador se guía por una gran variedad de factores, entre los cuales están, naturalmente, sus intereses y valores. Además de ser importante subjetivamente, el problema escogido debe también ser importante objetivamente, pero hay que reconocer que en las ciencias humanas existen muy pocos criterios precisos y «objetivos» para ayudar al investigador en su elección. La mayoría de los físicos tienden en general a estar de acuerdo acerca de en qué campos y en qué direcciones de investigación tienen que profundizar fructíferamente en un determinado período. Y ello porque la acumulación y la integración de sus conocimientos es mayor y tienen una visión más clara de las disponibilidades de fuentes y recursos para afrontar un estudio. En las ciencias humanas, por el contrario, los investigadores trabajan normalmente sobre problemas que ellos consideran subjetivamente importantes y están dispuestos a dejar a otros investigadores estudiar áreas y problemas que les parezcan importantes. Por otra parte, los científicos sociales no pueden limitarse a tratar sólo los problemas para los que disponen de exhaustivas fuentes y recursos. Muchos problemas interesantes e importantes social y políticamente tienen que afrontarse incluso en situaciones de carencia de fuentes y recursos de investigación, es decir, en condiciones en que tales problemas no son susceptibles de respuestas satisfactorias según los normales cánones científicos.

Sin embargo, también en las ciencias sociales el investigador puede encontrar una guía para la elección del problema a investigar si trata de maximizar la importancia teórica de su trabajo. En este caso deberá seleccionar un problema que, en un momento dado y en relación a un nivel determinado de acumulación de conocimientos de ese terreno, sea susceptible de ofrecer resultados nuevos, capaces de modificar las posturas científicas y teorías existentes en ese campo. Realizar una elección según este criterio requiere una gran familiaridad con un amplio espectro de los conocimientos y teorías existentes, así como ideas claras sobre los sectores en los que tales conocimientos y teorías son más débiles y tienen mayor necesidad de profundización y modificación.

En conclusión, pues, en las ciencias sociales las tres dimensiones de la importancia de un problema —la *personal-subjetiva*, la *externa político-social* y la *teórica* respecto al cuerpo de conocimientos acumulados— pueden ser motivos igualmente válidos para la elección.

Los interrogantes a que se enfrentan las ciencias humanas pueden tener un carácter predominantemente empírico, analítico o normativo. Los problemas empíricos están representados por interrogantes cuya respuesta depende y se busca en las experiencias sensoriales. Los problemas analíticos son, en cambio, aquellos cuya respuesta depende o se busca en un análisis del significado de las palabras que los expresan. Los problemas analíticos son esencialmente problemas lingüísticos y conceptuales, cuya solución no depende de la confrontación con la realidad empírica, sino de la definición de las palabras en las frases que los expresan: son, pues, problemas semánticos. Por último, los problemas normativos son aquellos problemas cuya resolución depende necesariamente de juicios de valor referidos a lo que es preferible, deseable o moral.

Sin adentrarnos en una discusión más profunda de esta distinción, basta aquí subrayar que los problemas afrontados por la ciencia política son predominantemente de carácter empírico. Sin embargo, en una disciplina que obligatoriamente recurre a términos y conceptos de empleo corriente en el lenguaje común y que no dispone —o sólo parcialmente— de su propio vocabulario especializado, se plantean continuamente problemas inherentes a la clarificación semántica de los

propios conceptos. Así pues, difícilmente en las investigaciones políticas los problemas empíricos pueden formularse bien sin recurrir a fases de clarificación analítica.

Por otra parte, es necesario subrayar que ciertos problemas políticos normativos —como los que desde hace siglos están en el centro del pensamiento y la filosofía política— pueden reformularse de manera que se obtenga el máximo de información empírica para sostener el juicio o la opción de valor. Por ejemplo, en la discusión clásica de si la democracia es un régimen mejor (o preferible a) que la dictadura, el problema se puede formular de tal modo que se refiera a una serie de consecuencias empíricas de los dos tipos de régimen, como la cantidad de censura ejercida sobre los medios de comunicación, el número de personas arrestadas por delitos de opinión, y otros aspectos. Estas informaciones de carácter empírico —en tanto no excluyen el juicio de valor del análisis, es decir, si una más estricta censura sea mejor o peor que una más blanda— amplían u objetivan el ámbito de las informaciones sobre las que se basan el juicio de valor o la indicación prescriptiva.

Por último, hay que recordar también que los resultados de investigaciones empíricas suelen utilizarse como base de prescripciones normativas por parte de los investigadores. Es evidente, pues, que los tres tipos de problemas se pueden mezclar en una misma investigación. Ello no supone un problema mientras que se mantengan claramente diferenciados dentro de ella. Es importante sobre todo evitar resolver un tipo de problemas recurriendo a respuestas de otro tipo: responder a preguntas empíricas con argumentos normativos, a preguntas analíticas con respuestas empíricas, y así sucesivamente.

De lo expuesto hasta ahora resulta claro que la *formulación* de un problema es un proceso tan importante como su elección. Existen algunas reglas que pueden ayudar al investigador. La formulación de un problema empírico para la investigación debería ser en primer lugar lo más *explícita* posible; es decir, explicitar al máximo la interrogante central de la investigación, separándola y distinguiéndola de las interrogantes periféricas o secundarias. En segundo lugar, la formulación del problema debe ser *clara* en el sentido de que los términos utilizados en ella no deben generar ambigüedad ni confusión intelectual acerca de su significado, ni deben encontrarse en ellos asertos o respuestas implícitas. En tercer lugar, el problema debe formularse de tal modo que pueda tener una *respuesta empírica*, excluyendo, pues, aquellas cuestiones que, indirectamente o de modo incompleto, no son susceptibles de tales respuestas. Por último, una correcta formulación del problema debe estar dotada de *valor teórico*, debe estar formulada de tal modo que se inserte en y contribuya al desarrollo de los conocimientos ya acumulados.

Puede parecer exagerado que nos detengamos en estas indicaciones que parecen de sentido común. Sin embargo, está claro que no es raro encontrarse con estudios políticos que se caracterizan por una pésima —y a veces ninguna— formulación del problema. A menudo, lo que se identifica es más el objeto de estudio que los problemas y los temas teóricos inherentes a él. Otras veces la formulación del problema está ampliamente implícita o diluida en medio de otros problemas no centrales o con consideraciones de otra naturaleza. También es frecuente la formulación poco clara de los problemas, hecha con términos ambiguos y sujeta a más de una interpretación de significado. En estos casos se descarga sobre el lector la pesada tarea de reconstruir el problema de base de la investigación. Una mala formulación del problema reduce, pues, el valor científico del trabajo en cuanto empece la comunicación intersubjetiva de sus resultados. Dispo-

ner como punto de partida de un problema importante, formulado de modo explícito, claro, empíricamente controlable y teóricamente relevante significa haber planteado de entrada de la mejor manera toda la investigación.

3. Las estrategias de investigación

Una formulación correcta del problema en el sentido expuesto más arriba ofrece la ventaja de indicar ya con la suficiente claridad qué unidades y qué propiedades o características de tales unidades deberá tomar en consideración la investigación. Las unidades son los tipos de objetos o acontecimientos de que se ocupa una determinada investigación científica; por ejemplo, elecciones, partidos, discursos televisados, individuos, programas electorales u otros. Las propiedades, en cambio, son las características o las dimensiones de las unidades que se consideran importantes para la investigación. Por ejemplo, para un estudio que tiene como unidad las elecciones políticas, una determinada formulación del problema

Propiedades \ Unidades	P1	P2	P3	P4	P5	P6	...	Pn
U1	S1-1	S1-2	S1-3	S1-4	S1-5	S1-6		S1-n
U2	S2-1	S2-2	S2-3					S2-n
U3								S3-n
U4								S4-n
U5								S5-n
...								
Un	Sn-1							Sn-n

FIG. 2.1.—La matriz de datos.

podría considerar importantes ciertas propiedades de las elecciones tales como el número de partidos que participan en ellas, el sistema electoral utilizado, la dimensión media de las circunscripciones.

La figura 2.1 muestra la matriz de datos que hay que construir mentalmente para toda investigación. Las casillas identifican el estado (o el valor) de cada determinada propiedad para cada unidad determinada. Concretamente, en el ejemplo de una investigación cuyas unidades de análisis sean los individuos miembros del Parlamento y entre cuyas propiedades figure su antigüedad parlamentaria, en la casilla de cruce de cada unidad (individuo) con la citada propiedad figurará el número efectivo de años que el individuo ha pasado en el Parlamento (o cualquier otra medida de esa antigüedad que se haya decidido utilizar)².

La matriz de los datos atrae claramente la atención —independientemente del problema de investigación a que se aplique— sobre dos problemas iniciales: a) la existencia de diferentes estrategias de investigación según la combinación entre *número de unidades* y *número de propiedades* que se consideran; b) la necesidad de realizar una selección de las unidades y de sus propiedades importantes para un estudio concreto.

El primer problema es una cuestión de *estrategia de investigación*; el segundo es un problema de *elección de los casos*. Veamos el primero de los dos. Según el número de unidades y de sus propiedades que se toman en consideración se pueden configurar esencialmente dos estrategias principales de investigación. La primera tiene carácter *intensivo*; considerando muy pocas unidades —o sólo una, al límite— se estudia en ellas un elevado número de propiedades y dimensiones. Para fines ilustrativos podemos caracterizar esta estrategia como típica del enfoque histórico, pero se suele recurrir a ella también en las ciencias sociales, en especial en el análisis de casos, estrategia sobre la que nos detendremos más adelante. La segunda estrategia tiene carácter *extensivo* en cuanto que se caracteriza por la toma en consideración de una o muy pocas dimensiones/propiedades en un gran número de casos. En su forma más extensa, este es el caso de los sondeos de tipo Gallup, en los que, precisamente, un gran número de ciudadanos son entrevistados sobre una o algunas cuestiones concretas, como, por ejemplo, si tienen confianza en el líder o en el partido en el poder. La figura 2.2 resume estas posibles estrategias y pone de manifiesto las direcciones en que puede moverse la investigación. También pone de relieve otra dimensión importante: la temporal. Desde este punto de vista la estrategia de investigación puede calificarse como *sincrónica* o *diacrónica* según que el estudio de las unidades y de sus propiedades se limite a un intervalo de tiempo relativamente corto o abarque un período de tiempo largo. Una investigación diacrónica aumenta notablemente el número de observaciones y de informaciones que el investigador debe recoger y analizar, en cuanto que los estados/valores de las propiedades/dimensiones de los casos se modifican y «varían» no sólo de caso en caso, sino también para el mismo caso de un período temporal a otro. El ejemplo más conocido de estrategias de investigaciones diacrónicas es el de los estudios sobre el desarrollo y la modernización.

Las diversas combinaciones posibles muestran la rica variedad de estrategias posibles de investigación. Naturalmente, no todas son perseguibles con la misma facilidad y, sobre todo, no todas sirven a los mismos fines. Existe una relación inversa entre el número de casos y de propiedades en el sentido de que, si se

² GALTUNG, J., *Theory and Methods of Social Research*, Oslo, Universitetsforlaget, 1967.

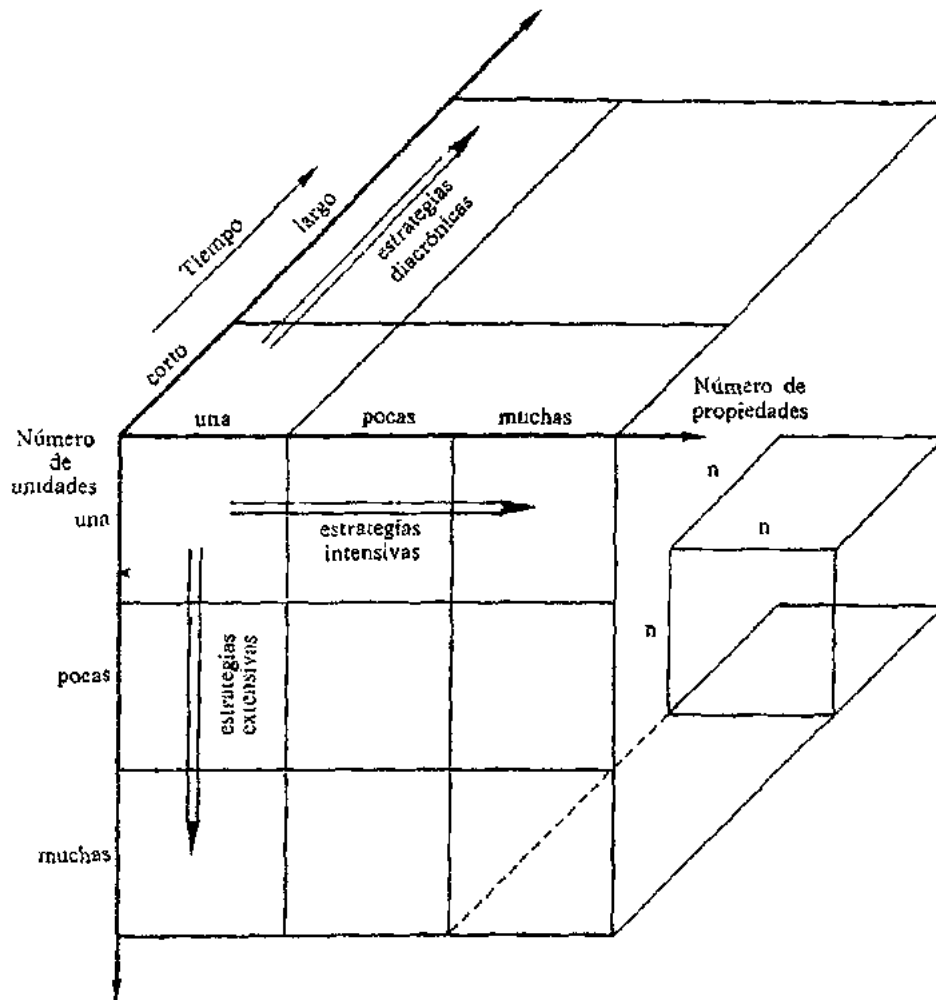


FIG. 2.2.—Matriz de datos y estrategias de investigación.

extiende considerablemente el número de unidades, para mantener la posibilidad de controlar mentalmente la investigación es necesario limitar el análisis a un número relativamente reducido de propiedades. Y viceversa, si se quieren considerar cuantas más propiedades sea posible de la unidad, es necesario limitar nuestro estudio a pocas unidades o a una sola. La lógica es simplemente la de los límites humanos de recursos físicos y mentales. Un estudio que tome en consideración muchos casos y propiedades en un amplio marco temporal resulta difícilmente organizable y controlable incluso para un equipo de investigadores, incluso con la ayuda que hoy se nos ofrece de la posibilidad de recurrir a ordenadores para organizar grandes cantidades de datos e informaciones.

El escoger una estrategia más o menos intensiva o extensiva, diacrónica o sincrónica depende, además de consideraciones prácticas de gran importancia, de los fines de la investigación. Se puede querer estudiar una sola unidad por la sencilla razón de que es precisamente esa la que se quiere conocer del modo más profundo posible. Si, en cambio, queremos llegar a conclusiones más generales sobre un fenómeno —por ejemplo, un cierto tipo de régimen político—, estudiar sólo uno de ellos supone restricciones enormes en lo que se refiere a la posibilidad de generalizar a otros casos las conclusiones a que se llega. Las investigaciones muy intensivas tendrán, en general, poca capacidad de formular propuestas generales, mientras que las muy extensivas no podrán examinar un número lo suficientemente alto de propiedades que permita controlar todas las condiciones importantes del contexto³. Volveremos sobre estos problemas más ampliamente en la parte dedicada a las relaciones entre variables y a los métodos de control. En cualquier caso, las distintas estrategias de investigación no están concebidas en términos contradictorios ni alternativos. Se puede tener que recurrir a distintas estrategias incluso en un mismo estudio según el tipo de fuentes o de recursos de que se disponga. Hipótesis que surgen de investigaciones de un tipo pueden ser sometidas a control de investigaciones basadas en estrategias diferentes.

4. La elección de los casos

Cualquiera que sea la estrategia de investigación escogida de antemano, si el número de unidades es mayor de uno puede surgir la necesidad de limitar el estudio a algunas de ellas; es decir, de seleccionar del universo de unidades potenciales un número reducido de ellas que serán los casos específicos de la investigación. No tiene ningún valor intrínseco el limitar el número de casos estudiados respecto a las unidades del universo si no es el práctico de ahorrar recursos y poder seguir estrategias de investigación más intensivas. Siempre que sea posible y deseable está bien estudiar el universo de las unidades. Sin embargo, cada vez que la situación impone que se centre la investigación sobre un número de unidades del universo se plantean dos problemas: a) el de cuántos casos se van a observar (número de casos); b) el de qué casos se van a observar (tipo de casos).

La elección de los casos, pues, afecta a la determinación de qué y cuántas observaciones tienen que hacerse cuando no es posible o conveniente hacer todas las que idealmente sería deseable. Mediante esta elección se reducen los costes de la investigación, pero se plantea inevitablemente un problema adicional que se refiere a la posibilidad de generalizar al universo de las unidades las propuestas expresadas por la muestra de los casos⁴.

Este problema tiene soluciones satisfactorias cuando operamos sobre grandes números. Entonces podemos recurrir a diversas técnicas estadísticas de muestreo que, sobre la base de las teorías de la probabilidad, minimizan la dimensión de la muestra (número de casos) en relación a un determinado nivel de probabilidad de error de las estimaciones que se pueden obtener de ella respecto a los valores efectivos del universo. En otras palabras, la lógica de las técnicas estadísticas de

³ GALTUNG, J., *Theory and Methods*, op. cit.

⁴ ACKOFF, R. L., *The Scientific Method. Optimizing Applied Research Decisions*, Nueva York, John Wiley and Sons, 1962.

muestreo es la de optimizar la relación entre número de casos a analizar y el error de estimación que puede provocar³.

Con frecuencia, sin embargo, y sobre todo por muchos problemas típicos de la investigación política, no es posible o conveniente recurrir a estas técnicas de elección de casos. Se debe, pues, recurrir a otros medios para obtener una muestra representativa, o bien abandonar simplemente este objetivo y seleccionar los casos en base a criterios específicos de no representatividad. En efecto, en muchas situaciones, el universo de las unidades es tan restringido y/o el número de los casos que de hecho podemos estudiar es tan bajo que una muestra establecida a partir de reglas prácticas que guían el juicio de elección del investigador puede dar errores de muestreo inferiores a los que, en el mismo caso, se producirían con una muestra probabilística. En esta hipótesis, pues, la elección del número y del tipo de casos es un procedimiento basado en el sentido común y en criterios pragmáticos, pero en el que también hay que respetar ciertas reglas.

Se puede buscar la representatividad de los casos escogidos mediante su selección en base a la clasificación de las unidades según algunas de sus propiedades, de modo que el subgrupo de casos sea «típico» del universo respecto a ellas. Se puede si no, y más fácilmente, renunciar a la representatividad (y con ella a la posibilidad de generalizar al universo los resultados obtenidos) a favor de elegir casos caracterizados por una no representatividad deliberada en relación a fines específicos de la investigación. En especial, es importante subrayar un criterio de elección de excepcional relevancia para la investigación comparada: el de la homogeneidad/heterogeneidad de los casos. Una muestra es heterogénea en la medida en que contiene casos con una alta dispersión de valores/estados para una o más de las propiedades que interesan; por el contrario, es homogénea cuando la dispersión es baja o nula. Se pueden, pues, componer muestras «adecuadas» no representativas del universo, caracterizadas por diversas combinaciones de homogeneidad/heterogeneidad en algunas propiedades.

Un ejemplo puede ilustrar en síntesis las distintas opciones posibles. Si queremos estudiar las causas de la participación electoral en los países democráticos, se puede decidir estudiar una muestra de casos representativos de los distintos niveles de participación que se dan de hecho; también se pueden elegir casos homogéneos en relación al problema, o sea caracterizados por niveles similares de participación; o bien una serie de casos en los que la heterogeneidad sea máxima, es decir, casos extremos de altísima y/o bajísima participación (casos desviantes respecto al universo). Obviamente, respecto a estas diversas estrategias, la investigación de las causas de los distintos niveles de participación electoral se plantea en términos diferentes. En el primer ejemplo se intentará definir propiedades que asumen estados/valores distintos en los casos y cuya variación puede explicar la de los niveles de participación electoral. En el segundo ejemplo buscaremos sobre todo aquellas propiedades comunes a los casos que puedan explicar su homogeneidad en términos de participación electoral. Por último, en el ejemplo de los casos extremos y desviantes se buscarán aquellas propiedades de los casos que resultan tan únicas y excepcionales como para explicar la desviación de los mismos casos respecto a la media del fenómeno.

Esta posibilidad de maximizar o minimizar la varianza mediante la elección de los casos es aplicable tanto al problema en cuestión (niveles de participación

³ ACKOFF, R. L. *The Design of Social Research*, Chicago, University of Chicago Press, 1965.

electoral, en el ejemplo) como a las otras propiedades de los casos que se examinan y que se consideran causas posibles de la primera. Así pues, la definición de la población de los casos obedece a lógicas distintas según la naturaleza del diseño de la investigación y las cuestiones a las que se quiera dar respuesta. Volveremos más en detalle sobre este punto en las secciones que tratan el problema de establecer y controlar relaciones entre las propiedades de los casos.

El mismo problema se plantea en la selección de las dimensiones potencialmente relevantes para la investigación concreta. Para cada caso hay un universo infinito de dimensiones que representan todo lo que se puede predicar respecto a él. Si los casos en examen son, por ejemplo, individuos, las propiedades que se incluyen normalmente en todos los estudios sobre comportamiento y las actitudes de los individuos son, en general, sexo, edad, ocupación, instrucción, etc. Pero para cada problema determinado de investigación pueden tener importancia otras muchas propiedades; por ejemplo, la tenencia o no de un televisor, el hecho de haber participado o no en la Resistencia partisana y así sucesivamente en una serie que es infinita. Otro tanto ocurre si en vez de individuos se estudian agregados sociales o instituciones. De este infinito universo de atributos de cada unidad hay que seleccionar necesariamente un número relativamente reducido de propiedades consideradas especialmente importantes respecto al problema específico de investigación formulado. Para este paso de la investigación no existen criterios precisos de guía. Una primera selección de las posibles variables importantes, es decir, dotadas de un papel significativo respecto al fenómeno objeto de estudio, viene guiada en general por determinadas asunciones heurísticas, por conocimientos heredados o por reconocimientos de investigación.

Trataremos más en detalle estos problemas en los apartados siguientes. Antes hay que prestar la debida atención al importante problema de la formación y tratamiento de los conceptos. De hecho, lo que hasta ahora hemos llamado «casos» y sus «propiedades» son obviamente conceptos a los que recurrimos para indicar más o menos directamente referentes empíricos; es decir, hechos, acontecimientos, estructuras, actividades, comportamientos. El problema de relación entre conceptos y referentes observables es de primordial importancia para la investigación que pretende ser empírica. Afirmar que se intentan estudiar algunos *partidos políticos* (casos) para evaluar la relación existente entre su *fuera organizativa* y su *apoyo electoral* (propiedades) supone establecer una relación entre determinados conceptos y determinados fenómenos reales. Esa relación no deja de tener problemas. Por el contrario, para que se establezca requiere procesos de definición, tratamiento y operativización de los mismos conceptos para transformarlos en contenedores no equívocos de los citados fenómenos reales.

5. Formación de los conceptos

Las exigencias de claridad en la formulación de los problemas plantean una dificultad que subyace en todas las fases de la investigación, la del lenguaje de las ciencias humanas y de la formación de sus conceptos. Los conceptos a que recurren normalmente las ciencias humanas son utilizados, en su mayoría, corrientemente en el lenguaje común. El problema de su claridad y univocidad de significado no puede por tanto darse por descontado, sino que tiene que afrontarse prioritariamente en cada uno de los pasos de la investigación.

Los conceptos de los que nos servimos en la investigación político-social son de dos tipos: empíricos (u observables) y teóricos (o no observables). Los primeros son conceptos que pueden referirse, de una manera más o menos directa, a fenómenos observables. Los segundos, por el contrario, no son reconducibles en modo alguno a referentes observables. Esos conceptos —como, por ejemplo, los de sistema, equilibrio, *feed-back*— tienen el papel de enlazar las teorizaciones y sólo son definibles en referencia a esa función en la teoría que hace uso de ellos.

En el lenguaje de las ciencias sociales son especialmente importantes los conceptos empíricos, los que tienen referentes más o menos directos de su significado en la experiencia sensorial; son los que se pueden referir a cosas observables. Dado que los significados de tales conceptos tienen en nuestra mente la forma de palabras y, a su vez, se refieren a un conjunto de objetos, de referentes, se plantean dos problemas de fondo: 1) si un determinado símbolo (palabra) refleja adecuadamente el significado de un concepto; y 2) si un concepto refleja adecuadamente los objetos o las propiedades de los objetos que intenta reflejar. El primero es el problema de la *ambigüedad* de los conceptos. Disponemos de pocas palabras para muchos significados; así que con frecuencia una palabra tiene varios significados. En el lenguaje común esta parquedad de términos respecto a los significados puede tener incluso aspectos positivos, pero en un lenguaje científico corre el riesgo de generar una notable confusión conceptual. El segundo problema es el de la *vaguedad* del concepto; el de conceptos que no indican con la suficiente claridad su propio referente, las cosas a que se refieren⁶.

Prioritariamente, pues, al comienzo de cada recogida de datos y de informaciones importantes, la investigación científica exige un esfuerzo analítico de formulaciones y definiciones de conceptos con los que indicamos las unidades de investigación y aquellas de entre sus propiedades que nos interesan. Sin este esfuerzo previo cada investigador podría dejarse guiar en la definición de sus conceptos por el sentido común y recoger sus informaciones a partir del nexo que ha establecido entre los conceptos y la realidad. Pero ese procedimiento reduce claramente la intersubjetividad de los resultados alcanzados, debido a la ambigüedad de los términos, y plantea arduos problemas en el transcurso de la investigación.

Para hacer frente al problema de la formación de conceptos empíricos, una primera regla de oro es la de recurrir explícitamente a definiciones verbales (o declarativas) que eliminan la ambigüedad del término, aclarando de entre los distintos significados posibles el seleccionado previamente y que se mantendrá a lo largo de la investigación. En estos casos, incluso para conceptos cuyo significado parezca consolidado, puede ser útil y necesario recurrir a las definiciones lexicográficas que se pueden encontrar en los diccionarios, que certifican el significado de cada término referido exclusivamente al uso que se viene haciendo de él en un contexto dado. Este recurso sirve al menos para establecer si la acepción del significado que queremos utilizar corresponde a la predominantemente usual o si se aleja de ella y necesita entonces una mayor especificación.

Es evidente que cada proceso de definición recurre a otros términos y conceptos, cuya definición lleva a retroceder en un proceso sin fin. La búsqueda de una solución a este problema en la identificación de «términos primitivos» cuyo significado no necesita definición y que por lo tanto se pueden usar en la definición de otros conceptos, no ha dado resultados satisfactorios. Pero el reconocer que existe este problema no debe legitimar el rechazo de todo intento de definición

⁶ SARTORI, G., *La Política. Lógica e método in scienze sociali*. Milán, SugarCo., 1979.

precisa y correcta de los términos, ni desalentar el recurso a procesos de definición cuyos términos tengan un significado claro, por lo menos más claro que el término que queremos definir⁷.

6. Escala de abstracción

De todo lo dicho se desprende que los conceptos utilizados en la investigación política no son entidades absolutamente independientes, sino que cabe concebir su colocación a lo largo de una escala de abstracción sobre la base de la relación que liga para cada uno de ellos su significado, entendido como el conjunto de las características y propiedades que lo constituyen, con su significado, entendido como el conjunto de objetos, fenómenos, acontecimientos a los que se aplica el concepto. El primer significado de un concepto se indica tradicionalmente como su *connotación* (o *intensión*); el segundo, como su *denotación* (o *extensión*).

Según la relación que existe entre la intención y la extensión de su significado, cada concepto se puede colocar a diversos niveles de abstracción. Más en particular, debe existir una relación inversa entre las dos: es decir, que cuanto más alta sea la intención de un concepto, tanto más baja resulta su extensión. En otras palabras, cuanto más rico y específico sea el elenco de las características y propiedades de un concepto, más reducida es la clase de objetos a los que puede aplicarse el concepto. Y al revés, cuanto más reducida sea la especificación de las características y propiedades del concepto, tanto más amplio será el conjunto de objetos a los que puede referirse. Así pues, los conceptos observables pueden usarse en un significado muy abstracto, en referencia a un amplio conjunto de objetos (intención mínima-extensión máxima); o bien de modo muy concreto y específico, en referencia a un conjunto de acontecimientos y hechos u objetos limitado. Se puede incluso llegar a conceptos ideográficos, con una especificación de las características y propiedades del concepto tan amplia y precisa (máxima intención) que el concepto resulta aplicable a un solo caso (extensión mínima)⁸.

Puede ser útil poner un ejemplo. Tomemos un concepto como el «clientelismo». Se puede definir con una intención tan baja —es decir, a través de la identificación de unas pocas de sus características y propiedades— que sea aplicable a una gran cantidad de comportamientos en todos los sistemas políticos del mundo. Por el contrario, podemos aumentar la especificidad y el número de propiedades del concepto de tal manera que sea aplicable y abarque sólo aquellos fenómenos de clientelismo político que se perciben como tales en la cultura política occidental. También podemos aumentar aún más la intención del concepto añadiendo otras propiedades que lo hagan aplicable sólo a los fenómenos específicos de clientelismo de Italia; podemos, por último, llegar a características de especificación descriptiva tales que hagan al concepto aplicable a un único caso concreto de comportamiento clientelar. Este ejemplo sólo tiene carácter indicativo. No siempre el mismo concepto puede utilizarse a lo largo de toda la escala de abstracción y a veces conviene recurrir a diversos términos para indicar distintas configuraciones intensivas de un concepto general.

Los conceptos organizados de este modo vertical corresponden, pues, a distintos niveles de análisis teórico. En toda investigación es preciso establecer a qué

⁷ RIGGS, F., *The Definition of Concepts*, en SARTORI, G.; RIGGS, F., y TEUNE, H., *Tower of Babel*, Pittsburgh, International Study Association, 1975, pp. 39-76.

⁸ SARTORI, G., *La Política ...*, op. cit.

nivel de abstracción queremos colocar nuestro aparato conceptual para evitar los peligros de usar conceptos a un nivel no apropiado con respecto a su definición. La regla que se tiene que respetar es la citada más arriba de la relación inversa entre intensión y extensión. De hecho, ampliar la extensión de un concepto sin reducir en la misma medida su intensión (es decir, sus características y propiedades del significado), o viceversa, provoca confusión. Las consecuencias prácticas de este problema resultarán más claras al tratar los errores de validez de los procesos de operativización de los conceptos. Este problema resulta especialmente grave en los análisis comparados de amplio alcance, en los que se transfieren conceptos típicos de un área cultural o política —y por tanto culturalmente cargados de una cierta intensión de significado— a áreas distintas.

Por último hay que señalar que cuanto más alto sea el nivel de abstracción en el que se coloca un concepto tanto más vasto, y por ello menos definido y preciso, es el conjunto de objetos al que se aplica. Cuanto más bajo sea el nivel de abstracción, pues, tanto mayor será la precisión con que se indican sus referentes, su capacidad empírica. También aquí un ejemplo puede ser útil. Si definimos el concepto de «partido político» como «un grupo organizado que compite por el poder político», esa definición es tan poco intensiva que el universo de referentes empíricos que abarca será muy amplio. Entrarán en él, por supuesto, las organizaciones que consideramos normalmente como partidos políticos, pero junto a ellos muchas otras cosas, como, por ejemplo, organizaciones terroristas y secretas, determinados grupos de presión, y hasta las fuerzas armadas en ciertos casos. Podemos aumentar la intensión del concepto mediante un incremento de sus atributos, por ejemplo añadiendo que la competición por el poder se realiza «presentando candidatos a las elecciones de cargos públicos». Esto reducirá automáticamente su extensión empírica, excluyendo claramente todas las otras organizaciones citadas arriba. Podríamos seguir por este camino añadiendo que las elecciones deben ser de tipo «competitivo», excluyendo, en este caso, a los partidos únicos del campo de referentes empíricos del concepto.

7. Indicadores y definiciones operativas

Las definiciones verbales (o declarativas) no resuelven del todo el problema. Mejor dicho, resuelven el problema de la relación entre palabra/término y significado, pero no el de la relación significado-referente. De hecho, como hemos dicho, los conceptos observables tienen que tener una relación con los referentes empíricos a los que se refieren. Tienen que ser contenedores que especifiquen los objetos a los que se refieren y que abarcan en su significado. En las ciencias sociales, para la mayor parte de los conceptos —incluso para los formulados con más cuidado y claridad— las definiciones declarativas no permiten al investigador juzgar ni ver con precisión lo que se indica empíricamente y lo que no. Muchos de estos conceptos son tan generales y de tan elevado nivel de abstracción que es difícil sacar de sus definiciones semánticas criterios válidos para identificar los referentes empíricos. Para colmar esta distancia entre intensión de un significado conceptual y sus referentes empíricos hay que dar algunos pasos intermedios. Esos pasos son la identificación de *indicadores* de los conceptos y de *definiciones operativas*. La elección de indicadores y definiciones operativas de los conceptos es uno de los momentos más complejos de la investigación.

La definición operativa de un concepto es un tipo de definición que incorpora en su interior la especificación del campo de referentes empíricos del concepto. Está dotada de una serie de instrucciones que indican cómo se propone etiquetar, medir o identificar el concepto dado⁹. Tomemos, a título de ejemplo, el concepto de «disciplina parlamentaria» de un partido, que no se presta por sí solo —por muy clara y unívoca que sea la definición semántica de los términos «disciplina» y «parlamento»— a ser observado directamente. Una posible operativización de tal concepto es la frecuencia con que sus miembros votan en sentido contrario a las decisiones del grupo parlamentario. Esta definición especifica la parte perceptible y medible del fenómeno que el concepto se supone representa. La propiedad «disciplina parlamentaria» de un partido puede ser, pues, evaluada para cada partido haciendo referencia a la definición operativa. Otro ejemplo puede ser el concepto de «cuadro» de partido. De él se puede dar una precisa definición operativa que, en relación con la anterior, identifique el universo empírico de los cuadros de partido declarando, por ejemplo, que por ellos se entiende a todos los que son miembros de los comités ejecutivos a nivel local o identificando de otro modo este grupo (en una reciente investigación sobre partidos políticos italianos ese concepto se definió operativamente referido a los delegados al congreso nacional del partido).

No todos los conceptos son directamente operativizables. Así, el universo de conceptos empíricos puede ubicarse idealmente sobre un continuum que va desde la mínima hasta la máxima posibilidad de operativización directa. Muchos conceptos —y con frecuencia los de mayor importancia para las explicaciones y las teorías de las ciencias humanas— son tan abstractos y generales que no son susceptibles de ninguna operativización empírica directa satisfactoria. En estos casos hay que recurrir a otros conceptos más específicos y menos abstractos, cuyos atributos ofrezcan la posibilidad de operativización directa. Esos conceptos más específicos se llaman *indicadores* del concepto más abstracto y sirven en un cierto sentido de puente entre ellos y las definiciones operativas¹⁰.

En general es necesaria una serie de indicadores para expresar completamente el significado y los referentes de un concepto complejo. Se trata, pues, de identificar las dimensiones o las características relevantes de un concepto y de establecer un indicador para cada una de ellas. La operación implica un proceso de segmentación del concepto complejo y los indicadores escogidos para cada segmento del concepto deberían corresponder en conjunto bastante bien al universo más amplio de características que tenemos en mente cuando utilizamos el concepto originario. Por ejemplo, el concepto de «movilización social» no es susceptible de una satisfactoria definición operativa directa y muchos estudiosos han recurrido a una serie de otros conceptos que se consideran indicadores de la movilización social como el «nivel de instrucción», la «urbanización», la «difusión de los mass-media» y otros. Cada uno de estos indicadores es fácilmente susceptible de definiciones operativas. Otro ejemplo es el concepto de «democracia» (o «régimen democrático»). Mucho mejor que intentar la definición operativa directa de un término tan complejo es recurrir a una definición que incluya los indicadores importantes de tal concepto, como «elecciones libres y competitivas», «libertad de prensa y de asociación», «independencia judicial», etc. Cada uno de estos indicadores se puede

⁹ RIGGS, F., *The Definition...*, op. cit.

¹⁰ PRZEWORSKI, A., y TEUNE, H., *The Logic of Comparative Social Inquiry*, Nueva York, Wiley, 1970; RIGGS, F., *The Definition...*, op. cit.; MARRADI, A., *Concetti e metodi in scienza politica*, Florencia, Giuntina, 1980

definir operativamente mucho más fácilmente o, a su vez, si esto resulta difícil, dar origen a otros indicadores. La lógica de la relación entre los conceptos y sus indicadores empíricos, pues, es que los segundos representan a los primeros y su variación está ligada a éstos.

La relación «concepto-indicador-operativización» constituye uno de los pasos más importantes de una investigación y plantea numerosos problemas. En primer lugar hay que subrayar que los conceptos —y sobre todo los de más elevada abstracción y riqueza de significado— no tienen una relación unívoca con sus indicadores. Por consiguiente, pueden ser operativizados y medidos de más de una manera. Los conceptos que no se pueden observar directamente sólo se podrán abordar mediante indicadores empíricos escogidos, que no son capaces de agotar y traducir completamente el rico significado de conceptos de alto valor teórico. Con frecuencia representan sólo un subconjunto de un conjunto mucho más amplio de posibles indicadores que hubieran podido elegirse para representar un determinado concepto.

Por ejemplo, el concepto de «clase social» se puede definir verbalmente o declarativamente como «un grupo de individuos colocados a lo largo de una jerarquía de posiciones sociales diferenciadas». Esta definición puede operativizarse de diferentes maneras. Implícitamente Marx operativizó ese concepto en referencia a dos indicadores principales, fácilmente operativizables: 1) una posición económica respecto a la propiedad de los medios de producción; 2) la conciencia de tal posición. Sin embargo, como es sabido, otros estudiosos han ofrecido diferentes operativizaciones del mismo concepto, considerando indicadores como la renta, la educación, la ocupación y su prestigio social como más importantes. Por tanto, han operativizado la clase social en términos de estatus socio-económico como una combinación de los mencionados indicadores.

De este ejemplo resulta claro que numerosos conceptos se prestan, según el contexto en que se usan, a ser operativizados de diversos modos y que distintas operativizaciones de un mismo concepto pueden justificarse en función de diferentes objetivos cognoscitivos. La operativización marxista del concepto de clase puede ser ideal en relación con el objetivo de establecer las condiciones de un comportamiento revolucionario de ciertos grupos, mientras que la realizada en términos de estatus socio-económico puede resultar mucho más significativa si quiere usarse el concepto de clase en el estudio del comportamiento electoral. De esto se desprende que la elección entre indicadores posibles de conceptos cuya operativización no es directa, así como la elección entre diversas posibilidades de definiciones operativas de un concepto o indicador dependen también de los fines teóricos para los que se va a emplear el concepto. «La validez de un indicador, en definitiva, depende de una implícita comprensión de la relación entre el proceso mediante el cual el indicador se produce y los objetivos teóricos del investigador»¹¹.

Un concepto operativizado siempre será un concepto empobrecido en su significado, en tanto que está amputado de las características que no resultan operativizables. Este problema está, en definitiva, en la base de la compleja relación entre teoría e investigación. El concepto de movilización social definido en términos no operativos y sin referencia a indicadores empíricos tiene un gran significado teórico y es importante para la explicación y la teoría. El mismo concepto definido operativamente a través de sus indicadores en términos de nivel de urbanización,

¹¹ SMELSER, N. J., *Comparative Methods in Social Sciences*, Englewood Cliffs, N. J., Prentice Hall, 1976.

instrucción, uso de los mass-media, etc., es un concepto muy preciso empíricamente, pero irremediablemente menos fecundo teóricamente.

Es bueno, pues, dedicar la máxima atención analítica a una correcta y explícita formulación de la relación entre conceptos, indicadores y definiciones operativas tratando de evitar aquellos errores capaces de reducir completamente la validez de nuestros indicadores. En primer lugar hay que evitar que el proceso de operativización lleve consigo un empobrecimiento excesivo del significado, mediante una reducción innecesaria del ámbito de referentes indicados por la definición declarativa del concepto. Con respecto al concepto de clase en Marx, un ejemplo muy a propósito —y también con el que nos encontramos a menudo— de este problema sería una operativización del concepto que haga referencia sólo a la dimensión «posición respecto a los medios de producción», descuidando la de la conciencia de la propia posición o viceversa. Una operativización de ese tipo no es válida en tanto que un indicador del concepto, y por tanto una parte de su significado o de sus referentes empíricos, se deja a un lado sin razón (y esto independientemente de los problemas técnicos que la operativización de la «conciencia de clase» puede provocar en una investigación concreta).

Otro caso importante de error de validez del proceso de operativización es aquel en el que el indicador y su operativización identifican un espectro de fenómenos más amplio que el que abarca el concepto y por tanto «mezclan» otras cosas junto a las que se supone que significa el concepto. Un ejemplo a título indicativo puede ser el caso en que se defina «renovación parlamentaria» como el proceso que identifica el surgimiento de nuevo personal en las Cámaras y se operativiza como el número de nuevos elegidos en una elección dada. En este caso, dado que los parlamentarios nuevos elegidos para una legislatura incluyen probablemente a personas que, aunque no elegidas en la legislatura anterior, lo fueron en la precedente, tenemos una operativización que identifica una serie de casos más amplia que aquella que queríamos abarcar con el concepto. Mezclamos al nuevo personal parlamentario —a los electos por primera vez— con otras cosas —los reelegidos—, lo que tendrá un efecto negativo sobre los resultados de la investigación y sobre sus interpretaciones.

A la luz de los simples ejemplos puestos, estos errores de validez pueden parecer banales o de fácil solución. Sin embargo, son frecuentes los procesos de operativización más complejos o menos directos. Pensemos en la dificultad de operativizar de manera válida ciertas actitudes en las investigaciones basadas sobre datos de opinión. Operativizar un concepto como el de «orientación autoritaria» de un individuo en estos casos implica establecer, a través de una serie de indicadores, una relación entre este concepto y un cierto tipo de respuesta a una serie de preguntas preestablecidas. La lógica de la operativización es que será más o menos autoritario el que responda de una determinada manera a las preguntas. En este caso la operativización es muy indirecta e implica referirse a hipótesis y teorías implícitas; las que están en la base del razonamiento en virtud del cual se mantiene que «si una persona tiene una orientación autoritaria, entonces debería estar de acuerdo con... o debería contestar de tal modo...». En el caso, pues, en que los indicadores inherentes a un concepto sean muy indirectos es preciso emplear o asumir una teoría o hipótesis auxiliar para sostener la opción. Las operativizaciones de este tipo son enormemente complejas y dejan continuamente abierto el problema de si los indicadores elegidos no olvidan dimensiones potencialmente operativizables del autoritarismo —o sea, si no serán incompletas—, o bien añaden algo que el autoritarismo no es según la definición que se ha dado de él.

En las investigaciones de opinión es continua la discusión sobre el hecho de si las respuestas a ciertas preguntas pueden interpretarse efectivamente como indicadores válidos de un determinado concepto.

8. El tratamiento de los conceptos

Una definición operativa incorpora una serie de instrucciones que indican cómo se puede etiquetar, medir o identificar el concepto al que se refiere. En otras palabras, la operativización transforma los conceptos de las propiedades de las unidades de investigación en *variables*, especificando la posible variabilidad de los casos a lo largo de una dimensión. Ello permite, pues, asignar valores o etiquetas al estado de cada unidad respecto a la propiedad a la que se refiere. Haciendo referencia a la matriz de datos presentada antes (véase la figura 2.1), los procesos de operativización transforman las propiedades de las unidades objeto de estudio en variables, permitiendo rellenar las casillas de la matriz atribuyendo a cada caso uno de los estados de las variables.

Hasta tal punto es complejo el debate entre los metodólogos sobre este tema, que podemos decir que la operativización indica —según la naturaleza de las propiedades de las unidades y nuestro nivel de información— tres modos de asignar valores o etiquetas al estado de una unidad respecto a una propiedad suya¹².

Para algunas propiedades la definición operativa divide la extensión del concepto en categorías o clases. Esto permite establecer si una propiedad dada existe o no en un caso dado y la operación que a través de ella se cumple asigna las unidades a las diferentes categorías; es decir, clasifica las unidades¹³. Por ejemplo, en un estudio sobre las instituciones de los referéndum populares en varios países (o en varios estatutos de las regiones o estados de un país) podemos querer utilizar la propiedad «efecto sobre la producción legislativa» de un referéndum operativizándola como «efecto propositivo», «derogativo» o «deliberativo». Los casos concretos de consultas referendarias se clasifican por tanto según su naturaleza respecto a esa propiedad. Otro ejemplo es el de la clasificación de unidades-individuos respecto al tipo de profesión religiosa (católica, protestante, judía, musulmana); el de la clasificación de los sistemas políticos respecto al régimen constitucional (parlamentario, presidencial, mixto).

En otros casos las propiedades y su concepto pueden operativizarse suponiendo que están compuestas por categorías ordenadas a lo largo de un *continuum* que va del más al menos y que indica el grado de posesión en cada caso de una determinada propiedad. La operación que cabe hacer en este tipo de operativizaciones consiste en asignar los casos a la categoría que el investigador considera más adecuada, teniendo en cuenta que las diferencias medidas así son de grado y no de cantidad. Los intervalos entre las categorías ordenadas a lo largo de un *continuum* son desiguales y su tamaño es indeterminado. Ejemplos de este tipo de operaciones son aquellos en que se pregunta a los entrevistados su opinión no en términos de un categórico «sí» o «no», sino más bien en términos del grado de acuerdo o

¹² TEUNE, H., *Measurement in Comparative Research*, en «Comparative Political Studies», 1, 1968, pp. 123-138; MARRADI, A., *Concetti*, op. cit.

¹³ Como ha señalado acertadamente Marradi (1980), con el término «clasificación» se entiende tanto la operación intelectual de categorización de un concepto como el procedimiento de atribución de los objetos a una de las categorías establecidas a partir de aquella. En esta obra no hemos considerado necesario distinguir terminológicamente los dos significados, pero sí hemos tenido en cuenta la diferencia.

desacuerdo ofreciéndoles, por ejemplo, una serie de respuestas ordenadas alternativas. En el estudio de macroagregados tenemos este tipo de medida cuando, por ejemplo, los indicadores permiten colocar a varios sistemas políticos en una serie de categorías ordenadas en relación a la mayor o menor centralización administrativa. Un ejemplo más corriente puede ser el que, en base a un examen, se asignan estudiantes a clases de estudio de lengua para «principiantes», «intermedios», «avanzados», asumiendo que éstas representan categorías sobre un *continuum* que va desde el desconocimiento absoluto de una lengua hasta su conocimiento avanzado.

Por último, el tercer modo de asignar valores a los estados de las variables a través de la operativización permite establecer la *cantidad* de la propiedad que posee cada caso; el grado numérico en que están presentes las propiedades. Si en una investigación tenemos que estudiar la propiedad «fuerza organizativa» de un partido político y damos a este concepto una definición operativa que hace referencia al número de afiliados y/o al número de las unidades locales en el territorio, el estado de cada caso respecto a las variables se identifica por cantidades numéricas precisas a las que se pueden aplicar todas las propiedades matemáticas de los números. En este caso, en comparación al anterior, la dimensión de los intervalos entre los posibles estados de las variables es conocido: si el número de los afiliados de un partido es doble que el de otro se puede concluir que —en relación a la definición operativa de «fuerza organizativa» adoptada— la fuerza organizativa del primero es dos veces la del segundo. Dicho de otro modo, la igualdad de los intervalos permite comparaciones entre caso y caso que se refieren al tamaño de las diferencias.

Los tres modos fundamentales de asignar valores y etiquetas a los estados de las propiedades según su definición operativa se indican normalmente con los términos de medida *nominal*, *ordinal* y *cuantitativa*, dando lugar a tres tipos de escalas. Sobre esta terminología se centra el debate entre metodólogos respecto a las propiedades de los tres tipos de tratamiento y a las técnicas estadístico-matemáticas que cada uno de ellos permite adoptar legítimamente¹⁴. A estos problemas se refieren ampliamente los textos especializados de metodología. Aquí importa subrayar que, desde el punto de vista del tratamiento lógico de los conceptos, existe una clara cesura entre el tratamiento conceptual discontinuo del primer tipo, que da lugar a clasificaciones de casos, y el continuo de los tipos segundo y tercero que, en distintos grados y con distintas consecuencias sobre las técnicas estadísticas utilizables según que se trate de escalas ordinales a intervalos o cardinales, da lugar a distintos tipos de medida de los conceptos operativizados.

9. Clasificaciones

Conviene subrayar que la lógica clasificatoria —es decir, la operación intelectual mediante la cual se divide la intención de un concepto en un cierto número de categorías— y las clasificaciones tienen una importancia fundamental en las ciencias sociales (y en otras ciencias mucho más consolidadas) y merecen una atención adicional¹⁵. En primer lugar, las buenas clasificaciones no son tan fáciles de formular como puede parecer a primera vista observando los ejemplos que

¹⁴ MARRADI, A., *Concetti*..., op. cit.

¹⁵ SARTORI, G., *La Política*..., op. cit.

hemos puesto antes. El tratamiento clasificatorio de los conceptos tiene reglas que hay que seguir. Las fundamentales son: 1) que se base sobre un único criterio explícito de distinción, y 2) que la clasificación sea exhaustiva y exclusiva¹⁶. La exhaustividad de las categorías de una clasificación implica que cada unidad debe atribuirse a una clase. La exclusividad exige que ninguna unidad se atribuya a más de una clase.

Respetar estos criterios no es fácil. En muchos casos la tarea se facilita porque recurrimos a clasificaciones de sentido común (como puede ser la distinción entre sexo masculino y femenino) o a otras ya existentes de forma estandarizada (por ejemplo, las ocupaciones definidas por los Institutos de Estadística) o basadas en informaciones *selfreported* (como la religión o la nacionalidad). Por otra parte, sin embargo, y sobre todo en los casos de mayor valor e interés heurístico, tenemos que crear nuestras propias clasificaciones sobre la base de una serie de datos e informaciones desagregados. Pensemos en el intento de clasificar según el principio de las motivaciones predominantes las variadas respuestas de los entrevistados a una pregunta referente a las razones por las que han cambiado su comportamiento electoral. La tarea de identificar una serie reducida de categorías nominales que puedan resumir y contener la gran variedad de respuestas individuales es un proceso complejo.

Un mismo concepto se puede operativizar en más de uno de los tres modos principales señalados e incluso en los tres. Tomemos una vez más un ejemplo creado *ad hoc*. Si queremos evaluar la propiedad «centralización administrativa» de varios sistemas políticos, se puede tratar nominalmente tal concepto proponiendo una clasificación dicotómica —centralizado-descentralizado— basada en la presencia o ausencia de ciertas características distintivas de la centralización (y sus indicadores). Podemos también tratar de medir el grado de centralización ordenando los sistemas políticos en un *continuum* que valore comparativamente la mayor o menor presencia de las características. Por último, podemos tratar de cuantificar la centralización midiendo cada característica individualmente y ajustándola en un *índice* de centralización, es decir, en una medida que combina los indicadores según una operación matemática¹⁷.

Naturalmente que estas tres posibilidades tienen sus ventajas e inconvenientes. Un tratamiento nominal no hace justicia a los casos intermedios, forzándolos a entrar en una de las categorías. Además, los tratamientos continuos permiten recurrir a técnicas estadísticas más poderosas para el análisis. Sin embargo, el tratamiento nominal-clasificatorio permite un mayor control mental del concepto cuando se pone en relación con otros. Es indicativo, de hecho, que, a veces, conceptos fácilmente operativizables de modo cuantitativo sean reconducidos, al final del análisis, a clasificaciones o escalas ordinales mediante la identificación de puntos de ruptura en el *continuum*.

Todo lo dicho acerca de la relación entre tratamiento clasificatorio y continuo de los conceptos operativizados lleva también a concluir que con mucha frecuencia la relación que existe entre tratamiento nominal-clasificatorio y tratamiento continuo no es sólo una relación alternativa impuesta por el nivel de las informaciones disponibles o por un cálculo de conveniencia basado en los costes y resultados de las diferentes posibilidades. En muchos casos el tratamiento clasificatorio

¹⁶ MILL, J. S., *A System of Logic. Ratiocinative and Inductive*, Londres, 1843.

¹⁷ LAZARSFELD, P. F., y BARTON, A. H., *Qualitative Measurement in the Social Sciences: Classification, Typologies and Indices*, en D. LERNER y H. D. LASSWELL (eds.), *The Policy Sciences*, Stanford, Stanford University Press, 1951; MARRADI, A., *Concetti ...*, op. cit.

se presenta como un requisito lógico de aquel continuo, sobre la base de la consideración de que la lógica del «más o menos» o del cuanto de una propiedad es posterior a la de la existencia o no de la propiedad¹⁶.

Este hecho es evidente en las fases iniciales de la investigación. Las unidades de investigación son mezclables o comparables entre sí sólo si por sus características definitorias pertenecen a una misma clase de objetos, fenómenos o acontecimientos. En este sentido, es, pues, evidente el carácter lógicamente anterior del análisis clasificatorio respecto a cualquier otro paso de la investigación. Todo lo dicho vale también para el nivel del tratamiento conceptual de las propiedades de las unidades: la explicación de los conceptos que permite la lógica clasificatoria pone en evidencia las divisiones en subgrupos de las unidades de investigación y permite identificar a cuáles de esas se pueden aplicar operativizaciones continuas.

Vamos a poner también en este caso un ejemplo que aclare. Consideremos un estudio de las instituciones parlamentarias que prevea el análisis de una propiedad como la del «poder legislativo de las comisiones». Naturalmente, en términos sustanciales tal poder varía en cada caso y no es difícil pensar en una operativización continua de ese concepto que haga referencia a uno o más índices (por ejemplo, el porcentaje de leyes aprobadas en comisión respecto al total de leyes aprobadas; el número de leyes que se atribuyeron a las comisiones en la fase deliberante pero se llevaron al pleno por diversos motivos, etc.). Este posible índice del poder legislativo de las comisiones parlamentarias tendrá su ámbito de variación partiendo de 0, valor que representa el caso en que tal poder es nulo, el caso en que las comisiones parlamentarias no tienen institucionalmente poder de producir leyes.

El problema se plantea en los siguientes términos: ¿es lícito reducir una diferencia institucional-estructural entre los diferentes parlamentos (atribución o no del poder legislativo a las comisiones) a una diferencia de cantidad o grado en su poder? La respuesta es que, en un ejemplo de este tipo, la reducción de la diferencia entre los casos a diferencia cuantitativa implica ocultación. Sería más apropiado un tratamiento clasificatorio previo de los poderes de las comisiones que distinga la presencia o ausencia del poder deliberativo en cada caso individual. Sólo después es oportuno y útil aplicar a la clase de parlamentos que prevén comisiones con poderes legislativos una lógica cuantitativa que lleve a evaluar las diferencias de grado y cantidad en que poseen tal poder.

En casos como este es necesario realizar previamente un ejercicio conceptual clasificatorio. Ello no supone la pérdida de ninguna información importante, sino que, al contrario, favorece una clarificación conceptual que la medida asumirá necesariamente como algo implícito (si esta diferencia no estuviese asumida al menos implícitamente en la medida, entonces deberíamos concluir, paradójicamente, que el tratamiento cuantitativo determina una pérdida de información). En conclusión, el tratamiento nominal y clasificatorio es una forma de operativización alternativa respecto a las formas continuas, pero en muchos casos es un tratamiento lógico necesario y prioritario respecto a los otros.

10. Tipologías

Un caso concreto de aplicación de la lógica clasificatoria especialmente importante en ciencia política y en las ciencias sociales en general es el de la formación

¹⁶ SARTORI, G., *La Política...*, op. cit.

de tipos y tipologías. Una tipología se suele definir como una «clasificación multidimensional», entendiéndose con ello que se trata de una organización de tipo clasificatorio establecida por la combinación y cruce de más de un criterio. Los tipos que vienen a componer una tipología constituyen una construcción caracterizada por una serie de atributos, cada uno de los cuales identifica una dimensión que puede ser representada por un tratamiento nominal o continuo, pero cada uno de los cuales entra en el tipo con su valor específico¹⁹.

Consideremos un ejemplo concreto de tipo que por su aplicación al caso italiano y por las discusiones que ha provocado debería ser conocido por el lector: el tipo de sistema de partidos de *pluralismo polarizado* (véase el capítulo sobre sistemas de partidos). Este tipo surge del ámbito de una tipología general de los sistemas de partidos construida combinando dos de sus dimensiones o prioridades importantes: 1) el número de partidos; y 2) la distancia ideológica que les separa. El tipo «pluralismo polarizado» surge como combinación específica de un valor/estado de cada una de las dos dimensiones-propiedades. Se caracteriza, pues, por la combinación de un *elevado* número de partidos con una *gran* distancia ideológica entre ellos. Por supuesto que distintas combinaciones de valores/estados de las dos dimensiones-propiedades constituyen otros tantos tipos posibles de sistema de partidos, cada uno de los cuales es concebible en abstracto (pocos partidos y corta distancia ideológica; pocos partidos y gran distancia ideológica, y así sucesivamente).

Este ejemplo de las modalidades constitutivas de los tipos y de las tipologías permite poner de manifiesto y discutir algunos de sus problemas. En primer lugar, las diversas dimensiones-propiedades que se combinan en las construcciones tipológicas pueden ser originariamente operativizadas tanto nominalmente como clasificaciones o como escalas de distinto tipo. Es evidente, sin embargo, que en presencia de numerosos casos los valores de una variable operativizada de modo continuo son tantos, o mejor dicho demasiados, como para ser utilizados todos, uno por uno, en una construcción tipológica. Estamos frente a uno de los casos en que un tratamiento continuo presenta inconvenientes en términos de coste y se reduce, mediante la identificación de uno o más umbrales, a un tratamiento disyuntivo, como mucho dicotómico o tricotómico. En términos concretos, para utilizar con fines de construcción de una tipología una dimensión como el nivel de desarrollo económico, puede ser más conveniente dividir los países en desarrollados y subdesarrollados identificando un umbral en los valores de los diferentes índices.

Incluso de esta manera, el número de tipos posibles que surgen de la combinación suele ser muy elevado. Para darnos cuenta basta calcular que, si queremos construir una tipología combinando cuatro dimensiones-propiedades operativizadas en términos dicotómicos, obtenemos dieciséis tipos posibles. Un número tan alto de tipos posibles hace difícil la interpretación y el mismo control mental de la realidad objeto de estudio. Por lo tanto, siempre es necesario llevar a cabo una reducción de las combinaciones posibles entre los distintos estados/valores de las dimensiones, operación que, en general, se realiza de tres maneras fundamentalmente.

La primera tiene carácter lógico, en el sentido de que una relación existente entre dos o más dimensiones de la tipología es lo que puede permitir deducir que ciertas combinaciones de sus valores no pueden lógicamente verificarse. Otras re-

¹⁹ LAZARSFELD, P. F., y BARTON, A. H., *Qualitative* ., op. cit.

ducciones se obtienen eliminando aquellas combinaciones de estados/valores que, aunque sean concebibles lógicamente, no se dan empíricamente o se dan con tan poca frecuencia que no hace falta establecer un tipo especial. Por último, ciertos grupos de combinaciones estados/valores se pueden reagrupar juntos en un solo tipo, en base a consideraciones pragmáticas ligadas a los objetivos específicos de la investigación²⁰. Las reducciones del segundo tipo, las realizadas de hecho en base a consideraciones de frecuencia empírica, son ciertamente legítimas y útiles. Hay que hacerlas, sin embargo, con cautela, en cuanto que tipos con baja o bajísima frecuencia empírica pueden ayudar a identificar los casos desviados, sobre cuya importancia volveremos al final del capítulo.

Las construcciones tipológicas se basan completamente sobre principios de la lógica clasificatoria y para ellos son válidos los criterios de exhaustividad y exclusividad ya mencionados. Además hay que subrayar que la reconstrucción de los procesos constitutivos de las tipologías (identificación de las dimensiones, reducción de los estados y reducción de las combinaciones) es un ejercicio lógico que presenta numerosas ventajas. De hecho, en la literatura se encuentran muchos más tipos que tipologías completas. Lo que es preocupante, en cuanto sugiere que algunos de ellos han sido creados a propósito, sin haber explicitado completamente la construcción tipológica de la que se desprenden y sin explicar, por tanto, de qué tipo se trata (tipos empíricos, populares, ideales). En estos casos, aplicar *a posteriori* la lógica de constitución al tipo permite explicitar completamente la tipología y poder, por tanto, identificar y evaluar las reducciones de las combinaciones que se han realizado, la posible existencia de superposiciones de casos (no exclusividad) y de casos desviantes.

En este apartado hemos dedicado mucho más espacio a las clasificaciones y tipologías que a los niveles de medida representados por las escalas. Esta opción se deriva del hecho que los textos de estadística y metodología dedican mucho espacio a estos últimos y de la convicción de que, en el estado actual, en la mayor parte de los campos de las ciencias sociales todavía es necesario saber clasificar y tipologizar bien más que medir y cuantificar. Naturalmente, saber operar bien con los primeros no debe significar abandonar los segundos; pero recurrir a los segundos no permite abandonar los primeros²¹. Las clasificaciones y las tipologías ordenan el universo de los objetos y los acontecimientos políticos; ayudan a la formación de los conceptos a medir y constituyen objetivos científicos en sí, cuya necesidad se siente. Además, son la base del método de control al que la investigación política recurre con mayor frecuencia: el método comparado.

11. Relaciones entre variables

Conviene recapitular los problemas discutidos hasta ahora y sus relaciones con la ayuda de un esquema (fig. 2.3). Hemos partido de la definición del problema y de la identificación de las unidades y de sus propiedades importantes para el estudio; hemos discutido los problemas inherentes a la definición conceptual y operativa de éstos a través de los procesos de indicación. El tratamiento operativo de los conceptos los transforma en variables operativas susceptibles de asumir estados diferentes en cada caso; estados que se definen por tratamientos

²⁰ LAZARSFELD, P. F. y BARTON, A. H., *Qualitative...*, op. cit.

²¹ SARTORI, G., *La Política*, op. cit.

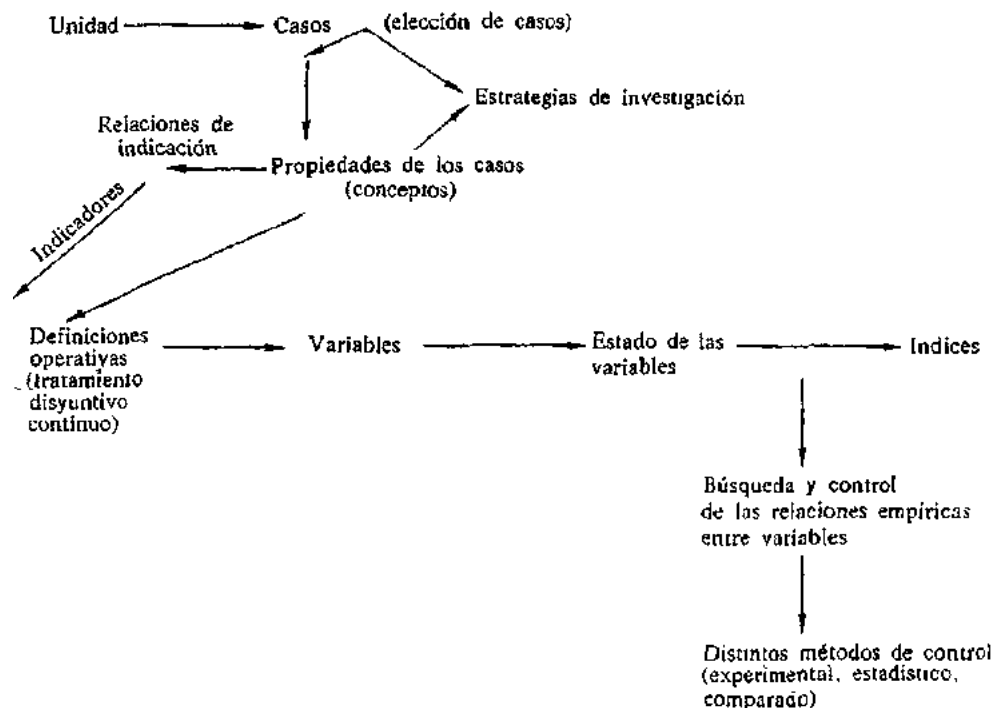


FIG. 2.3.—Esquema resumido de la exposición.

discontinuos en base a procesos de formación de clasificaciones y tipologías y por continuos en base a varios niveles de medida.

En este punto, la matriz inicial se compone idealmente de casos que hemos seleccionado de entre las unidades de investigación, de las variables y de sus estados. Ahora tenemos que pasar a discutir los procesos de investigación, de elaboración de hipótesis y confirmación de las relaciones empíricas entre las variables. En las ciencias sociales los tipos de relaciones que se buscan son, en general, causales. El objetivo principal es el de explicar en el sentido de atribuir una o más causas a un fenómeno, de identificar sus condiciones causales. Normalmente, se identifican con el nombre de *variables dependientes* los fenómenos que se intentan explicar y con el de *variables independientes* las condiciones causales que se supone los determinan²².

Sin embargo, incluso basándonos sólo en un juicio de sentido común, para cada fenómeno que se quiera explicar en el sentido señalado más arriba el número de las condiciones causales que pueden influir sobre él es extremadamente alto. Con palabras de un conocido estudioso, la situación inicial «es una de *multiplicidad* de condiciones (causales), de *confusión* de su influencia sobre lo que se tiene que explicar y de *indeterminación* respecto a los efectos de cada condición en combinación»²³. Lo que supone que en todo estudio científico la búsqueda de explicaciones causales, incluso de naturaleza probabilística, implica un esfuerzo

²² BROWN, R., *Explanation in Social Science*, Chicago, Aldine, 1963.

²³ SMELSER, N. J., *Comparative Methods...*, op. cit.

de reducción del número de las posibles condiciones causales y de aislamiento de una condición de otra, de tal modo que quede claro el papel de cada una de ellas.

En la práctica, la reducción de las posibles condiciones causales de un fenómeno se hace *in primis* mediante *procesos heurísticos* que simplifican el universo de variables potencialmente importantes. Se asume que ciertos conjuntos de variables no son influyentes; tanto si son conocidas como si son desconocidas, se les supone constantes y, por tanto, como no influyentes respecto al problema concreto. A esa operación se le da el nombre de recurso a la cláusula de *ceteris paribus*²⁴. De este modo se acepta implícitamente una estrategia de división del trabajo científico en la que aquello que es problema para un cierto tipo de estudio se considera como un dato que no influye en otros y viceversa. Cada enfoque se basa necesariamente en la congelación de un cierto número de variables en parámetros con el fin de disponer de un esquema interpretativo controlable mentalmente. Así, por ejemplo, los estudios sobre la relación entre valores, cultura y comportamiento recurren en general a procesos paramétricos sobre el contexto institucional en que se manifiesta este impacto; los que estudian la influencia de las estructuras sociales sobre el comportamiento reducen generalmente a parámetros los procesos psicológicos mediante los cuales se produce este impacto; los modelos de los economistas consideran la mayoría de las veces como constantes sin influencia las variables de naturaleza política y social.

El campo de las variables causales potencialmente operativas, aún después de reducirlo tan drásticamente, sigue siendo siempre bastante amplio en todo tipo de investigación. Se impone una organización de esta amplitud mediante la formulación de hipótesis precisas, que son, ni más ni menos, que juicios sobre la relación causal existente entre un número reducido de variables independientes y la o las variables dependientes. Alguna hipótesis está ya presente, de forma más o menos explícita, desde el comienzo de la investigación, y va siendo posteriormente afinada o modificada; otras pueden surgir *ex novo* durante la realización del estudio, de la intuición creativa del estudioso o de la bibliografía existente. El proceso de formulación de las hipótesis es, en definitiva, muy idiosincrásico y escapa a cualquier intento de sistematización. No hay tesón ni proceso lógico que pueda por sí sustituir la creatividad del investigador a este respecto. Cuando se sostiene que ciertos instrumentos estadísticos permiten «descubrir» hipótesis que se someten después a control se olvida subrayar que cada técnica debe, en cualquier caso, aplicarse a un conjunto relativamente reducido de variables, identificadas necesariamente a partir de alguna hipótesis implícita sobre su importancia.

12. Los métodos de control

Muy distinto es en cambio el problema del control de una hipótesis de relación entre una o más variables independientes y la variable dependiente. El proceso de comprobación²⁵ de las hipótesis puede y debe someterse a algunas reglas meto-

²⁴ URBANI, G., *Introduzione*, en G. URBANI (ed.), *La politica comparata*, Bologna, Il Mulino, 1973, pp. 7-39.

²⁵ La posición epistemológica dominante actualmente mantiene que en la investigación científica no es apropiado hablar de «verificación» de las hipótesis o teorías, en cuanto que el proceso cognoscitivo procede sobre todo mediante la «falsificación» de éstas. Esta tesis se aclarará al final del capítulo. A pesar de lo fundado de la crítica, en este texto se recurre al término de «verificar» y «controlar» en sentido general, para referirnos tanto a la verificación en sentido estricto como a la falsificación de las proposiciones. Y ello para no recurrir al tér-

dológicas. Respecto a cualquier hipótesis de relación causal entre varias variables se presentan, de hecho, constantemente dos problemas: 1) cómo aislar una condición causal de la otra de tal manera que se conozca el papel causal independiente de cada variable; y 2) cómo controlar si otros tipos de variables no tenidas en cuenta directamente en la hipótesis influyen en la relación que se estudia y cómo lo hacen. Este proceso, que permite ganar control empírico sobre posibles fuentes de variación de las variables independientes, se basa, en cada tipo de investigación científica, en la transformación en parámetros —o sea, en constantes que no influyen en la relación— de algunas de las variables operativas que se sospecha influyen en las variables dependientes. Mediante la transformación de variables operativas en constantes se actúa de tal modo que la mayor parte de las posibles condiciones causales no varíen, de manera que la influencia de una o algunas condiciones —a las que en cambio se deja variar— pueda aislarse y analizarse. Repitiendo esta operación de transformación de variables en constantes de modo sistemático para todas las variables que se sospecha puedan influir en la variable dependiente, la explicación se hace más afinada y generalizada²⁶.

Esta lógica de la parametrización de las variables independientes se encuentra en la base de toda investigación científica de relaciones causales. Los distintos métodos de investigación y control científico se distinguen respecto a esta lógica en el *modo* y en la *medida* en que las distintas condiciones causales son convertidas en parámetros.

Los métodos de control y comprobación de las hipótesis de relaciones causales a los que se recurre en la investigación científica son tres: el *método experimental*, el *método estadístico*, el *método comparado*. Algunos autores consideran también un cuarto método, el *método histórico*, mientras que otros excluyen de este elenco al método comparado. Trataremos estos problemas más adelante justificando la distinción de los tres métodos. Aquí interesa subrayar que los tres métodos se diferencian por la manera como resuelven el problema central de toda investigación causal: la transformación en parámetros de las variables operativas externas (llamadas a menudo *variables intervinientes*) e internas a la relación causal planteada como hipótesis, y que esta diferencia tiene repercusiones profundas en la «potencia» de cada método.

13. El control experimental

El método experimental se considera generalmente como el método más potente para establecer y controlar relaciones de causalidad entre variables, de tal modo que se suele considerar como equivalente al método científico. En el experimento, se introducen deliberadamente modificaciones de los estados/valores de las variables independientes e intervinientes para valorar su efecto sobre los estados/valores de las variables dependientes. La característica distintiva del método experimental es, pues, la posibilidad de manipular los estados de las variables directamente controlando su creación (es el investigador el que decide y determina qué varía y cuánto). Esto permite la máxima posibilidad de parametrizar todas las variables de las que se sospecha que influyen en la variación

mino «testar», anglicismo decididamente horrible que se está difundiendo en los textos de metodología.

²⁶ SMELSER, N. J., *Comparative Methods*, op. cit.

de la variable dependiente y, por tanto, aislarlas y evaluar sus influencias individuales. En otras palabras, refiriéndonos a nuestra matriz de casos y variables, *el método experimental basa su potencia en la posibilidad de cambiar el estado de una variable sin hacer variar también todos los estados de las otras variables*. De hecho, pues, la matriz de los datos la crea el investigador según sus exigencias cognoscitivas.

La parametrización mediante el método experimental permite, pues, maximizar los factores requeridos para efectuar inferencias causales correctas: 1) la *precedencia temporal* de las variables independientes respecto a las variables dependientes (en cuanto que la introducción del estímulo precede al efecto observado); 2) la *constante asociación* entre las dos variables (gracias a la posibilidad de repetir el experimento en diferentes condiciones); 3) la *validez de la relación establecida*, o sea el hecho de que sea realmente el producto de la variación de la variable independiente y no de otras variables externas (en cuanto es posible excluir la influencia de todas las demás variables, manteniéndolas constantes)²⁷.

En las ciencias sociales, el método de control experimental se utiliza sólo en campos bastante reducidos, en el estudio de microfenómenos y aplicado a grupos pequeños. Es imposible modificar con estímulos externos la naturaleza organizativa de un sindicato o el régimen constitucional de un país. En cambio, en el ámbito de problemáticas más reducidas sí es posible recurrir a diseños de investigación experimentales. En ellos nos acercamos a la lógica experimental confrontando los resultados de dos grupos de individuos, uno de los cuales se somete a estímulo (por ejemplo, la modificación del estado de una o varias variables independientes) y el otro no (grupo de control). Las diferencias halladas en las variables dependientes se atribuyen al efecto del estímulo²⁸.

Para poner un ejemplo concreto, pensemos en el caso en que se quiera controlar la relación que liga el nivel de participación electoral (variable dependiente) con determinados esfuerzos de propaganda política de los partidos, la propaganda casa por casa de los militantes o el envío de material propagandístico (variables independientes). Tras haber seleccionado según las reglas de muestreo estadístico dos grupos de electores, se somete a uno de ellos al estímulo de la propaganda militante. Las eventuales diferencias en el nivel de participación electoral de los dos grupos se atribuyen al impacto de los estímulos introducidos. Este ejemplo está muy simplificado. El diseño de investigación experimental puede alcanzar niveles de alta sofisticación, añadiendo otros grupos de comparación (sometidos al experimento) o de control (no sometidos al experimento), seleccionados con criterios de muestreo casual o no, observados antes y/o después del experimento. Estas diferentes soluciones pueden combinarse en el intento de maximizar la validez del experimento y la fiabilidad de las conclusiones obtenidas.

Aparte de los complejos problemas técnicos que implica esta estrategia, el objetivo es el de obtener el mayor grado posible de control sobre las fuentes de variación de las variables dependientes mediante la parametrización progresiva del mayor número posible de variables independientes. Sin embargo, el diseño de investigación experimental en las ciencias sociales no es equiparable a la experimentación en las ciencias naturales. Por más esfuerzos que hagamos, en el aná-

²⁷ BLALOCK, H. M., jr., *Causal Inferences in Nonexperimental Research* (1961), Nueva York, Norton Library, 1972.

²⁸ CAMPBELL, D. T., y STANLEY, J., *Experimental and Quasi-Experimental Research*, Chicago, Rand McNally, 1963; trad. castellana, *Diseños experimentales y cuasi-experimentales en la investigación*, Buenos Aires, Amorrortu, 1979.

lisis de los fenómenos socio-políticos no se consigue un control directo sobre todas las condiciones causales posibles.

14. El control estadístico

En la gran mayoría de los problemas a que se enfrentan las ciencias humanas no es posible manipular la realidad o introducir en ella estímulos externos. Para establecer relaciones entre fenómenos se recurre a métodos de estudio de su *covariación* en el conjunto de casos que se examinan. Poniendo en relación más variables en los distintos casos se trata de establecer cómo y cuánto covarian sus estados y si de la variación combinada que resulta se puede inducir una cierta relación entre las variables consideradas²⁹.

Naturalmente, en este método la variación de los estados/valores de una variable se obtiene a través de la toma en consideración de *casos diversos*. Esto implica que al cambio de las variables examinadas corresponde también un cambio, caso por caso, de todas las otras variables presentes en la matriz de datos construida en nuestra investigación (y, además, varían también las propiedades de los casos no incluidos en nuestra investigación). Así pues, frente a la caracterización de una relación causal debemos considerar que ella es sólo una dentro de una red de relaciones determinadas por el variar de todas las variables de caso a caso y que podrían existir otras relaciones aún más importantes que la descubierta y capaces de determinarla enteramente.

Esta es la situación «normal» a que debe enfrentarse el investigador. Si se dispone de un gran número de casos, el método estadístico permite la aplicación de técnicas matemáticas al universo o a una muestra de él para obtener los mismos efectos de transformación de variables operativas en parámetros que ya hemos citado como esencial esfuerzo cognitivo. Un amplio repertorio de técnicas estadísticas —algunas bastante complejas y sofisticadas— permite establecer relaciones de covariación entre variables y descontar y controlar el impacto sobre esa relación de otras variables influyentes, pero de menor importancia. Las técnicas estadísticas permiten, además, afrontar con mayor seguridad el problema de avanzar propuestas relativas a un fenómeno, a partir del análisis de una pequeña parte de las unidades del universo (los casos escogidos), mediante una serie de procedimientos capaces de potenciar la inducción de la muestra al universo, utilizando la teoría de las probabilidades para evaluar el error de estimación de esta inducción.

El método estadístico es, por tanto, un método muy potente de descubrimiento y control de las relaciones entre variables, aunque menos potente que el experimental. Dado que los estados de las variables no se manipulan en una dirección querida ni se mantienen constantes, no es posible establecer la precedencia temporal de una variable respecto a las otras. Por consiguiente, las asociaciones que se puedan establecer no permiten hacer afirmaciones seguras acerca de la dirección causal de la relación que se ha descubierto. Establecida una cierta relación entre X e Y, averiguar si el primero es causa del segundo o viceversa es una cuestión de interpretación teórica, efectuada a menudo en base a «otros conocimientos» que posee el investigador³⁰. Además, incluso en el método estadístico, la posibilidad de controlar la influencia potencial de los cambios de las variables externas es limitada. No es posible técnicamente, ni ventajoso heurísticamente, incluir en un modelo todas las variables que se supone que puedan ser influyentes.

²⁹ BLALOCK, H. M., jr., *Social Statistics*, Nueva York, McGraw-Hill, 1960.

³⁰ MARRADI, A., *Concetti*, op. cit.

15. El control comparado

Cuando el número de casos del estudio es insuficiente para dar veracidad al método y/o la operativización de las variables se presta mal a la utilización de técnicas estadísticas es preciso entonces recurrir al método comparado de control de las hipótesis: un método claramente menos potente que los anteriores y a la vez más complejo.

Algunos niegan la existencia de un método autónomo de control comparado en general, en base al argumento de que también hay comparación en el método experimental y en el estadístico. Ese argumento en esa forma es correcto. Mejor, se puede llevar más lejos hasta sostener que la actividad mental de comparación es inherente al pensamiento y como tal subyace en toda forma de conocimiento y juicio, incluso las que la vida presenta cotidianamente (desde la compra de un televisor a la elección de la cola más corta en las ventanillas de una oficina de correos). En este sentido, el ejercicio de la comparación es un proceso mental constitutivo e implícito del pensamiento y como tal no se puede considerar un método de control científico, independientemente de su fuerza o debilidad en cuanto tal.

Sin embargo, en lo que se refiere específicamente a la investigación política, a esta argumentación se pueden presentar dos objeciones de distinta naturaleza e importancia. La primera es que en esta acepción general la comparación es un proceso ampliamente implícito y la explicitación de los criterios de los métodos de comparación es esencial para el que quiere controlar sus generalizaciones de modo transparente intersubjetivamente. La metodología comparada se impone la tarea mínima de sustraer la actividad de comparación del reino de lo implícito y de lo no sistemático, para hacerla explícita y sistemática. Este proceso identifica sus puntos de fuerza y de debilidad, y por lo tanto las tareas y los casos de aplicación y validez³¹.

La segunda objeción —de mayor importancia— se refiere a que en los problemas de control de las ciencias sociales en el tema mencionado, se confunde la o las comparaciones con el método de control comparado. Se confunde la actividad genérica de comparar (que puede tener diversas finalidades) con un proceso cognoscitivo específico obtenido a través de la comparación: la parametrización de las variables externas a la relación que se considera como hipótesis. El problema central de todo intento de verificación y control de una relación causal entre dos variables consiste en anular y controlar la influencia de otras variables externas potenciales; y este problema sólo se puede resolver mediante una parametrización que las haga no influyentes. En relación a este aspecto central, el método experimental y el estadístico ofrecen diversas soluciones. Lo mismo hace el método comparado: el mismo objetivo de fondo, la parametrización, y, por tanto, el control de las variables externas, se afronta en el método comparado de manera diferente a los otros y en ello basa su autonomía como método de control.

Cuando no es posible manipular los estados de las variables externas ni controlar con técnicas estadísticas la influencia de sus cambios, sigue siendo posible reagrupar en clases o tipos los casos que se examinan en base a los estados de las variables externas, a fin de controlar la relación entre variables independientes y dependientes en cada clase separada caracterizada por la parametrización de

³¹ LIPHART, A., *Il metodo della comparazione*, en «Rivista Italiana di scienza politica», t. 1971, pp. 66-92; URBANI, G., *Introduzione* ., op. cit.; SARTORI, G., *La Politica...*, op. cit.

una o más variables externas potencialmente no influyentes. Dicho de otra manera, el proceso de clasificación asume el papel fundamental del proceso de parametrización en cuanto los casos que pertenecen a una misma clase son, en efecto, asimilados precisamente en base a la *constancia* de la propiedad que constituye el criterio de su agrupamiento. La lógica común a todo método de control —la parametrización de las variables— se manifiesta en el método comparado a través de procesos de clasificación y reclasificación en grupos de los casos, según las variables independientes o intervinientes cuyo efecto se quiere anular en cada ocasión.

Por ejemplo, si queremos controlar por un período determinado una hipótesis que liga causalmente el número de partidos en las coaliciones gubernamentales con la estabilidad (en el sentido de duración temporal) de éstas y disponiendo de un número reducido de casos —por ejemplo, los países democráticos europeos de la postguerra— se podrá establecer la existencia de una cierta relación más o menos fuerte entre las dos variables. Sin embargo, habrá que controlar también que esa relación no esté, a su vez, influida o incluso determinada por otros factores externos. Es razonable sospechar que la existencia o no de determinados mecanismos institucionales destinados a estabilizar el ejecutivo (como, por ejemplo, los que imponen de hecho el recurso a las elecciones en cada caída del gobierno) tenga un peso en la relación encontrada. Para controlar ese peso y anularlo de la relación descubierta, a nivel general es preciso clasificar los casos en grupos según la existencia o no de estos mecanismos (y de su mayor o menor importancia) y volver a estudiar la relación en el interior de cada subclase de casos. En cada una de ellas, la influencia de los citados mecanismos institucionales se parametriza; es decir, se mantiene constante. Comparando la naturaleza de la relación entre número de partidos en el gobierno y la duración de las coaliciones para cada clase se tiene la posibilidad de evaluar el impacto de las variables intervinientes sobre la relación entre variables independiente y dependientes, así como, al mismo tiempo, verificar la validez de la relación avanzada como hipótesis. La operación se puede repetir para otras variables reclasificando los casos respecto a sus estados.

Este ejemplo aclara la lógica del control comparado y permite introducir la discusión de una serie de problemas que sólo podremos enunciar. En primer lugar, el método de control comparado pone claramente en evidencia la importancia de los problemas de tratamiento vertical (escalas de abstracción) y horizontal (clasificación) de los conceptos. Los problemas de clasificación son parte esencial del método comparado, en tanto que constituyen un instrumento para controlar las fuentes de variación. Las entidades que pertenecen a una misma clase no varían respecto a la característica que define la clase. En consecuencia, al clasificar los casos para el análisis comparado el investigador hace así que ciertas características de un fenómeno se transformen en parámetros para los fines del análisis. En este sentido la clasificación ya convierte potenciales variables operativas en parámetros y toda posterior subclasificación añade otro control paramétrico adicional.

Incluso en lo que se refiere al tratamiento vertical a lo largo de escalas de abstracción, el método comparado pone de manifiesto algunos problemas generales de la investigación. Comparando unidades socio-políticas diferentes —en la mayor parte de los casos, países y sistemas políticos distintos— es preciso recurrir a conceptos comparados no peculiares de una determinada cultura; o bien a conceptos más abstractos y extensivos, cuyas operativizaciones necesitan procesos de indicación menos directa y aumentan los problemas de la validez de los

indicadores. En otras palabras, la naturaleza muy cualitativa del método comparado pone claramente en evidencia —sin que se pueda esconder tras soluciones cuantitativas— la tensión constante entre la escasa inclusividad (extensión) de categorías culturalmente específicas y la carencia de contenido de las muy inclusivas. Lo que exige una atención constante al análisis conceptual en términos de relaciones entre intensión y extensión, dado que la capacidad de comparación depende en definitiva del nivel de generalidad del lenguaje a que se recurre³².

Por último, hay que mencionar una serie de problemas intrínsecos al método comparado que identifican sus debilidades internas y que hacen que sea el menos «fuerte» de los métodos de control, aunque a veces es el único a que se puede recurrir. Desde la perspectiva del control paramétrico de las variables intervinientes existe una contradicción intrínseca entre la necesidad de ampliar lo más posible el número de casos a contemplar y la de reducirlo en la misma medida. Tanto para dar un carácter más general a las conclusiones como para poder realizar las parametrizaciones de más variables independientes o intervinientes, es necesario ampliar el número de casos, bien directamente o bien introduciendo la dimensión temporal —permitiendo que los estados de las variables cambien no sólo de caso en caso sino también de período en período para cada caso³³. De hecho, la posibilidad de ganar control sobre las variables intervinientes a través de clasificaciones y subclasificaciones se topa con el límite que le impone la restricción del número de casos que puede crear rápidamente clases con frecuencia empírica demasiado baja como para poder establecer conclusiones. Por otra parte, es evidente que el aumento del número de casos determina el aumento de las potenciales condiciones causales y, por tanto, el número de variables intervinientes que hay que controlar.

Para conseguir el máximo control sobre el contexto de la comparación habría que limitarla a áreas político-culturales reducidas, a casos homogéneos, o bien a pocos casos. En tal eventualidad los casos más homogéneos en relación a los contextos permiten considerar como parámetros todas sus semejanzas y homogeneidades, anulando como constante su influencia sobre la relación que se estudia. Dicho de otro modo, la parametrización de las variables intervinientes se realiza en una cierta medida *ante-facto* en el mismo proceso de selección de los casos³⁴.

Sin embargo, el número de casos homogéneos es tan limitado que no tendremos muchas posibilidades de parametrizar mediante su reagrupación a las variables independientes e intervinientes. Además, las conclusiones teóricas que se deriven de esta estrategia tienen un alcance limitado en cuanto que, precisamente, en cada caso el papel de las variables intervinientes no está controlado explícitamente. De hecho, en las comparaciones entre casos de máxima homogeneidad las variables contextuales intervinientes se convierten en parámetros sin influencia *sólo respecto a aquellos casos*. En cuanto se amplía el número de casos, algunas de esas variables parametrizadas, al cambiar de estado, vuelven a entrar como variables operativas potenciales. Lo que permite formular hipótesis y teorías más completas, pero que nos lleva al problema, recién discutido, inherente a la ampliación de los casos y de las observaciones. Por estos motivos se ha defendido la necesidad de una estrategia comparativa basada en la selección de casos muy

³² SARTORI, G., *La Política...*, op. cit.

³³ LIJPHART, A., *El método...*, op. cit.

³⁴ LIJPHART, A., *The Comparable Cases Strategy in Comparative Research*, en «Comparative Political Studies», 7, 1975

heterogéneos en los contextos. En este caso se aplica una lógica distinta: toda *similitud* manifestada en las relaciones entre variables en contextos heterogéneos lleva a considerar como no influyentes sobre la relación las heterogeneidades contextuales. A estas últimas se las hace intervenir en la interpretación sólo cuando se interrumpen las semejanzas³⁵. No nos detenemos más sobre este enfoque alternativo —comúnmente conocido como método de los *most dissimilar systems*—, que parece capaz de producir sus mejores resultados sobre todo en el campo de los análisis basados sobre datos de sondeos internacionales.

El problema inherente a la elección del número o del tipo de casos da origen de hecho a diversos tipos de diseños comparados de investigación —caracterizados por varias ventajas/inconvenientes en cuanto al problema del control de las variables intervinientes y que no podemos examinar en detalle. Hay que subrayar en primer lugar que el método comparado tiene debilidades evidentes como método de control y es claramente menos potente que los métodos anteriores. Pero, con frecuencia, es el único método al que el investigador puede recurrir y está dotado de sus reglas específicas y de procedimientos de control que lo distinguen del método experimental y del estadístico. Hacer comparaciones no es lo mismo, pues, que utilizar el método comparado de control, del mismo modo que el recurso a datos y fuentes estadísticas no significa utilizar el método estadístico. De ahí deriva la falta de fundamento de todas esas argumentaciones que confunden los dos métodos, afirmando que «cuando se compara se usan también instrumentos estadísticos» o que, al revés, «cuando se realizan análisis estadísticos hace falta de algún modo comparar».

En estos últimos apartados hemos discutido los distintos métodos desde la perspectiva de su capacidad de control de hipótesis ya formuladas. Hemos subrayado que ninguno de estos métodos se puede utilizar en una investigación si no es referido a algunas hipótesis de partida que, como son generales, no circunscriben el ámbito de aplicación. El recurso a estas metodologías, sin duda, puede no sólo llevar a reformulaciones y afinamientos de las hipótesis iniciales, sino también, en ciertos casos, a la formulación de hipótesis completamente nuevas. Y esta producción y refinamiento se producen en el contexto de una aplicación del método tendente a *controlar* la validez de las hipótesis iniciales a partir de las cuales ha empezado la investigación. Por este motivo hemos preferido caracterizar estos métodos sobre todo como métodos de control más que de formulación de hipótesis. Por otra parte, tales métodos no poseen en exclusiva el monopolio del proceso de formación de las hipótesis. Si tienen, en cambio —o deberían tenerlo en una buena investigación—, el monopolio de los procesos de control y verificación.

Estas observaciones permiten arrojar luz sobre el espacio y el papel reservado al *método histórico*. El método histórico al que se hace referencia en los estudios de ciencia política no es, claro está, el método historiográfico que utilizan los historiadores, sino más precisamente un método que trata de formular hipótesis y generalizaciones basándose en observaciones y ejemplos *extraídos* de la historia y/o apoyándose en ellos³⁶. Es el método utilizado por tantos clásicos de la ciencia política y que ha logrado algunos de sus resultados más importantes y estimulantes. Y que todavía hoy es utilizado ampliamente por estudiosos de las ciencias

³⁵ PRZEWORSKI, A., y TEUNE, H., *The Logic...*, op. cit.; MECKSTROTH, T. W., «*Most Different Systems*» and «*Most Similar Systems*»: A Study of the Logic of Comparative Inquiry, en «*Comparative Political Studies*», 7, 1975, pp. 132-157.

³⁶ SARTORI, G., *La Política...*, op. cit.

políticas y sociales y que caracteriza a una gran parte de la literatura³⁷. Por último, ese método ha dado abundantes pruebas, tanto en los clásicos como en los contemporáneos, de su fecundidad como método de producción de hipótesis. Desde el punto de vista que aquí nos interesa, sin embargo, el problema se plantea en términos distintos: ¿hasta qué punto el método histórico sirve y tiene su propia autonomía como *método de control* de las hipótesis?

En sentido general, el control histórico tiene esta función. Es mejor apoyar con ejemplos de casos históricos las propias hipótesis que no hacerlo. Sin embargo, cuando se va más allá de esta definición mínima de control, el método histórico muestra sus debilidades, que, para ser superadas, nos llevan inevitablemente a la lógica del método de control comparado. El método histórico extrae los casos de la experiencia histórica y es, por tanto, una manera de ampliar el número de observaciones a lo largo de una dimensión diacrónica. Sin embargo, dado que nadie aceptaría como método de control una elección de casos históricos hecha casualmente o con criterios no explícitos —incluso escogiendo aquellos casos que más convienen a las propias hipótesis y dejando a un lado los otros—, también en este método nos volvemos a encontrar con el problema del criterio que se sigue para la elección de los casos. También desde el punto de vista de la posibilidad de parametrizar posibles condiciones causales el método histórico no puede más que recurrir a aquellas técnicas de clasificación de los casos elegidos respecto a determinadas propiedades de ellos, si no quiere limitarse a una mera catalogación de ejemplos favorables o desfavorables. La lógica del control es la propia del método comparado. Desde el punto de vista metodológico, la elección de los casos y de las observaciones en una amplia dimensión histórica implica una única diferencia: aumenta la dificultad de formular hipótesis acerca de la constancia y la no influencia de las condiciones contextuales potencialmente importantes que, además, aumentan de número, dada la inevitablemente grande heterogeneidad de los contextos históricos de que se extraen los casos.

Como conclusión, el método histórico, que es tan fecundo desde la perspectiva de la formulación de hipótesis, no parece tan riguroso desde la del control. Para que ese control se realice con método se tiene que recurrir completamente a la lógica del control comparado y en condiciones particularmente difíciles.

16. El estudio de casos

El estudio de un caso singular —ya sea un sistema político, un grupo de presión, una ley electoral o la personalidad de un líder— ha sido tan olvidado por la reflexión metodológica como ampliamente usado en la práctica de la investigación; en la literatura politológica los trabajos que se centran sobre el estudio de un caso específico son decididamente mayoría. La razón es muy simple. Desde la perspectiva metodológica el estudio de un caso individual presenta límites insuperables en lo que se refiere a la posibilidad de establecer y verificar relaciones causales entre fenómenos. Dado que no hay ninguna variación (o como mucho una de tipo diacrónico, pero necesariamente limitada) en los efectos y presuntas causas, no hay ninguna base empírica para establecer una preferencia por alguna de las causas posibles ni para controlar la influencia de las posibles causas. Sobre

³⁷ KAPLAN, A., *The Conduct of Inquiry. Methodology for Behavioral Science*, Aylesbury, Bucks., International Textbook Company Ltd., 1973

este punto no debe quedar ninguna duda: en sí, el estudio de casos no tiene posibilidad de producir ninguna generalización o teoría causal dotada de alguna veracidad. Sin embargo, una serie de motivos válidos prácticos hacen que se recurra mucho a este tipo de análisis en la investigación. En primer lugar, los costes, en general inferiores en términos de acceso a los datos e informaciones, y la relativa simplicidad del diseño de la investigación. En el estudio de un caso se pueden soslayar gran parte de las cuestiones metodológicas típicas de la investigación comparada y en especial de la que se produce en culturas diferentes: menores problemas de equivalencia de los indicadores, posibilidad de recurrir a conceptos culturalmente claros sin problemas, posibilidad de recurrir con mayor confianza a ciertos procesos heurísticos como la cláusula *ceteris paribus* para conjuntos de variables que se intenta excluir del estudio. Además, el concentrarse sobre un caso único permite, en general, seguir una estrategia de investigación intensiva; es decir, tomar en consideración una gran cantidad de propiedades del caso y evaluar su importancia en la compleja red de relaciones que constituye todo fenómeno político-social³⁸.

Descontando, pues, las carencias desde el punto de vista de la producción de generalizaciones causales, los méritos reseñados más arriba constituyen otras tantas ventajas en lo que se refiere a otros pasos importantes del proceso cognoscitivo inherentes a la generalización de hipótesis y al afinamiento de las relaciones entre variables establecidas por otros estudios. Existen varios tipos de estudios de casos que, con una cierta simplificación, pero con más claridad, podemos resumir en el esquema de la figura 2.4³⁹, según sus características y, por tanto, sus funciones cognoscitivas⁴⁰.

El tipo de estudio de casos definido como *descriptivo a-teórico* se caracteriza por la falta de una formulación clara y explícita del problema en el sentido señalado al principio de este capítulo. El objetivo declarado no es el de afrontar un problema concreto formulado teóricamente, sino reconstruir imágenes configurativas generales de un fenómeno cuyos aspectos significativos se interpretan intuitivamente, apoyándose en el argumento de que un estudio intensivo y, a veces, una reflexión con empatía garantizan la «comprensión» del fenómeno.

Este tipo de estudio de casos no tiene posibilidad de contribuir autónomamente a la acumulación del conocimiento científico porque difícilmente produce resultados transmisibles intersubjetivamente (por mucha que sea la importancia de este tipo de estudios en la recogida y circulación de informaciones utilizables en distintos contextos de investigación).

La identificación y formulación de un problema concreto ofrecen en general al estudio de casos la posibilidad de formular *hipótesis interpretativas*. Un estudio de este tipo tiende a concentrar su atención en relaciones entre las propiedades del caso que se puedan generalizar. Aun cuando las hipótesis así formuladas no se pueden verificar en el interior del estudio, se pueden retomar en otras investigaciones de distinta naturaleza. Hoy en día, una gran parte de las hipótesis planteadas a partir de investigaciones comparadas tienen su origen en estudios inten-

³⁸ ECKSTEIN, H., *Case Study and Theory in Political Science*, en F. O. GREENSTEIN y N. W. POLSBY (eds.), *Handbook of Political Science*, vol. VII, *Strategies of Inquiry*, Reading (Mass.), Addison-Wesley Publishing Company, 1975, pp. 79-137; BROWN, B. E., *The Case Method in Comparative Politics*, en J. B. CRISTOPH y B. E. BROWN (eds.), *Cases in Comparative Politics*, Boston, Little, Brown, 1969, pp. 3-17.

³⁹ Tomo esta figura de un seminario celebrado en julio de 1979 en el Instituto Universitario Europeo por el profesor MORGENS PEDERSEN, al que agradezco que me permita publicarla.

⁴⁰ ECKSTEIN, H., *Case Study*, op. cit.

Tipo de estudio del caso	Estudio a-teórico investigación descriptiva	Estudio interpretativo generador de hipótesis	Estudio explicativo	Estudio de control de las teorías	Estudio del caso desviante
Identificación del problema teórico	NO	SI	SI	SI	SI
Conciencia de y recurso a teorías y generalizaciones existentes	NO	NO	SI	SI	SI
Control crítico de las teorías y generalizaciones existentes	NO	NO	NO	SI	SI
Conciencia y utilización de los casos desviantes de las teorías existentes	NO	NO	NO	NO	SI

FIG. 2.4.—Tipos de estudio de casos.

sivos de casos singulares. Esto también se debe al hecho de que la naturaleza intensiva del estudio del caso ofrece algunas ventajas: permite tomar en consideración muchos factores y, por tanto, formular diferentes hipótesis incluso alternativas.

El tipo *explicativo* de estudio de casos se caracteriza porque trata de explicar aspectos particulares recurriendo a generalizaciones y teorías existentes en la literatura. El acontecimiento o el fenómeno concreto se explican aplicando generalizaciones y teorías existentes a las condiciones específicas del caso. El modelo explicativo utilizado en este tipo de análisis es el deductivo-nomológico, aunque en las ciencias sociales tal tipo de explicaciones sólo son por lo general aproximativas.

Naturalmente, el intento de aplicar generalizaciones y teorías probabilísticas a la interpretación de casos singulares se transforma también en un ejercicio de control crítico, de modificación y revisión de aquellas frente a las condiciones específicas del caso. En esta función de *evaluación de la plausibilidad de teorías* formuladas o de la validez y utilidad de determinadas construcciones o esquemas teóricos, el estudio del caso ofrece evidentes ventajas. Es menos costoso que una investigación comparada y puede producir resultados útiles, en especial subrayando las condiciones de debilidad de una teoría.

Esta última función se desarrolla con mayor potencia aún en el *estudio del caso desviante*. Ese tipo de investigación se utiliza siempre referido a asociaciones generales establecidas mediante investigaciones de otro tipo (estadísticas comparadas, experimentales). Pero se diseña en base a un criterio preciso que maximiza su importancia teórica: la desviación del caso respecto a la tendencia general descrita y prevista por una teoría o generalización determinada. En la práctica, la lógica es la de la comparación entre el grupo de casos sobre los que se ha construido la teoría y/o se ha aplicado con éxito y el caso desviante, tratando de identificar las diferencias de este último respecto a los otros que puedan explicar la desviación. Se trata de un tipo de trabajo especialmente apto para identificar nuevas y añadidas variables respecto a las usadas normalmente y cuya efectiva influencia sólo se puede establecer más sólidamente recurriendo a otros métodos.

Como conclusión, cabe señalar que los diferentes tipos de estudio de casos se diferencian por las funciones que pueden desarrollar en el proceso cognoscitivo. Naturalmente, los de mayor complejidad pueden cumplir más de una función. En efecto, algunas funciones del estudio del caso que en el esquema se distinguen con fines analíticos, después no lo son en la actividad de investigación. El estudio de un caso dotado de ambiciones teóricas, y por tanto no limitado a objetivos descriptivos o a la formulación de hipótesis interpretativas, debería especificar al mismo tiempo lo que es explicable del caso refiriéndolo a generalizaciones y teorías existentes, lo que en cambio las contradice y falsifica y, por último, qué nuevos aspectos deberían ser tomados en consideración e incorporados en una versión revisada de la teoría para que ésta sea capaz de incluir también el caso estudiado.

17. Generalizaciones, teorías, explicaciones

En las últimas páginas hemos recurrido a los términos «generalización» y «teoría» sin detenernos en su papel y significado. Las generalizaciones y las teorías tienen en común la característica de ser, a la vez, el resultado y un punto de partida de la investigación. Son el resultado porque representan el objetivo más ambicionado, dado su potencial predictivo y explicativo y su capacidad de organizar los conocimientos de modo sistemático y eficaz; y son su base en cuanto que —según una orientación epistemológica que hoy prevalece— no se puede intentar ninguna explicación adecuada de un fenómeno sin recurrir a ellas.

Las *generalizaciones* son enunciados universales, que tienen la forma lógica del tipo «para todo X, si X entonces Y». En definitiva, se trata de hipótesis de relaciones entre fenómenos que han superado ya el estado de mera conjetura y han logrado un aceptable nivel de verificación a través de numerosas observaciones en casos y situaciones distintas. Se definen, pues, como universales por el carácter de su referente y como leyes o cuasi-leyes porque se enuncian como tales⁴¹.

Las generalizaciones son de distinto tipo. Según la base de la relación que liga a los elementos, la generalización puede ser determinista o probabilística. Según la dimensión temporal puede ser una generalización de sucesión o de concomitancia. En el primer caso, si se da X en el tiempo t , se dará Y en el tiempo t_1 , planteando, pues, una relación de sucesión temporal que es una condición importante para las leyes causales. En el segundo caso, en cambio, se establece una relación de implicación entre X e Y que introduce la posibilidad de generalizaciones de carácter funcional. Otra distinción se refiere al tipo de tratamiento de los conceptos que figuran en ellas (nominal o escalar) y, por tanto, la sintaxis que conecta a los términos. Por fin, una última diferencia entre tipos de generalizaciones originada por las propiedades semánticas de los términos que la forman, distingue entre generalizaciones teóricas y observativas.

Una *teoría* en sentido estricto está constituida por una serie de generalizaciones conectadas orgánicamente. El principio organizador puede ser de distinta naturaleza. El tipo de organización que se considera necesario en la literatura epistemológica que hace referencia a las ciencias físico-matemáticas es el de la conexión deductiva⁴². En este caso, para que un conjunto de generalizaciones pueda defi-

⁴¹ RUDNER, R. S., trad. castellana, *Filosofía de la Ciencia Social*, Madrid, Alianza, 1980.

⁴² NOVAK, S., *Understanding and Prediction. Essays in the Methodology of Social and Behavioral Theories*, Dordrecht, Reidel, 1976; HOLT, R. T., y RICHARDSON, J. M., *Competing*

nirse como teoría tiene que cumplir la condición de estar organizado deductivamente. Dicho de otro modo, las teorías tienen que tener la forma de sistemas axiomáticos, contruidos por generalizaciones con carácter axiomático que constituyen las premisas del razonamiento deductivo, el cual produce otras generalizaciones como teoremas que se desprenden lógicamente de los axiomas y sirven como conclusión del razonamiento deductivo.

El desarrollo de teorías axiomáticas en que estén explícitos tanto los términos primitivos —o postulados— como los axiomas y las reglas de transformación de éstos (el modo en que se pueden hacer deducciones de los axiomas) en teoremas es una posibilidad difícilmente alcanzable en las ciencias sociales actuales. Este significado «fuerte» del término teoría representa una ambición que puede tener, como máximo, el papel de un ideal al que aspirar tendencialmente⁴³. En algunos campos y enfoques de las ciencias sociales se pueden lograr niveles de semi-formalización de las construcciones teóricas en las que la explicitación afecta solamente a los términos primitivos y los axiomas, pero no a las reglas de transformación de éstos en teoremas. En general, sin embargo, en las ciencias sociales se tiende a identificar con el término «teoría» resultados de distinta naturaleza y de menor valor, que algunos consideran que sería mejor llamar con el nombre de pre-teorías o de inventarios de generalizaciones o, sencillamente, de esquemas conceptuales, y que, en todo caso, basan su validez en el grado de confianza de que gozan en la comunidad científica⁴⁴.

Se trata de conjuntos de enunciados/generalizaciones convalidados empíricamente y referidos al mismo problema, pero que están conectados de modo aproximado y no sistemático, y sólo para algunos de los cuales se puede explicitar el papel axiomático. También un conjunto de hipótesis no suficientemente verificadas puede estar organizado así, formulando estas pre o cuasi-teorías por las ventajas que ofrecen en la orientación de la investigación. Naturalmente, en el caso de pre-teorías no axiomáticas cada proposición no funda su validez en reglas precisas de deductibilidad, sino que cada una se mantiene en relación a las otras en una situación que puede ser de confirmación recíproca, pero también de debilitamiento recíproco.

El desarrollo de este tipo de pre o cuasi-teorías en las ciencias sociales no debe infravalorarse; tiene una importancia fundamental. En primer lugar, conjuntos más o menos fuertemente interrelacionados de generalizaciones referentes a un problema permiten organizar el conocimiento del modo más sistemático y eficaz posible, favoreciendo, al menos en parte, los procesos de acumulación y ofreciendo una guía y un principio fuerte de organización de la investigación. Aun cuando su axiomatización sea incompleta, una pre-teoría tiene la capacidad de generar nuevas hipótesis que someter a control empírico y su existencia permite identificar diferentes tipos de incongruencias dentro de la misma teoría, entre las diversas teorías y entre la teoría y la realidad observada. Cumplen, así, la importante función de orientar los ulteriores esfuerzos de investigación en direcciones teóricamente fecundas.

Además, las generalizaciones que componen las cuasi-teorías son la base del proceso de explicación de carácter nomológico-deductivo —al que se ha hecho

Paradigms in Comparative Politics, en R. T. HOLT y J. E. TURNER (eds.), *The Methodology of Comparative Research*, Nueva York, Free Press, 1970, pp. 21-71.

⁴³ KAPLAN, A., *The Conduct...*, op. cit.

⁴⁴ GALTUNG, J., *Theory and Methods*, op. cit.

referencia al hablar del estudio del caso de carácter explicativo— que «explica» un acontecimiento sobre la base de generalizaciones frecuenciales o probabilísticas y de circunstancias particulares que especifican aquella determinada generalización⁴⁵. Otros tipos de explicaciones, como la de naturaleza analógica o de la llamada «comprensión» no son plenamente satisfactorios para las ciencias sociales y tienen límites objetivos⁴⁶.

Las explicaciones analógicas se basan en el proceso de reducción de algo no familiar a algo familiar (con mucha frecuencia a modelos mecánicos), pero es problemático asegurar que los fenómenos sociales tengan estructuras lógicas isomórficas respecto a las de los modelos mecánicos o de otro tipo conocidos. Además, aun cuando las estructuras lógicas del modelo conocido y del fenómeno estudiado fueran isomórficas, podrían no serlo los contenidos empíricos sustantivos. Las explicaciones basadas en la llamada «comprensión» del fenómeno, concepción ligada al historicismo y a la filosofía de los valores, entienden por explicación el proceso lógico por el cual se reconoce una relación existente entre un tipo de comportamiento de hecho y un valor⁴⁷. También en este caso se han subrayado los límites y las carencias para las ciencias sociales de este método, inherentes sobre todo a la explicitación del valor que en el actor puede ser inconsciente, a la polivalencia del método de consideración del fenómeno mismo —dado que los referentes son valores— y al hecho de que, en definitiva, tal proceso explicativo es más un proceso de mera medida axiológica que de certificación lógica del acontecimiento.

No podemos profundizar más aquí este debate y la referencia a otros tipos de explicaciones sólo sirve para subrayar que, actualmente, el nomológico es el método explicativo al que se reconoce como más adecuado. Al mismo tiempo, sin embargo, hay que subrayar que en las ciencias sociales las explicaciones basadas en la deducción de una proposición que describe un acontecimiento de un conjunto de premisas que contengan descripciones de las condiciones específicas y generalizaciones empíricas tendrán en cualquier caso el carácter de la no adecuación, dado todo lo que hemos venido diciendo acerca de la dificultad de identificar todas las condiciones causales y las condiciones paramétricas del caso concreto. En la práctica, las explicaciones a que se pueda llegar serán a menudo incompletas, bien porque no se especifican enteramente las condiciones paramétricas o porque las explicaciones sólo ofrecen las líneas generales de lo que puede llegar a ser una explicación adecuada, elaborando y mejorando el razonamiento explicativo y especificando de manera más completa las generalizaciones y condiciones⁴⁸.

Es evidente en este punto la importancia del papel del estudio del caso —y en especial, del caso desviante— con vistas al control de teorías y generalizaciones. La verificación o falsificación de una teoría (o pre-teoría) —que con frecuencia se expresa con términos teóricos o en todo caso no directamente observables— se realiza directamente, a través del control de las hipótesis empíricas derivables de ella. En este sentido, toda teoría o generalización es un instrumento predictivo

⁴⁵ PASQUINELLI, A., *Nuovi principi di epistemologia*, Milán, Feltrinelli, 1971.

⁴⁶ BROWN, R., *Explanation...*, op. cit.

⁴⁷ STRAUSS, L., *An Epilogue*, en H. J. STORING (ed.), *Essays on the Scientific Study of Politics*, Nueva York, Holt, Rinehart y Wiston, 1962, pp. 310-320; ROSSI, P., *Spiegazione e comprensione nell'indagine sociale*, en *Atti del Convegno di Studi Metodologici*, Turin, Taylor, 1956, pp. 399-406.

⁴⁸ LEIGH, S.; JOHNSON, B. K. F.; PAULSEN, D. W., y SCHOCKET, F., *Political Research Methods. Foundations and Techniques*, Boston, Houghton Mifflin Co., 1976.

en tanto que prevé el valor de un determinado resultado como función del valor de un conjunto específico de variables y constantes.

La teoría debería controlarse comparando los valores del resultado previsto por ella con el de los fenómenos reales. En definitiva, una teoría de la estabilidad-inestabilidad democrática que especifique las variables importantes, las constantes, así como la relación (función) que las liga a la inestabilidad-estabilidad democrática se controla respecto a la congruencia entre las evoluciones reales o evoluciones previstas en base a los datos de una situación real. En este sentido, pues, capacidad predictiva no significa sólo capacidad de prever el futuro o evoluciones futuras. Las capacidades predictivas de una teoría pueden ser evaluadas retrospectivamente respecto a acontecimientos del pasado y también respecto a otros acontecimientos o casos contemporáneos aún no examinados. En estos casos se conocen los datos de la situación real. Por el contrario, el problema se complica cuando la teoría se usa prospectivamente, para prever desarrollos futuros, en cuanto los datos de la situación real tienen que ser estimados por sí mismos.

Dada la debilidad de los métodos de control de las ciencias humanas, es también débil la fiabilidad de los métodos de identificación de las variables importantes, de los parámetros y de las relaciones que les ligan y, por tanto, es frecuente el caso en que la extensión de la teoría a casos no examinados o a nuevos desarrollos no hace más que ayudar a aclarar cómo se modifica la misma teoría. Si los resultados previstos sobre la base de la teoría y de las circunstancias específicas del caso no corresponden a los reales, se impone una modificación o bien del número de variables importantes, o de los parámetros, o de las relaciones entre variables a partir de las nuevas informaciones.

Esta situación da testimonio de la constante naturaleza incierta e hipotética de los resultados científicos y diseña también un proceso de acumulación científica, mediante el cual las teorías formuladas se modifican en sus parámetros y en sus formas funcionales en base a informaciones de *feed-back* sobre su rendimiento. Este proceso es, en sustancia, la base de aquella posición epistemológica que pide para las teorías no la investigación de su comprobación, sino más bien la investigación de su falsificación como punto de partida para su desarrollo y mejora. Frente a resultados de nuevas investigaciones que no corresponden a los alcanzados a partir de las teorías existentes, el investigador debería modificarlas de manera que permita adaptarlas a los nuevos resultados, manteniendo al mismo tiempo su valor para los viejos⁴⁹.

18. Conclusión

Del desarrollo de una simple curiosidad intelectual a la identificación de generalizaciones y a la formulación de teorías, la investigación procede a través de estadios que son analíticamente distintos, pero que en la práctica se superponen y se entrecruzan. Cada nuevo paso tiene un efecto sobre los anteriores. Así como hipótesis nuevas o refinamientos de hipótesis existentes pueden desarrollarse en cualquier fase, el análisis de los datos y de las informaciones puede sugerir el recurso a nuevos conceptos, el desarrollo de nuevas variables y, a su vez, la bús-

⁴⁹ POPPER, K., *Logik der Forschung*, Viena, Springer, 1955; trad. castellana *Lógica de la investigación*, Barcelona, Laia, 1980.

queda de otros datos e informaciones. La formulación de generalizaciones y (pre-) teorías, que analíticamente puede considerarse como objetivo de la investigación científica, no es en realidad, desde otro punto de vista, más que el comienzo de identificación de nuevos problemas.

En la praxis de la investigación será también necesario recurrir a distintas combinaciones de los procedimientos diferenciados analíticamente en este capítulo. En la propia investigación se pueden elegir, o estar obligado a recurrir, a procesos distintos de identificación y operativización de los conceptos (y no siempre tendremos a nuestra disposición las técnicas necesarias de medida), a métodos de control diferentes y más o menos potentes, a diseños de estrategias de investigación de distinto tipo para afrontar aspectos diferentes de un mismo problema.

La importancia de estos problemas, como de todos los que hemos ido subrayando en las diversas partes del capítulo, puede ser, a veces, fuente de frustración para el investigador social y político, especialmente si éste asume, más o menos implícitamente, como modelo la potencia metodológica de las ciencias naturales y en particular de las físicas, que son las que más han contribuido a la definición dominante de «método científico». No es raro que el extremismo «de física» produzca por reacción el extremismo contrario: el rechazo por prejuicio de un riguroso discurso de método en las ciencias humanas. No es con esta intención con la que se han subrayado los problemas inherentes a la investigación político-social, sino más bien en la perspectiva de que la conciencia plena de ellos se convierta en el fundamento explícito de procedimientos lógicos capaces de garantizar el máximo de intersubjetividad, publicidad y transparencia de los resultados de las investigaciones.

Referencias bibliográficas

- ACKOFF, R. L., *The Scientific Method. Optimizing Applied Research Decisions*, Nueva York, John Wiley & Sons, 1962.
- , *The Design of Social Research*, Chicago, University of Chicago Press, 1965.
- BLALOCK, H. M., jr., *Social statistics*, Nueva York, McGraw-Hill, 1960; trad. italiana, *Statistica per la ricerca sociale*, Bolonia, Il Mulino, 1969.
- , *Causal Inferences in Nonexperimental Research* (1961), Nueva York, Norton Library, 1972.
- BOUDON, R., y LAZARSFELD, P. F., trad. castellana, *Metodología de las Ciencias Sociales*, Barcelona, Laia, S. A., 1985 (2 tomos).
- BROWN, R., *Explanation in Social Science*, Chicago, Aldine, 1963.
- BROWN, B. E., *The Case Method in Comparative Politics*, en J. B. CHRISTOPH y B. E. BROWN (eds.), *Cases in Comparative Politics*, Boston, Little, Brown and Co., 1969, pp. 3-17.
- BRUSCHI, A., *Scienza, tecnica e senso comune*, en A. BRUSCHI, U. GORI y P. ATTINA, *Relazioni internazionali. Metodi e tecniche di analisi*, Milán, Etas Kompass, 1975, pp. 5-33.
- CAMPBELL, D. T., y STANLEY, J., *Experimental and Quasi-Experimental Research*, Chicago, Rand McNally, 1963; trad. castellana, *Diseños experimentales y cuasiexperimentales en la investigación*, Buenos Aires, Amortortu, 1979.
- CARLEY, M., *Social Measurement and Social Indicators. Issues of Policy and Theory*, Londres, Allen & Unwin, 1981.
- ECKSTEIN, H., *Case Study and Theory in Political Science*, en F. I. GREENSTEIN y N. W. POLSBY, *Handbook of Political Science*, vol. VII, *Strategies of Inquiry* (Reading (Mass.)), Addison-Wesley Publishing Company, 1975, pp. 79-137.
- FISICHELLA, D. (ed.), *Metodo scientifico e ricerca politica*, Roma, La Nuova Italia Scientifica, 1985.
- GALTUNG, J., *Theory and Methods of Social Research*, Oslo, Universitetsforlaget, 1967.

C. EJERCICIOS

1. ¿Cuáles son las cuatro características que debe reunir la formulación de un problema empírico de investigación? Responda de acuerdo con el texto de Bartolini.

2. Roberto Hernández, Carlos Fernández y Pilar Baptista mencionan tres criterios para plantear el problema de investigación, explíquelos:

3. ¿A qué se refieren King, Kehoane y Verba cuando dicen que las diferencias entre las investigaciones cualitativas y cuantitativas son sólo de estilo?

4. Stefano Bartolini señala que el término *metodología* tiene tres acepciones, ¿cuáles son y en qué consisten?

5. Con base en la lectura de Roberto Hernández, Carlos Fernández y Pilar Baptista, relaciones las siguientes columnas:

- | | |
|---------------------------------------|--|
| () Objetivos de la investigación | 1. Factibilidad en la realización de un estudio en cuanto a la disponibilidad de recursos. |
| () Justificación de la investigación | 2. Orientan hacia las respuestas que se buscan con la investigación. |
| () Consecuencias de la investigación | 3. Indica el porqué de la investigación exponiendo sus razones. |
| () Planteamiento del problema | 4. Tienen la finalidad de señalar a qué se aspira en la investigación y deben expresarse con claridad pues son las guías de estudio. |
| () Viabilidad de la investigación | 5. Consecuencias positivas o negativas que el estudio implica en los ámbitos ético y estético. |
| () Preguntas de investigación | 6. Significa afinar y estructurar más las preguntas de investigación. |

6. Stefano Bartolini señala que la definición de conceptos plantea dos problemas de fondo, el de la ambigüedad y el de la vaguedad, ¿a qué se refiere con ello?

7. Mencione las características que deben tener las preguntas de investigación, según el texto de Roberto Hernández, Carlos Fernández y Pilar Baptista:

8. Con base en la lectura de King, Kehoane y Verba, desarrolle las diferencias entre la interpretación y la inferencia.

Interpretación	
Inferencia	

9. Con base en la lectura de Bartolini, ¿qué es la definición operativa de un concepto?

10. ¿Cuáles son y en qué consisten los criterios para juzgar las inferencias descriptivas?
Responda de acuerdo con la lectura de King, Kehoane y Verba.

D. AUTOEVALUACIÓN

1. Según Stefano Bartolini, ¿en qué nivel de abstracción se ubica la metodología de la investigación social?

- a) En el fundamento filosófico del conocimiento científico.
- b) En el nivel de las técnicas específicas de investigación.
- c) En un nivel intermedio entre la filosofía y las técnicas.
- d) En ninguno de los anteriores.

2. Según Gary King, Robert O. Keohane y Sidney Verba, ¿qué es la inferencia descriptiva?

- a) Un proceso en el que se utilizan hechos que conocemos para aprender sobre los hechos que desconocemos.
- b) Una simplificación de algún aspecto del mundo y una aproximación al mismo.
- c) Un proceso mediante el cual se comprende un fenómeno no observado a partir de un conjunto de observaciones.
- d) Es la distinción entre los procesos sistemáticos y no sistemáticos de los fenómenos que estudiamos.

3. Según la lectura de Roberto Hernández, Carlos Fernández y Pilar Baptista, ¿cuáles son los criterios para plantear el problema de investigación?

- a) Establecer una relación entre variables, claridad y no ambigüedad y la posibilidad de realizar una prueba empírica.
- b) Establecer una relación causal entre variables, claridad y no ambigüedad y el uso de conceptos con gran nivel de abstracción.
- c) Establecer relaciones explicativas y descriptivas, claridad y no ambigüedad y la posibilidad de realizar una prueba empírica.
- d) Establecer una relación entre variables independiente de que el estudio sea cualitativo o cuantitativo, claridad y no ambigüedad y la posibilidad de realizar una prueba empírica.

4. Según el texto de Stefano Bartolini, ¿qué son las unidades de una investigación?

- a) Las características que se consideran importantes para una investigación.
- b) Cada uno de los elementos que componen propiamente un proyecto de investigación.
- c) Los tipos de objetos o acontecimientos de que se ocupa una determinada investigación.
- d) Los recursos que determinarán los alcances de una investigación

5. Según Stefano Bartolini, ¿qué son las propiedades en una investigación?

- a) Los atributos o cualidades esenciales de un proyecto de investigación.
- b) Los aspectos relativos a los objetivos, las preguntas y la justificación de un proyecto de investigación.
- c) Los elementos sistemáticos y no sistemáticos de las unidades de observación.
- d) Las características o dimensiones de las unidades que se consideran importantes para la investigación.

6. Según Luis Jesús Galindo, ¿cuáles son las dimensiones de organización de la información?

- a) la sociedad, la cultura y la ecología.
- b) la sociedad, la cultura y la personalidad.
- c) la sociedad, la política y la economía.
- d) la sociedad, la política y la ecología.

7. Según Gary King, Robert O. Keohane y Sidney Verba, ¿cuáles son las características del diseño de investigación?

- a) Tienen por objetivo la inferencia, los procedimientos son públicos, las conclusiones son ciertas, su principal contenido es el método.
- b) Tiene por objetivo la inferencia, los procedimientos son públicos, las conclusiones son inciertas, su principal contenido es el método.
- c) Tiene por objetivo la inferencia, los procedimientos son insondables, las conclusiones son ciertas, su principal contenido es el método.
- d) Tiene por objetivo la inferencia, los procedimientos son públicos, las conclusiones son inciertas, su principal contenido es su materia.

8. Según la lectura de Roberto Hernández, Carlos Fernández y Pilar Baptista, ¿cuáles son las características de las preguntas de investigación?

- a) Deben constituir ideas iniciales del planteamiento y deben ser lo suficientemente generales para abarcar la complejidad del problema.
- b) Deben estar sujetas a más de una interpretación en su significado y ser útiles para generar argumentos normativos.
- c) Deben resumir lo que habrá de ser la investigación, establecer los límites temporales y espaciales del estudio y esbozar un perfil de las unidades de observación.
- d) Deben contar con fundamentos empíricos que sirvan de base para las prescripciones normativas del investigador.

9. Según Gary King, Robert O. Keohane y Sidney Verba, cuáles son los criterios para juzgar las inferencias descriptivas.

- a) Falta de sesgo, eficiencia y coherencia.
- b) Falta de sesgo, ponderación y coherencia.

- c) Falta de sesgo, validez empírica y coherencia.
- d) Falta de sesgo, valor teórico y coherencia.

10. Según la lectura de Stefano Bartolini, ¿cuáles son las estrategias de investigación intensivas?

- a) Las que consideran un elevado número de unidades y una o pocas propiedades.
- b) Las que consideran una o pocas unidades y un elevado número de propiedades.
- c) Las que consideran el mismo número de unidades y propiedades.
- d) Ninguna de las anteriores.

Módulo II.
**Investigación cuantitativa
y cualitativa**

OBJETIVOS ESPECÍFICOS

- Identificar las características del análisis cuantitativo y cualitativo de datos, así como las diferencias entre estos tipos de análisis.
- Desarrollar las habilidades necesarias para el diseño de instrumentos para la investigación en derechos humanos.

A. SÍNTESIS

I. Enfoques en la investigación social

Los enfoques cuantitativo y cualitativo representan dos grandes vertientes epistemológicas en las ciencias sociales. Como señala Roberto Hernández, a lo largo de la historia de la ciencia han surgido diversas corrientes de pensamiento. Sin embargo, debido a las diferentes premisas que las sustentan, desde la segunda mitad del siglo xx tales corrientes se han polarizado en dos enfoques principales: el enfoque cuantitativo y enfoque cualitativo de la investigación (Hernández, p. 224). En términos generales, los dos enfoques utilizan cinco fases similares y relacionadas entre sí:

- Llevan a cabo observación y evaluación de fenómenos.
- Establecen suposiciones o ideas como consecuencia de la observación y evaluación realizadas.
- Prueban y demuestran el grado en que las suposiciones o ideas tienen fundamento.
- Revisan tales suposiciones o ideas sobre la base de las pruebas o el análisis.
- Proponen nuevas observaciones y evaluaciones para esclarecer, modificar, cimentar y/o fundamentar las suposiciones e ideas; incluso para generar otras (Hernández, p. 225).

Desde la perspectiva más tradicional de los enfoques cuantitativo y cualitativo, el primero procede por deducción y el segundo por inducción, es decir, la investigación cuantitativa va de la teoría a los hechos, mientras la cualitativa va de los hechos a la teoría. En este sentido se distinguen ambos enfoques del siguiente modo:

El *enfoque cuantitativo* utiliza la recolección y el análisis de datos para contestar preguntas de investigación y probar hipótesis establecidas previamente, y confía en la medición numérica, el conteo y frecuentemente en el uso de estadística para establecer con exactitud patrones de comportamiento de una población (Hernández, p. 225).

El *enfoque cualitativo*, por lo común, se utiliza primero para descubrir y refinar preguntas de investigación. A veces, pero no necesariamente, se prueban hipótesis. Con frecuencia se basa en métodos de recolección de datos sin medición numérica, como las descripciones y las observaciones. Por lo regular, las preguntas e hipótesis surgen como parte del proceso de investigación y éste es flexible, y se mueve entre los eventos y su interpretación, entre las respuestas y el desarrollo de la teoría. Su propósito consiste en “reconstruir” la realidad tal y como la observan los actores de un sistema social previamente definido (Hernández, p. 225).

En términos generales, en un *estudio cuantitativo* regularmente se elige una idea, que transforma en una o varias preguntas de investigación; luego, de éstas se derivan hipótesis y variables; se desarrolla una investigación para probarlas; se miden las variables en un determinado contexto; analiza las dimensiones obtenidas (con frecuencia utilizando métodos estadísticos), y establece una serie de conclusiones respecto de la(s) hipótesis (Hernández, p. 226).

Por su parte, las investigaciones *cualitativas* también son guiadas por áreas o temas significativos de investigación. Sin embargo, en lugar de que la claridad sobre la(s) pregunta(s) de investigación e hipótesis preceda a la recolección y el análisis de los datos, los estudios cualitativos pueden desarrollar preguntas o hipótesis antes, durante o después de la recolección y el análisis (Hernández, p. 228). En resumen, las diferencias entre ambos enfoques de investigación se destacan en el cuadro siguiente:

Cuadro II.1. Diferencias entre los enfoques de investigación

Enfoque cualitativo	Enfoque cuantitativo
Basado en la deducción	Basado en la inducción
Sus fundamentos son fenomenológicos	Sus fundamentos son positivistas
Pretende construir la teoría durante el proceso de investigación	Pretende poner a prueba la teoría
Intenta comprender el sentido que los actores le atribuyen a sus acciones	Intenta identificar las regularidades en la vida social
Observación libre. Pretende conocer la realidad desde la perspectiva de los actores	Medición penetrante y controlada
Intenta explorar, comprender, describir, interpretar	Pretende medir, confirmar datos, explicar relaciones causales
Los datos que recoge son complejos, profundos, de grandes dimensiones	Los datos que recoge son simples y repetibles
No pretende generalizar sus resultados	Pretende generalizar sus resultados
Sus técnicas de investigación son la entrevista a profundidad y la observación participante	Su técnica de investigación por excelencia es la encuesta con instrumentos estandarizados
Utilizan estrategias intensivas. Consideran muchas propiedades en pocos casos	Utilizan estrategias extensivas. Consideran pocas propiedades en muchos casos

Estas son las principales características de ambos enfoques. Aun cuando en apariencia sean incompatibles, ambos deben seguir las reglas de cualquier investigación científica.

Gary King, Robert Keohane y Sydney Verba señalan que 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 cualitativistas como los cuantitativistas se beneficiarían si se les concediera una atención más explícita al diseñar investigaciones (King, Keohane y Verba, p. 60).

Los estilos de la investigación cualitativa y cuantitativa son muy diferentes. La primera se sirve de números y métodos estadísticos. Suele basarse 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 detallados de materiales históricos; utiliza un método discursivo e intenta estudiar de forma global o exhaustiva un acontecimiento o unidad. Los cualitativistas, aunque sólo tengan un número reducido 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, institución, lugar, problema o ley (King, Keohane y Verba, p. 60).

Durante varias décadas –señalan King, Keohane y Verba–, 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 idea. Esta diferencia de opiniones genera una viva polémica, pero, por desgracia, también divide la disciplina en dos ramas, una cuantitativa-sistémica-generalizadora y otra cualitativa-humanista-discursiva. A medida que la primera aumenta la complejidad de sus análisis estadísticos, 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 (King, Keohane y Verba, p. 60).

No obstante, en las ciencias sociales, la investigación, cuantitativa o cualitativa, tiene dos objetivos: describir y explicar. Unos investigadores se proponen describir el mundo, y otros explicarlo, lo cierto es que ambos propósitos son esenciales. No podemos dar explicaciones causales con sentido sin una buena descripción y ésta, a su vez, pierde gran parte de su interés si no está vinculada a una relación causal (King, Keohane y Verba, p. 90).

En síntesis, las posiciones más conservadoras o tradicionales de los enfoques cualitativos y cuantitativos consideran que ambas posturas son incompatibles porque sus supuestos epistemológicos son distintos. Sin embargo, las vertientes más recientes consideran que las diferencias entre ambos enfoques son sólo de estilo, pues cualquier investigación que se precie de ser científica debe ser capaz de obtener conclusiones veraces a partir de los datos que se recogen.

Una vez conocidos los supuestos y posiciones de ambos enfoques, a continuación se describirán dos técnicas correspondientes a cada uno de ellos, la encuesta y el grupo de discusión.

II. La metodología de la encuesta

De acuerdo con Heriberto López Romo, la encuesta se ha convertido en una herramienta fundamental para el estudio de las relaciones sociales. Esta técnica tiene características indiscutibles que explican su creciente demanda y aceptación entre los estudiosos de las diversas caras de la sociedad contemporánea. Algunas de estas características son:

- a) *Variación y flexibilidad de aplicaciones.* La encuesta ha demostrado ser útil en una gran variedad de campos y aplicaciones.
- b) *Comparabilidad.* El protocolo de la encuesta, por naturaleza sistemático, permite comparar y contrastar los resultados de manera directa y objetiva entre áreas, países y aplicaciones.

- c) *Desarrollo tecnológico.* Si bien los principios de la encuesta siguen siendo los mismos, los sistemas de levantamiento, codificación, captura, validación, procesamiento y presentación de la información se han modificado significativamente.
- d) *Oportunidad.* Los estudios realizados con esta técnica permiten ofrecer resultados con gran velocidad.
- e) *Comprensión.* Los usuarios pueden comprender con facilidad el método de trabajo y las implicaciones de los procedimientos y resultados.
- f) *Precisión.* La técnica permite conocer con precisión los resultados y el riesgo implícito de las decisiones que se toman con base en esos resultados.
- g) *Costo/eficiencia.* Debido al conocimiento y al desarrollo tecnológico la encuesta proporciona resultados con un alto índice de costo/eficiencia (López, p. 244).

Desde un enfoque metodológico, las encuestas pueden servir para describir o explicar un fenómeno. Es decir, con ellas se puede obtener promedios y porcentajes que den cuenta de las características de un fenómeno, o bien, establecer relaciones causales entre dos o más variables. Las etapas de la encuesta son:

1. Identificación del problema
2. Selección de la técnica
3. Diseño de la muestra
4. Diseño del instrumento
5. Recolección de la información
6. Procesamiento y análisis de la información
7. Generación de informes

El contenido de estas etapas se verá en detalle en el módulo IV, dedicado exclusivamente a la estadística. Ahora describiremos brevemente una técnica de investigación cualitativa utilizada con frecuencia en los estudios de derechos humanos, los grupos de discusión.

III. Grupos de discusión

Los grupos de discusión representan una técnica de investigación cualitativa que tiene por objeto conocer las opiniones o creencias de un grupo de personas sobre un tema específico. A diferencia de la encuesta, los grupos de discusión trabajan en el nivel del diálogo, es decir, mediante la discusión de los participantes se busca reconstruir un discurso grupal. Además, los grupos de discusión, a diferencia de la encuesta, no buscan generalizar sus resultados a toda la población, sino profundizar en un tema específico.

Esta técnica se puede utilizar para fines muy diversos; por ejemplo, conocer la opinión sobre un producto que será lanzado al mercado, para conocer las reacciones ante una política pública, para saber las percepciones en torno a temas específicos como la diversidad sexual, la violencia hacia la mujer, etcétera. Para finalizar la síntesis de este módulo, se describirán las principales características del grupo de discusión.

a. Selección de los participantes

Este proceso se realiza de acuerdo con los objetivos de la investigación. Como señala Russi Alzaga, los objetivos del estudio y el tema que tratarán éstos determinan en gran medida el tipo y el procedimiento de configuración que tendrá el grupo o los grupos que se vayan a realizar (Russi, p. 314). Por ejemplo, si en el campo de los derechos humanos se desea conocer las opiniones de la ciudadanía sobre una ley a favor del aborto, ¿a quién se le pregunta? En principio se puede realizar una encuesta dirigida a la población pero si se quisiera conocer una opinión profunda y especializada se podría invitar a representantes de organizaciones de la sociedad civil a grupos de enfoque en donde se discutan las implicaciones de dicha ley.

b. Integrantes del grupo de discusión

En los grupos de discusión participan las y los invitados y el moderador o moderadora, que es un integrante del equipo de investigación y es la persona responsable de dirigir la sesión. Las intervenciones del moderador están en función de los objetivos específicos que se persigan en el grupo de discusión. Regularmente se encarga de introducir los temas, evitar la dispersión durante la sesión, conceder la palabra y limitar los tiempos de las intervenciones. En cualquier caso no asume una posición durante la discusión, sólo la dirige.

c. Tamaño del grupo

Algunos autores sugieren que el grupo sea de 4 a 10 integrantes, ello permite el manejo eficiente de la discusión. Sin embargo, el número está en función de la complejidad del tema. Si éste es muy complejo, entonces es conveniente conformar los grupos con no más de cuatro personas. En realidad, más importante que el número de participantes es su adecuada selección y el manejo del grupo.

B. MATERIAL DE ESTUDIO

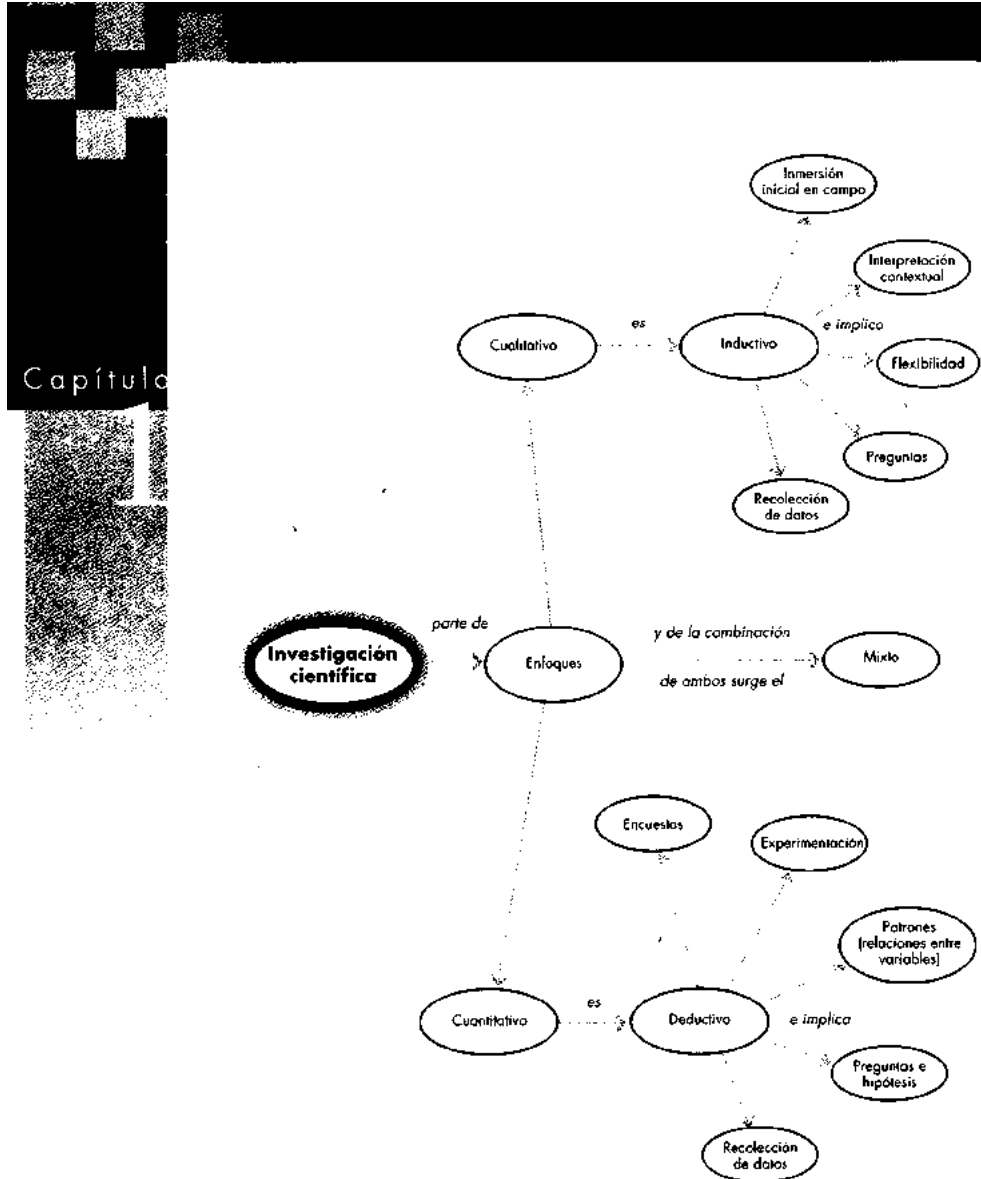
Bibliografía obligatoria

Hernández, Roberto *et al.*, "El proceso de investigación y los enfoques cuantitativo y cualitativo: hacia un modelo integral", en *Metodología de la investigación*, México, McGraw-Hill, 2003, pp. 3-26 **[Lectura 1]**.

López Romo, Heriberto, "La metodología de la encuesta", en Luis Jesús Galindo Cáceres (coord.), *Técnicas de investigación en sociedad, cultura y comunicación*, México, Pearson, 1998, pp. 33-73 **[Lectura 2]**.

Russi, Bernardo, "Grupos de discusión. De la investigación social a la investigación reflexiva", en Luis Jesús Galindo Cáceres, pp. 75-115 **[Lectura 3]**.

Lectura 1



Tomada de: Roberto Hernández *et al.*, "El proceso de investigación y los enfoques cuantitativo y cualitativo: hacia un modelo integral" en *Metodología de la investigación*, México, McGraw-Hill, 2003, pp. 3-26.

El proceso de investigación y los enfoques cuantitativo y cualitativo: hacia un modelo multimodal

Síntesis

El capítulo define los enfoques cuantitativo y cualitativo de la investigación. Así mismo, presenta las etapas del proceso de investigación de manera genérica y los aplica a ambas perspectivas. Además, propone una visión respecto de la investigación que implica la posibilidad de mezclar las dos modalidades de generación de conocimiento en un mismo estudio, a cual se denomina enfoque "multimodal" de la investigación.

Capítulo

1

Objetivos de aprendizaje

Al terminar este capítulo, el alumno será capaz de:

- Fortalecer su conocimiento del proceso de investigación.
- Conocer la perspectiva o el enfoque multimodal en la investigación científica.
- Determinar las diferencias entre los enfoques cuantitativo y cualitativo de la investigación.
- Aprender a mezclar los enfoques cuantitativo y cualitativo dentro del proceso de investigación.

En este capítulo se mencionan conceptos con los cuales quizás algunos estudiantes no estén familiarizados, tales como hipótesis, teoría, análisis estadístico, entrevistas, etcétera. Dichos conceptos se definen y explican en el transcurso del libro; por ello, los lectores no deben preocuparse ahora por su cabal entendimiento, lo importante es que queden claros los mensajes esenciales respecto de los enfoques cuantitativo y cualitativo, así como su proceder general.

¿QUÉ ENFOQUES SE HAN PRESENTADO PARA LA INVESTIGACIÓN?

A lo largo de la Historia de la Ciencia han surgido diversas corrientes de pensamiento tales como el Empirismo, el Materialismo Dialéctico, el Positivismo, la Fenomenología y el Estructuralismo, las cuales han originado diferentes rutas en la búsqueda del conocimiento. No se profundizará por ahora en ellas, ya que se tratan de manera extensa en antologías y textos sobre Sociología.¹ Sin embargo, y debido a las diferentes premisas que las sustentan,² desde la segunda mitad del siglo XX tales corrientes se han polarizado en dos enfoques principales: el enfoque cuantitativo y el enfoque cualitativo de la investigación.

A continuación comentaremos brevemente cada uno y después propondremos esquemas para comprender su inserción en el proceso de investigación e, incluso, para visualizar que pueden formar parte de un mismo estudio o de una misma aplicación de dicho proceso, lo cual se denomina enfoque integrado "multimodal". Sabemos que enfrentará escepticismo entre algunos colegas, en

¹ Horton, Paul B. y Chester L. Hunt, *Sociology*. McGraw-Hill, N.Y., 1985, y Coser, Lewis A. y Rosenberg, Bernard, *Sociological Theory: A Book of Readings*, Waveland Press, N.Y., 1994.

² El enfoque cuantitativo en las ciencias sociales se origina en la obra de Auguste Comte (1798-1857) y Emile Durkheim (1858-1917). Ellos proponen que el estudio sobre los fenómenos sociales requiere ser "científico"; es decir, susceptible a la aplicación del mismo método científico que se utilizaba con considerable éxito en las ciencias naturales. Sostenían que todas las cosas o fenómenos pueden medirse. A esta corriente se le llamó *Positivismo*.

El enfoque cualitativo tiene su origen en otro pionero de las ciencias sociales, Max Weber (1864-1920), quien introduce el término "verstehen" o "entendimiento", reconociendo que además de la descripción y medición de variables sociales deben considerarse los significados subjetivos y el entendimiento del contexto donde ocurre un fenómeno. Weber propone un método híbrido, con herramientas como los tipos ideales, en donde los estudios no sean únicamente de variables macrosociales, sino de instancias individuales.

especial entre quienes se muestran radicales ante alguna de estas posturas. No obstante, desde hace varios años creemos firmemente que ambos enfoques, utilizados en conjunto, enriquecen la investigación. No se excluyen, ni se sustituyen. Nuestra posición es incluyente y en toda América Latina quienes han compartido experiencias con nosotros han sido testigos de ello.

En términos generales, los dos enfoques (**cuantitativo** y **cualitativo**) utilizan cinco fases similares y relacionadas entre sí (Grinnell, 1997):

- a) Llevan a cabo observación y evaluación de fenómenos.
- b) Establecen suposiciones o ideas como consecuencia de la observación y evaluación realizadas.
- c) Prueban y demuestran el grado en que las suposiciones o ideas tienen fundamento.
- d) Revisan tales suposiciones o ideas sobre la base de las pruebas o del análisis.
- e) Proponen nuevas observaciones y evaluaciones para esclarecer, modificar, cimentar y/o fundamentar las suposiciones e ideas; o incluso para generar otras.

Así, el investigador de organizaciones buscará observar y evaluar aspectos de las empresas o instituciones, tales como el grado de satisfacción en los obreros. El investigador en Derecho Fiscal hará lo mismo con los fenómenos tributarios e intentará explicar la captación de impuestos en épocas de crisis. El investigador en Ingeniería de Construcción —por ejemplo— observará y/o evaluará los nuevos materiales para estructuras. El investigador en Ciencias de la Comunicación aplicará dichas fases para conocer más los fenómenos comunicativos, como el surgimiento de rumores cuando una fuente emite mensajes contradictorios. Sin embargo, aunque ambos enfoques comparten esas etapas generales, cada uno tiene sus propias características.

El **enfoque cuantitativo** utiliza la recolección y el análisis de datos para contestar preguntas de investigación y probar hipótesis establecidas previamente, y confía en la medición numérica, el conteo y frecuentemente en el uso de la estadística para establecer con exactitud patrones de comportamiento en una población.

El **enfoque cualitativo**, por lo común, se utiliza primero para descubrir y refinar preguntas de investigación. A veces, pero no necesariamente, se prueban hipótesis (Grinnell, 1997). Con frecuencia se basa en métodos de recolección de datos sin medición numérica, como las descripciones y las observaciones. Por lo regular, las preguntas e hipótesis surgen como parte del proceso de investigación y éste es flexible, y se mueve entre los eventos y su interpretación, entre las respuestas y el desarrollo de la teoría. Su propósito consiste en “reconstruir” la realidad, tal y como la observan los actores de un sistema social previamente definido.

A menudo se llama "holístico", porque se precia de considerar el "todo",³ sin reducirlo al estudio de sus partes.

Enfoque cuantitativo: usa recolección de datos para probar hipótesis con base en la medición numérica y el análisis estadístico para establecer patrones de comportamiento.

Desde nuestro punto de vista, ambos enfoques resultan muy valiosos y han realizado notables aportaciones al avance del conocimiento. Ninguno es intrínsecamente mejor que el otro, sólo constituyen diferentes aproximaciones al estudio de un fenómeno. Pensamos que la controversia entre las dos visiones ha sido innecesaria y no está exenta de dogmatismo. La posición asumida en esta obra es que son enfoques complementarios; es decir, cada uno sirve a una función específica

para conocer un fenómeno, y para conducirnos a la solución de los diversos problemas y cuestionamientos. El investigador debe ser metodológicamente plural, y guiarse por el contexto, la situación, los recursos de que dispone, sus objetivos y el problema de estudio. En efecto, se trata de una postura pragmática.

A continuación ofreceremos ejemplos de investigaciones que, utilizando uno u otro enfoque, se dirigieron al mismo objeto-sujeto de estudio.

En lo fundamental, ¿qué características destacan en el enfoque cuantitativo de la investigación? En términos generales, un estudio cuantitativo regularmente elige una idea, que transforma en una o varias preguntas de investigación relevantes (capítulo 3); luego de éstas deriva hipótesis y variables; desarrolla un plan para probarlas; mide las variables en un determinado contexto; analiza las mediciones obtenidas (con frecuencia utilizando métodos estadísticos), y establece una serie de conclusiones respecto de la(s) hipótesis.

Enfoque cualitativo: utiliza recolección de datos sin medición numérica para descubrir o afinar preguntas de investigación y puede o no probar hipótesis en su proceso de interpretación.

Si nos fijamos en el cuadro 1.1, los estudios cuantitativos proponen relaciones entre variables con la finalidad de arribar a proposiciones precisas y hacer recomendaciones. Por ejemplo, la investigación de comunicación en que Rogers y Waisanen (1969) proponen que la comunicación interpersonal resulta más eficaz que, la comunicación de medios en las sociedades rurales. Se espera que, en los estudios cuantitativos, los investigadores elaboren un reporte con sus resultados y ofrezcan recomendaciones que servirán para la

solución de problemas o en la toma de decisiones.

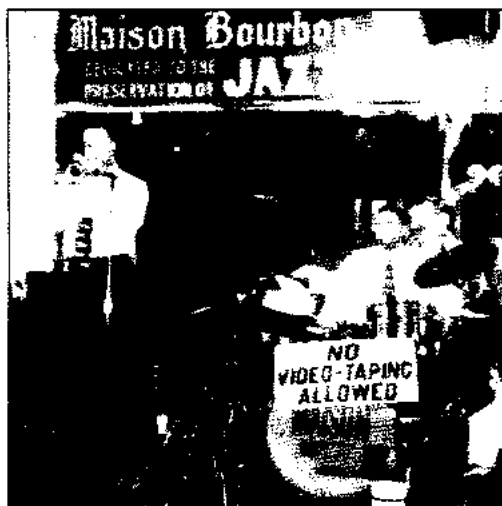
Esencialmente, ¿qué características destacan en el enfoque cualitativo a la investigación? Las investigaciones cualitativas también son guiadas por áreas o

³ Aquí el "todo" es el fenómeno de interés. Por ejemplo, en su libro *Police Work*, Peter Manning (1997), se sumerge por semanas en el estudio, la información y el análisis del trabajo policíaco. Le interesa estudiar las relaciones y lealtades que surgen entre sujetos que se dedican a este trabajo. Lo logran sin "medición" de actitudes: tan sólo captando el fenómeno mismo de la vida en este trabajo.

TEMA-OBJETO DE ESTUDIO	ESTUDIOS CUALITATIVOS	ESTUDIOS CUANTITATIVOS
La familia	Gabriel Careaga (1977): <i>Mitos y fantasías de la clase media en México.</i>	Ma. Elena Oto Mishima (1994): <i>Las migraciones a México y la conformación paulatina de la familia mexicana.</i>
Alcance del estudio	El libro es una aproximación crítica y teórica al surgimiento de la clase media en un país poco desarrollado. El autor combina los análisis documental, político, dialéctico y psicoanalítico con la investigación social y biográfica para reconstruir tipologías o familias tipo.	Descripción de la procedencia de los inmigrantes a México; su integración económica y social a diferentes esferas de la sociedad.
La comunidad	Luis González y González (1995): <i>Pueblo en vilo.</i>	Everett Rogers y Frederick B. Waisanen (1969): <i>The impact of communication on rural development.</i>
Alcance del estudio	El autor describe con detalle la microhistoria de San José de la Gracia, donde se examinan y entrelazan las vidas de sus pobladores con su pasado y otros aspectos de la vida cotidiana.	Se determina cómo ocurre el proceso de comunicación de innovaciones en comunidades rurales, y se identifican los motivos para aceptar o rechazar el cambio social. Asimismo, se establece qué método de comunicación es el más beneficioso.
Las ocupaciones	Howard Becker (1951): <i>The Professional dance musician and his audience.</i>	Linda D. Hammond (2000): <i>Teacher quality and student achievement.</i>
Alcance del estudio	Narración detallada de procesos de identificación y otras conductas de músicos de jazz con base en sus competencias y su conocimiento de la música.	Establece correlaciones entre estilos de enseñanza, desempeño de la ocupación docente y éxito de los alumnos.
Organizaciones de trabajo	William D. Bygrave y Dan D'Heilly (editores) (1997): <i>The Portable MBA Entrepreneurship Case Studies.</i>	P. Marcus, P. Baptista y P. Brandt (1979): <i>Rural Delivery Systems.</i>
Alcance del estudio	Compendio de estudios de caso que apoyan el análisis sobre la viabilidad de nuevas empresas y los retos que enfrentan en los mercados emergentes.	Investigación que demuestra la poca coordinación que existe en una red de servicios sociales. Establece las políticas a seguir para lograr que los servicios lleguen a los destinatarios.
El fenómeno urbano	Manuel Castells (1979): <i>The Urban Question.</i>	E. Wirth (1964): <i>¿Cuáles son las variables que afectan la vida social en la ciudad?</i>
Alcance del estudio	El autor critica lo que tradicionalmente estudia el urbanismo, y argumenta que la ciudad no es más que un espacio donde se expresan y manifiestan las relaciones de explotación.	La densidad de población y la escasez de vivienda se establecen como influyentes en el descontento político.

Ejemplos de estudios cualitativos y cuantitativos sobre un mismo tema

Cuadro 1.1



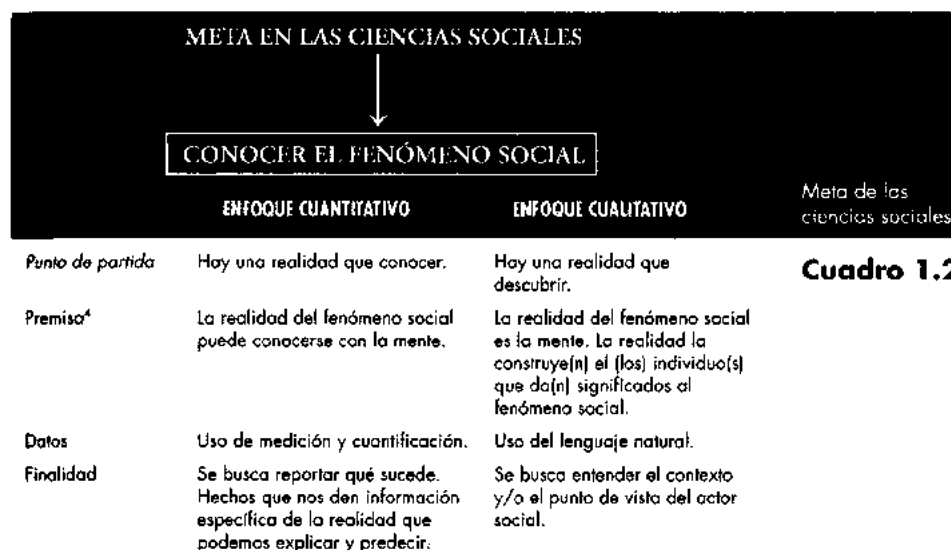
En una investigación cualitativa, lo importante es comprender el fenómeno, como por ejemplo las normas y condiciones propias que rigen la profesión de un músico de jazz.

■ **Modelo multimodal (triangulación):** convergencia o fusión de los enfoques de investigación cuantitativo y cualitativo.

temas significativos de investigación. Sin embargo, en lugar de que la claridad sobre la(s) pregunta(s) de investigación e hipótesis preceda (como en la mayoría de los estudios cuantitativos, al menos en intención) a la recolección y el análisis de los datos, los estudios cualitativos pueden desarrollar preguntas e hipótesis antes, durante o después de la recolección y el análisis. Con frecuencia, estas actividades sirven, primero, para descubrir cuáles son las preguntas de investigación más importantes; y, después, para refinarlas y responderlas (o probar hipótesis). El proceso se mueve dinámicamente entre los "hechos" y su interpretación en ambos sentidos. Como se observa en el cuadro anterior, su alcance final muchas veces consiste en comprender un fenómeno social complejo. El énfasis no está en medir las variables involucradas en dicho fenómeno, sino en entenderlo.

Tomando como ejemplo el estudio de las ocupaciones y sus efectos en la conducta individual, en el cuadro anterior notamos la divergencia a la que nos referimos. En el clásico estudio de Howard Becker (1951) sobre el músico de jazz, el autor logra que comprendamos las reglas y los ritos en el desempeño de esta profesión. ¿Y la utilidad de su alcance?, se preguntarán algunos. Pues no está solamente en comprender ese contexto, sino en que las normas que lo rigen se pueden extrapolar a otras situaciones de trabajo. Por otro lado, el estudio cuantitativo de Hammond (2000) trata de establecer con claridad variables personales y del desempeño de la profesión docente, que sirvan para formular políticas de contratación y de capacitación para el personal docente. ¿Para qué? Con la finalidad última de incrementar el éxito académico en los estudiantes.

Lo principal hasta ahora es que el lector se abstenga de evaluar si un enfoque es mejor que el otro. En cambio, necesita comprender, hasta el momento, que tradicionalmente se han aproximado de distinta manera al estudio de diversas cuestiones. Dicha aproximación parte de diferencias en cuanto a lo epistemológico (o teoría del conocimiento), que en pocas palabras significa qué postura toma un investigador o toma cada quien ante la realidad. El conocimiento de lo social es la meta de las ciencias sociales en general; pero en lo particular se presentan divergencias que se sintetizan en el cuadro 1.2.



Cuadro 1.2

¿Qué otras características poseen ambos enfoques y cómo se diferencian?

El enfoque cualitativo busca principalmente “dispersión o expansión” de los datos o información; mientras que el cuantitativo pretende intencionalmente “acotar” la información (medir con precisión las variables del estudio, tener “foco”).⁵

⁴ Becker (1993) dice: “la ‘realidad’ es el punto más estresante en las ciencias sociales. Las diferencias entre los dos enfoques han tenido un tinte eminentemente ideológico. El gran filósofo alemán Karl Popper (1965) nos hace entender que el origen de visiones conflictivas, sobre lo que es o debe ser el estudio del fenómeno social, se encuentran desde las premisas de diferentes definiciones de lo que es la realidad. El realismo desde Aristóteles establece que el mundo llega a ser conocido por la mente. Kant introduce que el mundo puede ser conocido porque la realidad se asemeja a las formas que la mente tiene. En tanto que Hegel va hacia un idealismo puro y propone: “El mundo es mi mente”. Esto último es ciertamente confuso, y así lo considera Popper, advirtiendo que el gran peligro de esta posición es que permite el dogmatismo (como lo ha probado, por ejemplo, el materialismo dialéctico). El avance en el conocimiento, dice Popper, necesita de conceptos que podamos refutar o probar. Esta característica delimita qué es y qué no es ciencia.

⁵ Usemos el ejemplo de una cámara fotográfica: en el estudio cuantitativo se define lo que se va a fotografiar y se toma la foto. En el cualitativo es como si la función de “zoom in” (acercamiento) o “zoom out” (alejamiento) se utilizaran constantemente para capturar en un área cualquier figura de interés.

De acuerdo con M. A. Rothery (citado por Grinnell, 1997), para generar conocimiento el enfoque **cuantitativo** se fundamenta en el método hipotético-deductivo, considerando las siguientes premisas:

1. Delineamos teorías y de ellas derivamos hipótesis.
2. Las hipótesis se someten a prueba utilizando los diseños de investigación apropiados.
3. Si los resultados corroboran las hipótesis o son consistentes con éstas, se aporta evidencia en su favor. Si se refutan, se descartan en busca de mejores explicaciones e hipótesis.

Cuando los resultados de diversas investigaciones aportan evidencia en favor de las hipótesis, se genera confianza en la teoría que las sustenta o apoya. Si no es así, se descartan la hipótesis y, eventualmente, la teoría.

Además de estas premisas, se toman en cuenta otras consideraciones en la investigación *cuantitativa*. Grinnell (1997) y Creswell (1997) señalan que no se desechan la *realidad subjetiva* ni las experiencias individuales. Asimismo,

1. Hay dos realidades: “la primera” consiste en las creencias, presuposiciones y experiencias subjetivas de las personas. Éstas llegan a variar: desde ser muy vagas o generales (intuiciones) hasta ser creencias bien organizadas y desarrolladas lógicamente a través de teorías formales. “La segunda realidad” es objetiva e independiente de las creencias que tengamos hacia ella (la autoestima, una ley, los mensajes televisivos, una presa, el SIDA, etcétera, ocurren, es decir, constituyen realidades en forma independiente de lo que pensemos de ellas).
2. Esta “realidad objetiva” (o realidades) es susceptible de conocerse. Bajo esta premisa, resulta posible conocer una realidad externa e independiente del sujeto.
3. La “realidad objetiva” es necesario conocerla o tener la mayor cantidad de información sobre ella. La realidad del fenómeno existe, y sí, a los eventos que nos rodean los conocemos a través de sus manifestaciones. Para entender nuestra realidad, el porqué de las cosas, hay que registrar y analizar dichos eventos (Lesser, 1935). Desde luego, para este enfoque, lo subjetivo existe y posee un valor para los investigadores; pero de alguna manera se aboca a demostrar qué tan bien se adecua a la **realidad objetiva**. Documentar esta coincidencia constituye un propósito central de muchos **estudios cuantitativos** (que los efectos que consideramos que tiene una enfermedad sean verdaderos, que capturemos la relación “real” entre las motivaciones de un sujeto y su conducta, que un material que se supone posea una determinada resistencia auténticamente la tenga, etcétera).

4. Cuando las investigaciones creíbles establezcan que la “realidad objetiva” es diferente de nuestras creencias, éstas deben modificarse o adaptarse a la primera.

Para este enfoque, la forma confiable para conocer la **realidad** es a través de la recolección y el análisis de datos, de acuerdo con ciertas reglas lógicas. Si éstas se siguen cuidadosamente y los datos generados poseen los estándares de validez y confiabilidad (capítulo 10), las conclusiones derivadas tendrán validez; es decir, la posibilidad de ser refutadas, o de replicarse con la finalidad de ir construyendo conocimiento.

Por lo común, en los **estudios cuantitativos** se establece una o varias hipótesis (suposiciones acerca de una **realidad**), se diseña un plan para someterlas a prueba, se miden los conceptos incluidos en la(s) hipótesis (variables) y se transforman las mediciones en valores numéricos (datos cuantificables), para analizarse posteriormente con técnicas estadísticas y extender los resultados a un universo más amplio, o para consolidar las creencias (formuladas en forma lógica en una teoría o un esquema teórico).⁶

Tales estudios llevan la esencia en su título: cuantificar y aportar evidencia a una teoría que se tiene para explicar algo; la teoría se mantiene hasta que se refute o se alcance una mejor explicación. Un estudio se basa en otro. Los **estudios cuantitativos** se asocian con los experimentos, las encuestas con preguntas cerradas o los estudios que emplean instrumentos de medición estandarizados. Además, en la interpretación de los estudios hay una humildad que deja todo inconcluso e invita a seguir investigando y mejorar el conocimiento, poniendo a disposición de otros investigadores todos los métodos y los procedimientos.

Por su parte, el **enfoque cualitativo**, a veces referido como investigación naturalista, fenomenológica, interpretativa o etnográfica, es una especie de “paraguas”, en el cual se incluye una variedad de técnicas y estudios **no cuantitativos** (Grinell, 1997).

Dentro de la variedad de enfoques cualitativos existe un común denominador que podríamos situar en el concepto de *patrón cultural* (Colby, 1996), que parte de la premisa de que toda cultura o sistema social tiene un modo único para entender cosas y eventos. Esta cosmovisión, o manera de ver el mundo, afecta la conducta humana. El estudio de los modelos culturales —que son marcos de referencia para el actor social y que están contruidos por el inconsciente, lo transmitido por otros y la experiencia personal— son entidades flexibles y maleables que se tornan en el objeto de estudio de lo cualitativo.

⁶ Los investigadores que se guiaron por este enfoque metodológico pretendían expandir el alcance de los estudios sociales más allá de lo que consideraban aproximaciones meramente especulativas.

En términos generales, los **estudios cualitativos** involucran la recolección de datos utilizando técnicas que **no** pretenden medir ni asociar las mediciones con números, tales como observación no estructurada, entrevistas abiertas, revisión de documentos, discusión en grupo, evaluación de experiencias personales, inspección de historias de vida, análisis semántico y de discursos cotidianos, interacción con grupos o comunidades, e introspección.

M. A. Rothery y R. Grinnell (Grinnell, 1997), y Creswell (1997) describen estas investigaciones como estudios:

- Que se conducen básicamente en ambientes naturales, donde los participantes se comportan como lo hacen en su vida cotidiana.
- Donde las variables no se definen con el propósito de manipularse ni de controlarse experimentalmente (desde luego, se observan los cambios en diferentes variables y sus relaciones).
- En los cuales las preguntas de investigación no siempre se han conceptualizado ni definido por completo, es decir, en la manera como van a medirse o evaluarse (aunque a veces sí es posible).
- En que la recolección de los datos está fuertemente influida por las experiencias y las prioridades de los participantes en la investigación, más que por la aplicación de un instrumento de medición estandarizado, estructurado y predeterminado.
- Donde los significados se extraen de los datos y se presentan a otros, y no necesitan reducirse a números ni necesariamente deben analizarse de forma estadística (aunque el conteo, el análisis de contenido y el tratamiento de la información utilicen expresiones numéricas para analizarse después).

Parton (1980, 1990) define los datos cualitativos como descripciones detalladas de situaciones, eventos, personas, interacciones, conductas observadas y sus manifestaciones.

Un **estudio cualitativo** busca comprender su fenómeno de estudio en su ambiente usual (cómo vive, se comporta y actúa la gente; qué piensa; cuáles son sus actitudes, etcétera).

Neuman (1994) sintetiza las actividades principales del **investigador cualitativo** con los siguientes comentarios:

- El investigador observa eventos ordinarios y actividades cotidianas tal y como suceden en sus ambientes naturales, además de cualquier acontecimiento inusual.
- Está directamente involucrado con las personas que se estudian y con sus experiencias personales.

- Adquiere un punto de vista “interno” (desde adentro del fenómeno), aunque mantiene una perspectiva analítica o una distancia específica como observador externo.
- Utiliza diversas técnicas de investigación y habilidades sociales de una manera flexible, de acuerdo con los requerimientos de la situación.
- Produce datos en forma de notas extensas, diagramas, mapas o “cuadros humanos” para generar descripciones bastante detalladas.
- Sigue una perspectiva holística (los fenómenos se conciben como un “todo” y no como partes) e individual.
- Entiende a los miembros que son estudiados y desarrolla empatía hacia ellos; no solamente registra hechos objetivos “fríos”.
- Mantiene una doble perspectiva: analiza los aspectos explícitos, conscientes y manifiestos, así como aquellos implícitos, inconscientes y subyacentes. En este sentido, la **realidad subjetiva** en sí misma es objeto de estudio.
- Observa los procesos sin irrumpir, alterar o imponer un punto de vista externo, sino tal y como son percibidos por los actores del sistema social.
- Es capaz de manejar paradoja, incertidumbre, dilemas éticos y ambigüedad.

Los **estudios cualitativos** no pretenden generalizar de manera intrínseca los resultados a poblaciones más amplias, ni necesariamente obtener muestras representativas (bajo la ley de probabilidad); incluso, no buscan que sus estudios lleguen a replicarse. Asimismo, se fundamentan más en un proceso inductivo (exploran y describen, y luego generan perspectivas teóricas). Van de lo particular a lo general.

Durante varias décadas se ha considerado que los **enfoques cuantitativo y cualitativo** son perspectivas opuestas, irreconciliables y que no deben mezclarse. Los críticos del **enfoque cuantitativo** lo acusan de ser “impersonal, frío, limitativo, cerrado y rígido”. Por su parte, los detractores del **enfoque cualitativo** lo consideran “vago, subjetivo, inválido, meramente especulativo, sin posibilidad de réplica y sin datos sólidos que apoyen las conclusiones”.

La base del divorcio entre ambos enfoques se ha centrado en la idea de que un estudio con un enfoque puede neutralizar al otro. Se trata de una noción que ha impedido la reunión de los enfoques **cuantitativo y cualitativo**. Superar una conceptualización, a nuestro juicio “fundamentalista”, lleva a concebir la unión de ambos enfoques, a lo cual Denzin (1978) denomina “triangulación”.

Si revisamos los estudios científicos realizados en los últimos años (buscando en las revistas producidas en distintas disciplinas y campos), observaremos una tendencia creciente en este sentido: la fusión, el matrimonio **cuanti-cuali**. Ello se debe tal vez a que nos hemos dado cuenta de que, más que beneficiarlo, las luchas ideológicas y las posiciones dogmáticas han impedido el avance del conocimiento; por ello, es menester buscar la convergencia o la triangulación.

La triangulación es complementaria en el sentido de que traslapa enfoques y en una misma investigación mezcla diferentes facetas del fenómeno de estudio. Dicha unión o integración añade profundidad a un estudio y, aunque llegan a surgir contradicciones entre los resultados de ambos enfoques, se agrega una perspectiva más completa de lo que estamos investigando.

Ante la oportunidad de fusionar ambos enfoques, Grinnell (1997) se formula una serie de cuestionamientos: ¿los paradigmas inductivo y deductivo deben estar vinculados con enfoques específicos? Por ejemplo, si empleamos un esquema inductivo, basándonos en una **postura cualitativa** para un estudio, ¿esto significa que también debemos utilizar procedimientos de recolección de datos usualmente asociados con investigaciones cualitativas?

Alternativamente, ¿un estudio basado en un esquema deductivo y guiado por una teoría producto de **investigación cuantitativa** siempre tendrá que vincularse con procedimientos de recolección de datos y diseños ligados a tal tipo de investigación, como los experimentos y las encuestas? Las respuestas no son sencillas. Los “puristas” demandan el divorcio entre los **enfoques cuantitativo y cualitativo**, como si el viejo enemigo del positivismo atacara de nuevo, aduciendo que aquello llamado **objetividad** no existe. Por su parte, los “situacionales” aseguran que cada enfoque resulta apropiado para situaciones específicas; los “pragmáticos” integran ambos enfoques, sobre todo cuando es apropiado en situaciones concretas. Esta obra se adhiere a la última posición.

Creemos que debe hacerse más énfasis en las bondades que en las limitaciones de cada enfoque; en todo caso, una situación de investigación particular nos dirá si debemos utilizar un enfoque u otro, o bien, ambos.

Por otro lado, es necesario aclarar que el **enfoque** seleccionado (**cuantitativo, cualitativo o algún tipo de mezcla entre ambos**) no necesariamente tiene que ver con los métodos de recolección de datos. Por ejemplo, un experimento clásico (definido como **cuantitativo**) puede utilizar métodos de recolección de datos tanto cualitativos como cuantitativos (aplicar como preprueba y posprueba un cuestionarios cerrado, estructurado, y entrevistas abiertas). O una **investigación cualitativa** (un estudio que trate de entender los sentimientos de enfermos terminales de SIDA) puede llegar a recolectar datos a través de entrevistas abiertas y de aplicar una prueba estructurada del sentido de vida.

Nuestra sugerencia a los estudiantes sería que conozcan ambos métodos, los dominen, reflexionen sobre sus ventajas y limitaciones, para así decidir qué enfoque o mezcla es más útil para cada caso.

Para que el lector que se inicia en estos menesteres tenga una idea de la diferencia entre ambos enfoques, utilizaremos un ejemplo muy sencillo y cotidiano relativo a la atracción física, explicado para su entendimiento, aunque a algunos profesores podría parecerles simple.

*Un ejemplo
para ayudar
a comprender
los enfoques
cuantitativo
y cualitativo
de la
investigación*

Supongamos que un(a) estudiante se encuentra interesado(a) en saber qué factores influyen en que una persona sea definida y percibida como “conquistadora” (que cautiva a individuos del sexo opuesto y logra que se sientan atraídos hacia él o ella y se enamoren). Entonces, decide llevar a cabo un estudio (su idea para investigar) en su escuela.

Bajo el enfoque cuantitativo-deductivo, el estudiante plantea su problema de estudio definiendo su objetivo y su pregunta de investigación (lo que quiere hacer y lo que quiere saber).

Por ejemplo, el objetivo podría ser “conocer los factores que determinan el hecho de que una persona sea percibida como atractiva y conquistadora”; y la pregunta de investigación, ¿qué factores determinan el hecho de que una persona sea percibida como atractiva y “conquistadora”?

Posteriormente, revisaría estudios sobre la atracción física en las relaciones heterosexuales, los elementos que intervienen en el inicio de la convivencia amorosa, la percepción de los jóvenes en torno a dichas relaciones, las diferencias por sexo de acuerdo con los atributos que les atraen de los demás, etcétera.

Precisaría su problema de investigación; seleccionaría una teoría que explicara satisfactoriamente —sobre la base de estudios previos— la atracción física y el enamoramiento en las relaciones entre jóvenes; y, de ser posible, establecería una o varias hipótesis (por ejemplo: “los(as) chicos(as) que logran más conquistas amorosas son aquellos(as) que tienen mayor prestigio social en la escuela, que son más seguros(as) de sí mismos(as) y más extrovertidos(as)”).

Después, podría entrevistar a compañeras y compañeros de su escuela y los interrogaría sobre el grado en que el prestigio social, la seguridad en uno mismo y la extroversión influyen en la conquista de personas del otro sexo. Incluso, llegaría a utilizar cuestionarios ya establecidos, bien diseñados y confiables. Tal vez entrevistaría solamente a una muestra de estudiantes. También sería posible preguntarles a las personas que tienen reputación como conquistadoras sobre qué piensan al respecto.

Y analizaría los datos e información producto de las entrevistas para obtener conclusiones acerca de sus hipótesis. Quizá también experimentaría eligiendo a personas que tuvieran un perfil “conquistador”, lanzándolas a conquistar a otros(as) chicos(as).

Su interés sería generalizar sus resultados, al menos, a lo que ocurre en su comunidad estudiantil. Busca probar sus creencias y si resulta que no consigue demostrar que el prestigio, la seguridad en sí mismo y la extroversión son factores relacionados con la conquista, intentaría otras explicaciones; tal vez agregando otros factores como dimensiones de la atracción física, la manera en que se visten, etcétera.

En el proceso va deduciendo de la teoría a lo que encuentra en su estudio. Desde luego, si la teoría que seleccionó es inadecuada, sus resultados serán pobres.

Bajo el enfoque cualitativo-inductivo, el estudiante, más que plantear el problema de investigación, lo que haría sería sentarse en la cafetería a observar a personas que tienen fama de ser conquistadoras, tomando en cuenta sus características, la manera en que se comportan, sus atributos y la forma de relacionarse con los demás (en particular,

con personas del sexo complementario). De ahí, podría derivar algún esquema que explique las razones por las cuales estas personas conquistan a otras.

Después entrevistaría con preguntas abiertas a algunas de éstas y también a quienes han sido conquistadas(os) por ellos(as). De ahí derivaría conclusiones y contrastaría sus hallazgos con los de otros estudios. No sería indispensable obtener una muestra representativa ni generalizar sus resultados.

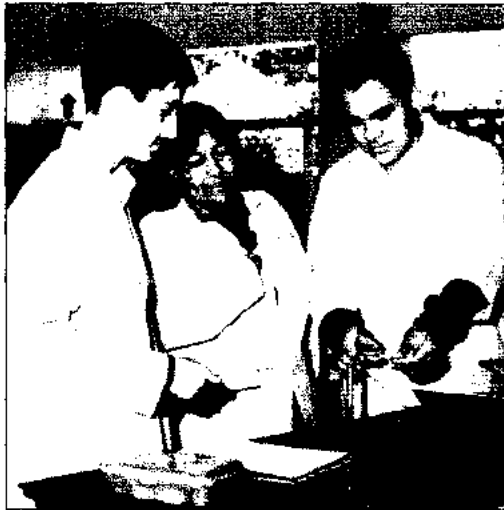
Si proceder sería inductivo: de cada caso de conquistador(a) estudiado obtendría quizás el perfil buscado.

También podría mezclar ambos enfoques y proceder simultáneamente de las dos maneras (encargarle a un grupo de amigos que realicen el primer estudio, basado en la teoría producto de investigaciones previas; y a otro grupo, que inicien observando a los conquistadores en la cafetería (enfoque mixto cuantitativo-cualitativo).

Los métodos cuantitativos han sido los más utilizados por los investigadores de las llamadas "ciencias exactas".

Desde luego, en el ejemplo no se consideran las implicaciones paradigmáticas que se encuentran detrás de cada enfoque; pero sí se enfatiza que en términos prácticos ambos estudios contribuyen al conocimiento de un fenómeno.

EL PROCESO DE INVESTIGACIÓN



Desde la primera edición de este libro se ha insistido en la premisa de concebir la investigación como un proceso constituido por diversas etapas, pasos o fases, interconectadas de una manera lógica, secuencial y dinámica. Ello no implica que no sea posible regresar a una etapa previa o visualizar las etapas subsiguientes.

En términos generales, este proceso se aplica tanto al **enfoque cuantitativo** como al **cualitativo**, desde luego, con sus diferencias, las cuales se ampliarán y analizarán en cada etapa dentro del libro.

Por ahora, comentaremos que, en el caso de la mayoría de los **estudios cuantitativos**, el proceso se aplica secuencialmente: se comienza con una idea que va acotándose y, una vez delimitada, se establecen objetivos y preguntas de investigación, se revisa la literatura y se construye un marco o una perspectiva teórica. Después se ana-

lizan objetivos y preguntas, cuyas respuestas tentativas se traducen (no siempre) en hipótesis; se elabora o selecciona un plan para probar las hipótesis (diseño de investigación) y se determina una muestra. Por último, se recolectan datos utilizando uno o más instrumentos de medición, los cuales se estudian (la mayoría de las veces a través del análisis estadístico), y se reportan los resultados.

Cabe señalar que en la recolección de los datos podría involucrarse un instrumento de naturaleza **cuantitativa** como la aplicación de una entrevista abierta.

Por su parte, en las investigaciones **cuantitativas** el proceso no necesariamente se aplica de manera **secuencial** (aunque puede aplicarse de esta forma). En la mayoría de tales estudios la secuencia sería como se muestra en la figura 1.1.

El planteamiento del problema (objetivos del estudio, las preguntas de investigación y la justificación) y las hipótesis consecuentes surgen en cualquier parte del proceso en un estudio **cuantitativo**: desde que la idea se ha desarrollado hasta, incluso, al elaborar el reporte de investigación.

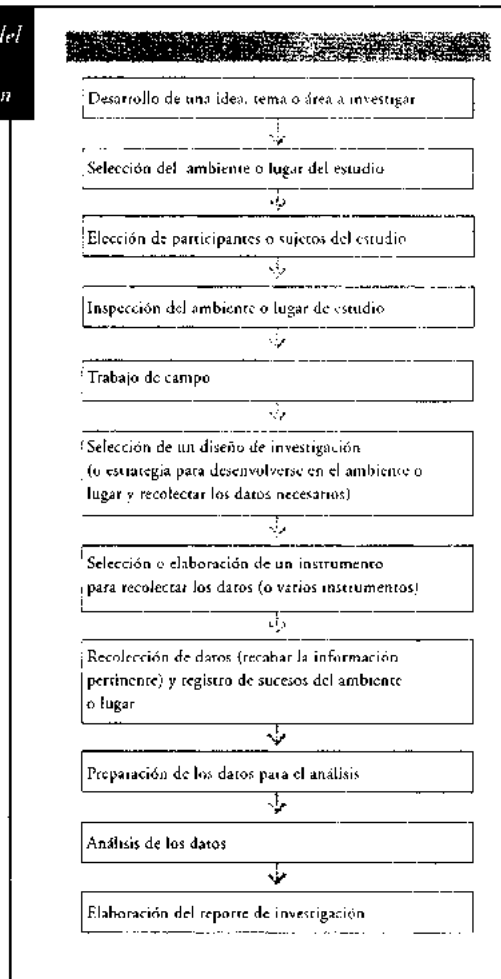
Y, al igual que en la investigación **cuantitativa**, tal planteamiento es susceptible de modificarse, como se menciona a lo largo del libro.

El trabajo de campo significa sensibilizarse con el ambiente o lugar, identificar informantes que aporten datos adicionales, adentrarse y comprenderse con la situación de investigación, además de verificar la factibilidad del estudio.

Aquí las técnicas de recolección de los datos, al igual que en la investigación **cuantitativa**, pueden ser múltiples (entrevistas, pruebas proyectivas, cuestionarios abiertos, sesiones de grupos, análisis de episodios, biografías, casos, grabaciones en audio o video, registros, revisión de archivos, observación, etcétera).

Diagrama del proceso de investigación

Figura 1.1



En la fase de recolección de información se suele llevar una bitácora o un diario personal —escrito o utilizando otras tecnologías—, donde se registran hechos, interpretaciones, creencias y reflexiones sobre el trabajo de campo y la obtención de datos. Asimismo, se anexan documentos y discusiones del equipo de trabajo.

¿Qué bondades tienen cada uno de los enfoques cuantitativo y cualitativo?

Como se ha insistido con anterioridad, ambos resultan fructíferos. La **investigación cuantitativa** nos ofrece la posibilidad de generalizar los resultados más ampliamente, nos otorga control sobre los fenómenos y un punto de vista de conteo y magnitudes de éstos. Asimismo, nos brinda una gran posibilidad de réplica y un enfoque sobre puntos específicos de tales fenómenos, además de que facilita la comparación entre estudios similares.

Por su parte, la **investigación cualitativa** da profundidad a los datos, la dispersión, la riqueza interpretativa, la contextualización del ambiente o entorno, los detalles y las experiencias únicas. También aporta un punto de vista “fresco, natural y holístico” de los fenómenos, así como flexibilidad.

Por ello, la mezcla de los dos modelos potencia el desarrollo del conocimiento, la construcción de teorías y la resolución de problemas. Ambos son empíricos, porque recogen datos del fenómeno que estudian. Tanto el uno como el otro requieren seriedad, profesionalismo y dedicación. Emplean procedimientos distintos que es posible utilizar con acierto.

Desde luego, los **métodos cuantitativos** han sido los más usados por ciencias como la Física, Química y Biología. Por ende, son más propios para las ciencias llamadas “exactas”. Los **cualitativos** se han empleado en disciplinas humanísticas como la Antropología, la Etnografía y la Psicología Social. No obstante, ambos tipos de estudio son de utilidad para todos los campos, como lo demostraremos a lo largo de la presente obra. Por ejemplo, un ingeniero civil puede llevar a cabo un estudio para construir un gran edificio. Emplearía estudios cuantitativos y cálculos matemáticos para levantar su edificación, y analizaría datos estadísticos sobre resistencia de materiales y sobre estructuras similares construidas en subsuelos iguales bajo las mismas condiciones. Pero también puede enriquecer el estudio realizando entrevistas a ingenieros muy experimentados.

Un estudioso de los impactos de una devaluación en la economía de un país complementaría sus análisis **cuantitativos** con sesiones en profundidad con expertos y llevaría a cabo un análisis histórico (tanto **cuantitativo** como **cualitativo**) de los hechos.

Un analista de la opinión pública, al investigar sobre los factores que más inciden en la votación para una próxima elección, utilizaría grupos de enfoque con discusión abierta (cualitativos), además de encuestas por muestreo (cuantitativos).

Un médico que indague sobre qué elementos debe tener en cuenta para tratar a pacientes de enfermedades en fase terminal, y lograr que enfrenten su crítica situación de una mejor manera, revisaría la teoría disponible, consultaría investigaciones cuantitativas y cualitativas al respecto para conducir una serie de observaciones estructuradas de la relación médico-paciente en casos terminales (muestreando actos de comunicación y cuantificándolos), y entrevistaría a enfermos y a médicos mediante técnicas cualitativas, organizaría grupos de enfermos para que hablen abiertamente de dicha relación y del trato que desean. Al terminar puede establecer sus conclusiones y obtener preguntas de investigación, hipótesis o áreas de estudio nuevas. Siempre es mejor tener lo mejor de los dos mundos.

¿CÓMO SE UTILIZAN AMBOS ENFOQUES EN UNA MISMA INVESTIGACIÓN O ESTUDIO?

Antes de responder a la interrogante, cabe mencionar que los autores han trabajado con los dos enfoques. El cuadro 1.3 muestra ejemplos de ello.

FASE CUALITATIVA	FASE CUANTITATIVA	Ejemplos de estudios de autores con enfoque integrado
Hernández Sampieri, Roberto: "El sentido de vida de los afectados por un siniestro" (2002)	Entrevistas profundas y no estructuradas con heridos y lisiados que resultaron de explosiones con cohetes	Aplicación de una escala de sentido de vida estructurada (Prueba en Celaya de Carmen Núñez, 2000)
Fernández Collado, Carlos: "Autoinversión en el trabajo" (1982)	En un estudio piloto se codifican las expresiones y frases de los obreros y empleados, cuando se refieren a sus experiencias de trabajo	Resultado de entrevistas, fue la materia prima para construir el cuestionario que se aplicó a 800 sujetos con la finalidad de ubicar el grado de involucramiento con sus trabajos
Baptista Lucio, Pilar: "Percepciones del director de empresa en México" (1986)	Entrevistas profundas y no estructuradas con directores de empresas medianas para establecer sus percepciones del ambiente político y empresarial	Con base en la investigación cualitativa, se establecieron el tipo y la frecuencia de conductas de comunicación que buscaban conocer el medio ambiente para tomar decisiones informadas

Cuadro 1.3

Nau (1995) y Grinnell (1997) plantean diferentes modalidades en las que es posible mezclar los dos enfoques:

El modelo de dos etapas

Aquí primero se aplica un enfoque y luego el otro, de manera relativamente independiente, dentro del mismo estudio. Uno precede al otro y los resultados se presentan de manera independiente o en un solo reporte.

En cada etapa se respetan los métodos inherentes a cada enfoque. Por ejemplo, el caso de un estudio en materia de Derecho Fiscal realizado por Acero (2001) para evaluar la "cultura fiscal en México". El investigador primero hizo un estudio **cuantitativo**, donde analizó estadísticas referentes al pago de impuestos y la evasión tributaria.

Partió de la base de que las cifras disponibles en esta materia serían un indicador confiable del grado en que la ciudadanía poseía una cultura fiscal. Aunado a ello, revisó otros estudios antecedentes sobre conocimiento de la materia tributaria por parte de los contribuyentes, actitudes hacia ésta, problemas en la recaudación y áreas que le permitieran deducir el grado en que estaba desarrollada la cultura fiscal mexicana.

Una vez concluido su **estudio cuantitativo**, procedió a realizar entrevistas con informantes clave acerca de la cultura fiscal (hizo entrevistas dirigidas entre magistrados de los tribunales fiscales, funcionarios gubernamentales, expertos en Derecho Fiscal, asesores en la materia y contribuyentes); pero con preguntas generales y abiertas tales como: ¿qué es la cultura fiscal? ¿Qué elementos la integran? ¿Existe una cultura fiscal en México? Las entrevistas siguieron el **método cualitativo** clásico (observó y registró las respuestas, así como el contexto en que se llevaron a cabo; interpretó; buscó significados; etcétera). Reportó los resultados de sus entrevistas y obtuvo sus conclusiones.

Primero enfocó **cuantitativamente** el estudio y, luego, **cualitativamente**; en su reporte final, incluyó ambas fases del proceso de investigación.

El modelo de enfoque dominante

En este modelo, el estudio se desarrolla bajo la perspectiva de alguno de los dos enfoques, el cual prevalece, y la investigación mantiene un componente del otro enfoque.

Por ejemplo, un estudio sociológico para conocer las secuelas del terrible acto de violación sexual a jóvenes de ambos sexos, realizado en una ciudad de Colombia, Valledupar.

Modelo de dos etapas: dentro de una misma investigación, se aplica primero un enfoque y después el otro, de forma casi independiente, y en cada etapa se siguen las técnicas correspondientes a cada enfoque.

La investigación podría enfocarse **cualitativamente**, utilizando tres herramientas: *a)* entrevistas en profundidad con víctimas de violación sexual; *b)* sesiones de grupo con discusión abierta sobre el tema, también con jóvenes que sufrieron una agresión de este tipo, y *c)* revisión de registros en los juzgados.

El investigador profundiza en las experiencias de las víctimas, sus traumas, el impacto en su percepción del mundo y la vida cotidiana; en fin, en otros temas que pudieran surgir. Comienza su estudio sin preguntas de investigación, mucho menos con hipótesis. Sólo establece tópicos generales para tratar en las entrevistas y en las sesiones, así como puntos a revisar en los expedientes. O si se quiere, completamente abierto con una pregunta general en entrevistas o sesiones relacionadas con el significado general de la experiencia, como detonante de las respuestas y los comentarios.

Se trata de una **investigación cualitativa**, pero a la que podría agregársele un componente **cuantitativo**: administrar una prueba estandarizada para medir la ansiedad en los jóvenes, posterior al difícil evento.

Otro caso sería el de un experimento para evaluar el grado en que un nuevo método para enseñar computación favorece el aprendizaje y la autoestima en los niños de 10 a 12 años de Valencia, Venezuela, estableciendo hipótesis que aseguran que el método incrementará estas dos variables. Es un estudio **cuantitativo**, al que podría agregársele un componente **cualitativo**: sesiones con los pequeños para elaborar mapas cognitivos de lo que aprendieron y registrar sus experiencias.

La ventaja de este modelo, según Grinnell (1997), consiste en que presenta un enfoque que en ningún caso se considera inconsistente y se enriquecen tanto la recolección de los datos como su análisis. La desventaja que los “fundamentalistas”, de uno u otro bandos, le encontrarían es que su enfoque se estaría subutilizando.

Modelo de enfoque dominante: se lleva a cabo bajo la perspectiva de alguno de los enfoques, el cual prevalece, y el estudio conserva componente(s) del otro enfoque.

El modelo mixto

Este modelo representa el más alto grado de integración o combinación entre los enfoques **cualitativo** y **cuantitativo**.

Ambos se entremezclan o combinan en todo el proceso de investigación, o al menos, en la mayoría de sus etapas. Requiere de un manejo completo de los dos enfoques y una mentalidad abierta. Agrega complejidad al diseño de estudio; pero contempla todas las ventajas de cada uno de los enfoques.

La investigación oscila entre los esquemas de pensamiento inductivo y deductivo, además de que por parte del investigador necesita un enorme dinamismo en el proceso. Lleva a un punto de vinculación lo **cualitativo** y lo **cuantitativo**, que suele resultar inaceptable para los “puristas”.

Un ejemplo lo sería un estudio mercadológico realizado por los autores para una cadena de centros comerciales en México y otros localizados en Centroamérica (Comunicometría, S.C., 2000). La investigación tenía como objetivo general conocer lo que los clientes regulares de cada centro comercial pensaban acerca de éste, en términos de si la plaza era funcional, les agradaba y satisfacía sus expectativas. Además, existía un interés por parte de los directivos respecto de si habría la necesidad de efectuar cambios en la decoración, o el giro de las tiendas, y si se requería revertir alguna debilidad.

■ **Modelo mixto:** constituye el mayor nivel de integración entre los enfoques cualitativo y cuantitativo, donde ambos se combinan durante todo el proceso de investigación.

Esta idea de investigación se transformó en una serie de objetivos de investigación de mercados y en preguntas que guiarían la investigación. Algunos ejemplos de los objetivos fueron: *a)* Conocer y analizar las percepciones y conductas que tienen los clientes que asisten regularmente a los centros comerciales, sobre la funcionalidad de la plaza, sus atributos y características; *b)* considerar las sugerencias de cambio en los centros comerciales que proponen los clientes para incrementar su asistencia a la plaza y aumentar el nivel de permanencia o estancia; *c)* llevar a cabo un análisis de fortalezas y debilidades de cada centro comercial ante la óptica de los clientes, y *d)* obtener información que ayude a definir la estrategia de cambio de cada centro comercial.

No fueron todos los objetivos, aunque sirven para ilustrar el ejemplo. Los investigadores no establecieron hipótesis e iniciaron con un estudio dirigido a describir las opiniones de los clientes.

El estudio comprendió dos vertientes o diseños: uno típicamente **cuantitativo** y otro **cualitativo**. Se realizaron ocho investigaciones, una por cada centro comercial. El cuantitativo consistió en una encuesta realizada en una muestra representativa de los clientes que asistían al centro comercial. El tamaño de la muestra se obtuvo de la estadística referente al registro de asistencia (conteo) dentro de una semana típica (no en épocas de gran afluencia como Navidad) y fue de 420 personas (60 por cada día de la semana). Se aplicó un cuestionario con preguntas cerradas y algunas abiertas. De manera paralela, se hicieron sesiones de grupo en profundidad (con una guía de tópicos que abarcaba la forma en que definían al centro comercial, lo que les gustaba o no de él, la evaluación de la mezcla de tiendas, el tiempo de estancia, los cambios que debían efectuarse en la plaza, la evaluación de los servicios que se ofrecían y otras dimensiones). Cada sesión implicó un grupo de ocho personas y las sesiones abarcaron individuos de diferentes segmentos entre los 18 y 60 años (agrupados por edad y se incluyeron grupos de hombres, mujeres y mixtos, en un total de diez sesiones por centro). Todas se grabaron en video, posteriormente se transcribieron y dos equipos de investigación independientes analizaron los datos (uno a partir de los videos y el otro a partir de las transcripciones).

Se utilizó una técnica cualitativa para analizar los datos, además de que el conductor de las sesiones y un observador que estaba detrás de una cámara de Gesell (con vidrio polarizado) llevaron una birácora de cada una y registraron los sucesos que se presentaron durante las sesiones, así como toda la interpretación que es propia del enfoque cualitativo.

Después se llevó a cabo el análisis estadístico con los datos de la encuesta y se aplicó un análisis interpretativo a los datos resultantes de las sesiones.

Al finalizar el estudio del primer centro comercial, surgieron nuevas preguntas de investigación. Una de ellas fue: ¿los centros comerciales deben orientarse a ser, más que lugares de compra, espacios de diversión?

Ello a causa de que el primer estudio arrojó suficiente información para pensar que los centros comerciales, al menos en los países estudiados, se han convertido en lo que eran las plazas públicas o los zócalos de antaño, y la gente va tanto a comprar como a entretenerse. En este sentido, cada vez encontraremos mayores espacios de diversión en ellos.

Así, el estudio para el segundo centro comercial refinó el cuestionario de la encuesta y la guía de tópicos de las sesiones. El resultado fue un estudio que se enriqueció con los dos enfoques, los cuales convivieron dentro del mismo proceso de investigación.

A lo largo de este texto se mencionará otro ejemplo mercadológico que combina ambos enfoques.

- En los últimos años ha surgido una controversia entre dos enfoques para la investigación: el cuantitativo y el cualitativo.
- Los defensores de cada uno argumentan que el suyo es el más apropiado y fructífero para la investigación.
- El enfoque cuantitativo se fundamenta en un esquema deductivo y lógico, busca formular preguntas de investigación e hipótesis para posteriormente probarlas, confía en la medición estandarizada y numérica, utiliza el análisis estadístico, es reduccionista y pretende generalizar los resultados de sus estudios mediante muestras representativas. Además de que parte de la concepción de que existen dos realidades: la del entorno del investigador y la constituida por las creencias de éste; por ende, fija como objetivo lograr que las creencias del investigador se acerquen a la realidad del ambiente.
- Los experimentos y las encuestas basadas en cuestionarios estructurados son ejemplos de investigación centrada en este enfoque.
- El enfoque cualitativo, por su parte, se basa en un esquema inductivo, es expansivo y por lo común no busca generar preguntas de investigación de antemano ni probar hipótesis preconcebidas, sino que éstas surgen durante el desarrollo del estudio. Es individual, no mide numéricamente los fenó-



menos estudiados ni tampoco tiene como finalidad generalizar los resultados de su investigación; no lleva a cabo análisis estadístico; su método de análisis es interpretativo, contextual y etnográfico. Asimismo, se preocupa por capturar experiencias en el lenguaje de los propios individuos y estudia ambientes naturales.

- Las entrevistas abiertas y la observación no estructurada son ejemplos asociados con el enfoque cualitativo.
- Estos dos enfoques son formas que han demostrado ser muy útiles para el desarrollo del conocimiento científico y ninguno es intrínsecamente mejor que el otro.
- Ambos llegan a mezclarse y a incluirse en un mismo estudio, lo cual, lejos de empobrecer la investigación, la enriquece; son visiones complementarias.
- Tanto el enfoque de investigación cualitativo como el cuantitativo, con sus diferencias, tienen cabida en el proceso de investigación científica.
- Es posible seguir, al menos, tres modelos para mezclarlos: 1. el modelo de dos etapas, 2. el modelo de enfoque dominante y 3. el modelo mixto.
- En el modelo de dos etapas se aplica primero un enfoque y luego el otro, de manera independiente, en un mismo estudio; en el modelo de enfoque dominante, una de las modalidades prevalece sobre la otra y se incluye un componente de esta segunda; en el modelo mixto se entremezclan ambos enfoques durante todo el proceso de investigación.



Enfoque cuantitativo	Esquema inductivo	Proceso de investigación
Enfoque cualitativo	Modelo de dos etapas	Recolección de los datos
Enfoque o modelo multimodal	Modelo de enfoque dominante	Triangulación
Esquema deductivo	Modelo mixto	

(Una vez que se haya leído todo el libro)

1. Desarrolle una propuesta de investigación para cada modelo (dos etapas, enfoque dominante y mixto) que en un solo estudio mezcle los enfoques cuantitativo y cualitativo.

Sin embargo, aún quedan investigadores y docentes que gustan de adoptar poses radicales. Se comportan como el niño del martillo, quien, habiendo conocido esta herramienta, toma todo aquello que encuentra a su paso como un clavo, sin la posibilidad de preguntarse si lo que necesita es un serrucho o un desarmador.

CARLOS G. ALONZO BLANQUETO
Profesor-investigador titular
Facultad de Educación
Universidad Autónoma de Yucatán
Mérida, México
Facultad de Odontología
Universidad Autónoma de Nayarit
Nayarit, México

Los estudiantes que se inician en la investigación comienzan planteándose un problema en un contexto general, luego ubican la situación en el contexto nacional y regional para, por último, proyectarlo en el ámbito local, es decir, donde se encuentran académicamente ubicados (campo, laboratorio, salón de clases, etcétera).

En la Universidad de Oriente, en Venezuela, la investigación adquirió relevancia en los últimos años por dos razones: el crecimiento de la planta de profesores y la diversificación de carreras en Ingeniería, área en la cual, por lo general, las investigaciones son cuantitativas-positivistas, con resultados muy satisfactorios.

De igual forma, en el estudio de fenómenos sociales y en ciencias de la salud, el enfoque cualitativo, visto como una teoría de la investigación, presenta grandes avances. Es una herramienta metodológica que se utiliza de manera frecuente en estudios doctorales de Filosofía, Epistemología, Educación y Lingüística, entre otras disciplinas. Las aportaciones de tales estudios se caracterizan por su riqueza en descripción y análisis.

Los enfoques cualitativo y cuantitativo, vistos como teorías filosóficas, son completamente diferentes; sin embargo, como técnicas para el desarrollo de una investigación, pueden mezclarse sobre todo en relación con el análisis y la discusión de resultados.

MARIANELIS SALAZAR DE GÓMEZ
Profesor titular
Escuela de Humanidades
Universidad de Oriente
Anzoátegui, Venezuela

Lectura 2

La metodología de encuesta

Heriberto López Romo

LA ENCUESTA SE HA CONVERTIDO EN una herramienta fundamental para el estudio de las relaciones sociales. Las organizaciones contemporáneas, políticas, económicas o sociales, utilizan esta técnica como un instrumento indispensable para conocer el comportamiento de sus grupos de interés y tomar decisiones sobre ellos.

Debido a su intenso uso y difusión, la encuesta es la representante por excelencia de las técnicas del análisis social. Este panorama la ubica dentro de varias situaciones paradójicas.

En primer lugar, si bien esta situación ha generado un ámbito favorable para la difusión y reflexión sobre ella, también ha producido diversos mitos y confusiones que con frecuencia han llevado a utilizar e interpretar la técnica de manera equivocada.

En segundo lugar, la enorme difusión que se ha hecho de la técnica dentro de las instituciones académicas y en los medios de comunicación al publicar resultados, ha generado una *cultura de encuesta*. Esta cultura, si bien ha abierto enormes posibilidades, también ha permitido que casi cualquier persona se erija como un experto de la encuesta. Para los profesionales de la encuesta es común enfrentar discusiones con personas que por haber tomado un curso de técnicas de investigación o haber seguido las encuestas electorales, se consideran expertos en la materia.

En tercer lugar, al mismo tiempo que existen fanáticos que pretenden erigir a la encuesta como la única técnica de análisis social científicamente válida, existen también los fóbicos que la repudian y la reducen a un instrumento de control ideológico.

En cuarto lugar, mientras una visión contemporánea plantea la relación complementaria de los métodos, los fóbicos y los fanáticos se empeñan en enfrentar las técnicas como si se tratase de formas de conocimiento distintas e irreconciliables.

En quinto lugar, si bien la encuesta ha contribuido significativamente al avance del conocimiento en muy diversas áreas, es paradójico observar cómo en muchas organizaciones la encuesta es usada para otros fines. Algunos de los usos de esta seudoinvestigación son ganar poder y visibilidad en la organización, justificar decisiones ya tomadas, inculpar a la investigación de los resultados, utilizarla como

Tomada de: Heriberto López Romo, "La metodología de la encuesta" en Luis Jesús Galindo Cáceres (coord.), *Técnicas de investigación en sociedad, cultura y comunicación*, México, Pearson, 1998, pp. 33-73.

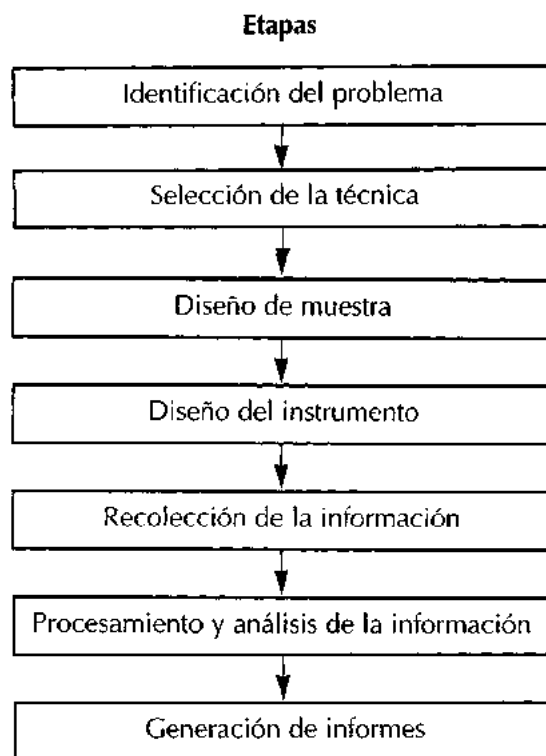
una herramienta en la promoción de los productos y servicios y calmar los ánimos de la impaciente alta gerencia.¹

Por otro lado, independientemente de las paradojas a las que se enfrenta actualmente la encuesta, esta técnica tiene características indiscutibles que explican su creciente demanda y aceptación entre los estudiosos de las diversas caras de la sociedad contemporánea. Algunas de estas características son:

- a) *Variación y flexibilidad de aplicaciones.* La encuesta ha demostrado ser útil en una gran variedad de campos y aplicaciones. La mercadotecnia, la comunicación, las organizaciones, la política, la salud y la demografía utilizan con éxito esta metodología, lo mismo que las universidades y los centros de investigación académica. En los diversos campos donde se usa la encuesta, también se le utiliza en una gran variedad de aplicaciones, que van desde el nivel estratégico, hasta el más táctico.
- b) *Comparabilidad.* El protocolo de la encuesta, por naturaleza sistemático, permite contrastar y comparar los resultados de manera directa y objetiva entre áreas, países y aplicaciones.
- c) *Desarrollo tecnológico.* Si bien los principios de la encuesta siguen siendo los mismos, en los últimos años hemos observado grandes desarrollos tecnológicos. Los sistemas de levantamiento, codificación, captura, validación, procesamiento y presentación de la información se han modificado significativamente e impactado a los destinatarios de la información.
- d) *Oportunidad.* Debido al enorme incremento en el conocimiento de las encuestas, lo mismo que al desarrollo tecnológico y de infraestructuras de investigación por encuesta, los estudios realizados con esta técnica permiten ofrecer resultados con gran velocidad.
- e) *Comprensión.* La significativa difusión de la técnica ha permitido que los usuarios puedan comprender con facilidad el método de trabajo y las implicaciones de los procedimientos y resultados.
- f) *Precisión.* La naturaleza de la técnica permite conocer el nivel de precisión de los resultados y, por lo tanto, el riesgo implícito en las decisiones que se tomarán con base en esos resultados.
- g) *Costo/Eficiencia.* El avance en el conocimiento y en el desarrollo tecnológico, a la par con la experiencia en la delimitación de problemas de investigación, han permitido a la encuesta proporcionar resultados con un alto índice de costo/eficiencia.
- h) *Experiencia profesional organizada.* La gran demanda de la encuesta ha permitido generar una oferta amplia y sólida, con enorme experiencia profesional.

¹ T. Kinnear y James Taylor, *Investigación de mercados*, McGraw Hill, México, 1981.

Este texto tiene como propósito contribuir al conocimiento y difusión de la técnica. De manera sintética se exponen las etapas contempladas en una encuesta. Estas etapas se presentan en el siguiente diagrama:



En este artículo se desarrollará el tema de la encuesta hasta la recolección de la información.

Tipología de estudios por encuesta

Como se señaló en el apartado anterior, independientemente del alto nivel de difusión que tiene actualmente la encuesta, existe un enorme desconocimiento acerca de esta técnica. Una de las principales fuentes al respecto proviene de la confusión en la nomenclatura o forma de referirse a los distintos tipos de encuestas.

Frecuentemente nos referimos a una encuesta como *probabilística*, *de mercado* o *por cuotas*. Esta forma de describir los tipos de encuesta, sin lugar a dudas eficiente en cuanto a comunicación, es simplista e imprecisa en cuanto a su

poder descriptivo. Esta forma de nombrar a las encuestas puede fácilmente distorsionar las implicaciones del estudio.

Con el objeto de tener una visión amplia sobre los alcances y limitaciones de esta técnica, en este apartado se revisarán los principales criterios de clasificación de la encuesta y la relación que existe entre ellos. Esta revisión pretende poner en claro la multiplicidad de criterios que existen para referirse a una encuesta y señala, asimismo, la confusión que puede existir a partir de usar un sólo criterio como definición del tipo de técnica usada. Adicionalmente, estas clasificaciones servirán como base de un modelo de clasificación de estudios por encuesta que se propondrá al final de este apartado.

Existen por lo menos diez criterios para referirse a los estudios por encuesta. Éstos proceden tanto de las características epistemológicas como de las propiamente técnicas y operativas relevantes a la encuesta. A continuación se revisarán estos criterios.

Área de interés

Un primer criterio frecuentemente usado para definir a las encuestas, está relacionado con el campo de aplicación o área de interés de los resultados. Los campos de aplicación más comunes son los siguientes:

- a) *Mercadotecnia*. Dentro de este campo están los estudios dirigidos a apoyar a las empresas en su labor de desarrollo, comercialización y venta de productos y servicios.
- b) *Medios de comunicación y publicidad*. Aquí se encuentran las encuestas que estudian la emisión y difusión de los mensajes, así como el impacto de los medios de comunicación y el comportamiento de las audiencias.
- c) *Opinión pública*. Este apartado está formado por las encuestas que estudian el comportamiento político del ciudadano.
- d) *Salud*. En este campo se encuentran las encuestas dirigidas a estudiar los conocimientos, actitudes y conductas de la población con relación a los temas relevantes de salud.
- e) *Cultura y sociedad*. Aunque el estudio de la cultura y la sociedad contempla, desde luego, las actividades comerciales, políticas y de comunicación, formalmente las encuestas sobre este tema se definen como aquellas que, desde un punto de vista más amplio, no tiene un interés pragmático. Generalmente son estudios académicos.
- f) *Organizaciones*. A pesar de que es uno de los campos más recientes en el uso de las encuestas, rápidamente cobra importancia. Las encuestas de

organizaciones están dirigidas a entender el comportamiento de las entidades morales. Este tipo de encuestas se conocen como encuestas *business to business*.

- g) *Demografía*. Las encuestas realizadas dentro de este campo están dirigidas a entender la dinámica de la población.

Propósito

El segundo criterio para referirse a las encuestas tiene que ver con el propósito o la aplicación que se dará a los resultados. En todos los campos de aplicación mencionados, al elaborar y desarrollar proyectos, existe una etapa estratégica y una táctica. Dentro de la primera etapa se diseña y se planea la actividad. En la etapa táctica se implementa y da seguimiento al plan desarrollado en la etapa estratégica. En este sentido, de acuerdo a la etapa en que serán aplicados los resultados, existen dos tipos de encuestas: las que proporcionan información amplia de carácter estratégico y las que proporcionan información de carácter específico o táctico. Las primeras son de carácter exploratorio; en sí mismas no proponen una conclusión. Las segundas son de carácter concluyente y los resultados proponen una conclusión.

De acuerdo a varios autores, las encuestas tácticas se pueden dividir, a su vez, en estudios que proporcionan información para el desarrollo del plan o proyecto, y encuestas para la evaluación y seguimiento de proyectos. De esta forma existen tres tipos de encuestas conforme al propósito:

- a) De planeación o estratégicas.
- b) Para el desarrollo de proyectos.
- c) Evaluativas o de monitoreo.

Para ilustrar esta clasificación, podemos usar el ejemplo de un típico programa de investigación en mercadotecnia. Como parte de la etapa estratégica se realiza una encuesta de usos y hábitos de la categoría de producto a la que se pretende ingresar. Esta encuesta permite identificar la situación de ese mercado en particular, así como las oportunidades y posibilidades del mismo. Después de realizar un análisis se decide probar el nuevo producto, fórmula o marca. La investigación de desarrollo permite conocer el desempeño de cada una de las alternativas, así como de los distintos elementos que compone este nuevo producto o concepto. Finalmente, una vez que el producto ganador ha sido lanzado al mercado, la investigación evaluativa o de monitoreo permite conocer cuál es el desempeño de este producto en el mercado real.

El análisis de la investigación evaluativa o de monitoreo, observada de una manera sistemática y más amplia, puede aportar a su vez elementos que sirvan para una nueva etapa estratégica. De esta forma se establece un círculo interactivo entre las tres etapas.

Enfoque metodológico

Si el criterio anterior se refiere al uso o aplicación que se le dará a los resultados, el enfoque metodológico clasifica a las encuestas considerando el propósito lógico o nivel de conocimiento que pretenden alcanzar. Desde este punto de vista existen cuatro criterios:

- a) *Encuestas exploratorias*. El propósito de estas encuestas es tener un primer acercamiento al fenómeno o tema estudiado. Sirven para identificar las características generales o dimensiones del problema, así como para establecer hipótesis y alternativas de trabajo. Frecuentemente, las sesiones de grupo o grupos de discusión cumplen esta función preliminar al levantamiento de una encuesta.
Las encuestas exploratorias, además de ayudar a identificar las categorías de análisis o el esquema conceptual, también permiten obtener información para el diseño de muestras y cálculos estadísticos que se usan en las partes subsiguientes de los programas de investigación.
- b) *Encuestas descriptivas*. El propósito de éstas es describir con precisión las características del fenómeno observado. Dicha descripción puede o no estar relacionada con una hipótesis de trabajo. La mayor parte de las encuestas al describir el fenómeno con porcentajes o promedios cumplen con este propósito de conocimiento.
- c) *Encuestas explicativas*. Su objetivo es describir la frecuencia con que ocurre un fenómeno asociado a otro. Este tipo de encuestas señala relaciones entre los distintos factores o características que constituyen un fenómeno, sin establecer causalidad.
- d) *Estudios causales*. Éstos pretenden marcar una relación causal entre las variables. En sentido estricto sólo los diseños experimentales permiten establecer esta relación de causalidad entre las variables. Sin embargo, en la práctica cotidiana de la investigación social, con frecuencia se utilizan diseños experimentales o cuasi-experimentales combinados con metodología de encuesta. En este tipo de proyectos, las poblaciones de estudio que han de ser sometidas al control experimental son seleccionadas a través de la técnica de encuesta. De hecho, con frecuencia existe confusión al identificar y describir este tipo de estudios que combinan las dos meto-

dologías. La confusión puede eliminarse si en el proyecto se identifica y separa la información que es tratada como parte del diseño experimental de la información que, a su vez, es tratada como parte de la encuesta.

Para ilustrar las diferencias entre los cuatro tipos de encuestas de acuerdo al propósito de conocimiento, tomaremos como ejemplo un programa de investigación ligado a un típico programa de comunicación.

Supongamos que una institución social desea elaborar un programa de comunicación dirigido a promover la planificación familiar entre una población específica. Para identificar cuáles son las principales motivaciones y frenos asociados a la planificación familiar, la institución realiza una encuesta exploratoria. En esta encuesta, típicamente de pocos casos, se realizan preguntas abiertas que permiten identificar las grandes dimensiones del problema.

Una vez identificadas las dimensiones y características del problema, se realiza una encuesta entre una muestra representativa en cuanto número y forma de seleccionar a los sujetos. Esto permite identificar y cuantificar con precisión la relevancia de cada uno de los aspectos estudiados en la etapa exploratoria.

Ya que se han determinado los aspectos relevantes, se diseña una campaña de comunicación. Con una encuesta evaluativa de tipo explicativo de la campaña se podría concluir que la campaña tiene mayor o menor aceptación entre ciertos grupos de la población, pero no que existe una relación de causalidad entre la exposición y la aceptación. Así, si encontramos que existe un mayor uso de métodos anticonceptivos entre las mujeres que están en riesgo de embarazarse y que estuvieron expuestas, no significa que la campaña sea la causa del uso, simplemente que existe una asociación entre haber visto la campaña y usar anticonceptivos. En esta conclusión no hay dirección; no se apunta hacia la causalidad.

Para encontrar si existe causalidad entre la campaña y el uso es necesario un diseño experimental o cuasi-experimental. Por ejemplo, una situación controlada, en la cual, el mensaje es transmitido en una localidad y en otra, poblacionalmente comparable, no. Esta situación de contraste permitiría inferir causalidad.

Los mercados de prueba de productos utilizados con frecuencia en mercadotecnia son ejemplos de diseños experimentales que utilizan la técnica de la encuesta.

Tipo de muestreo

El tipo de muestreo utilizado para seleccionar a las personas es uno de los criterios más usados para describir a las encuestas. Desde el punto de vista técnico, los resultados de una encuesta deberían ser representativos de la población. Como veremos adelante, en el apartado de muestra, para lograr esta característica la

muestra debe ser elegida de manera probabilística. Sin embargo, cuando se busca solamente explorar un fenómeno sin obtener estimaciones precisas, la encuesta no probabilística es de mucha utilidad.

Una explicación más amplia de las técnicas de muestreo se presentará en el apartado correspondiente.

Unidad estudiada

Un criterio para tipificar a las encuestas, muy directamente relacionado con el área de interés, es la unidad estudiada. Este criterio se refiere a la unidad de análisis. En la investigación social aplicada se pueden distinguir los siguientes tipos de encuestas, de acuerdo a la unidad estudiada:

- a) *Consumidores o usuarios.* En los estudios de mercadotecnia las encuestas están dirigidas a entender las conductas de los usuarios o consumidores de los productos o servicios.
- b) *Audiencia.* En los estudios de medios las encuestas están dirigidas a entender el funcionamiento y las características de la audiencia
- c) *Ciudadanos.* Desde el punto de vista político, las encuestas en los estudios de opinión conciben a la población como ciudadanos.
- d) *Población.* Los estudios de demografía, los culturales y los sociales estudian a las personas como integrantes de grupos.
- e) *Entidades comerciales.* Especialmente en mercadotecnia, se realizan con frecuencia estudios cuya unidad de análisis es el establecimiento comercial que distribuye los productos o servicios.
- f) *Organizaciones.* Dentro de la mercadotecnia industrial o para la comprensión del comportamiento macrosocial, se realizan encuestas cuyo énfasis está en las instituciones económicas, políticas o sociales. En este tipo de encuesta, aunque la unidad de análisis es la organización, se selecciona un informante calificado que represente el interés y el punto de vista de la entidad.

Tipo de levantamiento

Tal vez este sea el criterio más usado para describir la naturaleza de una encuesta, ya que se refiere a la forma como son obtenidos los datos. Existen básicamente tres formas, que por el momento sólo serán enunciadas. En el apartado de levantamiento de la información se ampliarán estas técnicas. Las encuestas pueden ser:

- a) Personales
- b) Telefónicas
- c) Por correo

Forma de registrar la información

En cuanto a la forma de registrar la información, existen dos grandes tipos de encuestas. Las que utilizan papel y lápiz, ya sean aplicadas por un encuestador o de manera autoaplicada, y las que utilizan ayudas automatizadas para obtener esta información. En estas últimas se encuentran las encuestas aplicadas a través de CATI (*Computer Assisted Telephone Interview*) y CAPI (*Computer Assisted Personal Interview*).

Temas abordados

Otro criterio para describir la naturaleza de una encuesta es el tipo de tópicos incluidos. Con frecuencia escuchamos referencias a encuestas de valores, conductas, estilos de vida, etcétera. Tal vez este sea el criterio más impreciso para definir una encuesta. Es difícil encontrar un estudio, sin importar el área o propósito, que se concentre solamente en un tópico. En el apartado de diseño y construcción de instrumentos, se aborda ampliamente este tema y su relación con las técnicas de medición. Por el momento, solamente se enumeran:

- a) Valores.
- b) Necesidades, gustos e intereses.
- c) Conocimientos.
- d) Actitudes e imágenes.
- e) Opiniones.
- f) Intenciones.
- g) Conductas, usos y hábitos.
- h) Demográficos.

Periodicidad

En cuanto a la frecuencia con que es levantada la información, existen básicamente dos tipos de encuestas: únicas en el tiempo y de seguimiento.

- a) *Encuesta unitaria*. Son realizadas de manera única y responden generalmente a un problema específico.

- b) *Encuestas de seguimiento*. En éstas los temas son explorados a través del tiempo. Este seguimiento puede ser continuo o discontinuo en el tiempo. Las encuestas de seguimiento pueden ser levantadas en muestras independientes cada vez, o en la misma muestra durante todo el tiempo. Las primeras encuestas son conocidas como transversales y las segundas como longitudinales o de panel.

Destino de la información

Considerando el destino de los resultados de la encuesta, así como el patrocinio, se pueden señalar tres tipos de encuestas:

- a) *Encuestas ad-hoc*. En este tipo de encuestas un patrocinador o cliente único solicita y cubre en su totalidad los costos del estudio. Los resultados del mismo le pertenecen y no pueden ser difundidos a otras instancias.
- b) *Encuestas de multicitiente o sindicadas*. Éstas, ofrecen información de interés amplio. Los resultados pueden ser compartidos por varios usuarios, los cuales sufragan los gastos de manera compartida.
- c) *Encuestas de difusión pública*. Son realizadas para fines académicos. El propósito de estos estudios es el conocimiento en sí mismo. Sus resultados, generalmente financiados con fondos públicos, se difunden a través de medios públicos.

Identificación del problema

En cualquier proyecto, para obtener los resultados buscados, es necesario partir de una definición clara de cuál es el objetivo. En el trabajo organizado, especialmente en el que se refiere a la búsqueda de respuestas científicas, esta meta se expresa como un problema. Al definir un problema se formula el camino que habrá de tomar el proyecto, las circunstancias que lo anteceden y acotan, así como las implicaciones relacionadas. Mediante la definición del problema también determinamos el qué y el cómo vamos a medir la problemática de interés y si esto es factible de someterse a prueba. En este sentido, el problema de investigación puede observarse desde dos dimensiones: la lógica de la formulación del problema y el análisis de la situación en que éste se da.

La formulación del problema expresa de manera lógica la naturaleza de lo buscado. Tal como lo señala Rojas Soriano, "en la formulación científica del problema

radican las posibilidades de su solución.”² Al definir un problema se deben especificar los siguientes aspectos:

- a) Naturaleza del conocimiento buscado.
- b) Partes constitutivas del problema.
- c) Forma de operacionalizar o medir las partes.

En primer lugar, al formularse el problema debe expresarse la naturaleza del proceso de conocimiento que hemos de seguir para encontrar los resultados buscados. En este sentido, en la formulación deben existir términos tales como describir, explorar, encontrar la relación causal entre dos circunstancias, etcétera. En el apartado anterior, al hablar del enfoque metodológico, se planteó la tipología básica de estudios con relación a la forma de conocimiento. Toda definición de problema debe referirse a alguna de las formas de conocimiento, y este sentido define el método que ha de seguirse para encontrar la solución.

En segundo lugar, en la formulación del problema deben estar incluidas las partes que dan origen al estudio. Es decir, el o los temas relacionados.

Finalmente, en tercer lugar, estos temas o conceptos deben ser traducidos de manera precisa a la forma como han de ser tratados o medidos dentro del marco del método de conocimiento seleccionado: “Normalmente, el proceso de pasar de un tema de investigación a un problema investigable implica un afinamiento progresivo de conceptos y un estrechamiento progresivo del campo.”³

En el siguiente cuadro se presentan algunos ejemplos de formulación de problemas de investigación a partir de un tema de investigación.

Tema de investigación	Problema de investigación
Imagen de un producto	Explorar las verbalizaciones asociadas a la descripción de un producto.
Situación de una marca	Describir el nivel atribución publicitaria de la campaña X.
Diferencias en la votación	Conocer cuál es la relación entre la intención de voto y la pertenencia a un partido político.
Impacto de una campaña publicitaria	Conocer la relación causal entre la exposición a un mensaje publicitario y la compra del artículo o servicio publicitario.

² Raúl Rojas Soriano, *Métodos para la investigación social. Una propuesta dialéctica*, Plaza y Valdés, México, 1992, p. 33.

³ Sellitz, C., M. Jahoda, M. Deutsch y S.W. Cook, *Métodos de investigación en las relaciones sociales*, Ediciones Rialp, Madrid.

En cuanto al análisis de la situación, al formular un problema es fundamental considerar cuatro aspectos:

- a) El origen del problema.
- b) Los requisitos.
- c) La estructura conceptual.
- d) El valor de la información.

En lo que se refiere al origen del problema, es necesario conocer los antecedentes; es decir, quién, cuándo y ante qué situación se planteó la necesidad de realizar una investigación. Esta información permitirá plantear con visión amplia la resolución del problema. Dentro de los antecedentes, es también fundamental conocer cuál ha sido la experiencia anterior en relación al tema estudiado.

Por otra parte, al definir un problema hay que partir de los requisitos esenciales de éste. Los aspectos imprescindibles que deben definirse *a priori* son el tipo de conocimiento buscado, el universo o población a estudiar, la naturaleza y forma de medir los temas y conceptos, el tratamiento estadístico requerido, así como la forma de entregar los resultados. Todos estos puntos deben ser discutidos y definidos con precisión.

De esta manera, al hablar de población a estudiar se requiere una definición en términos de edad, sexo, nivel socioeconómico, uso o compra de algún producto específico, ocupación o posición o actitud ante un estímulo determinado. Asimismo, es necesario definir *a priori* cuál es la forma en que se debe operacionalizar un concepto. Un mismo concepto puede ser definido de varias formas dependiendo del ámbito en que se use o del interés específico del estudio.

Por otra parte, al definir el problema es necesario considerar el modelo conceptual o teórico que fundamenta esta formulación. En muchos estudios, sobre todo de carácter táctico, no existe una formulación explícita de un modelo de análisis. Sin embargo, siempre es conveniente revisar la literatura o la experiencia previa, a fin de construir un modelo que permita enclavar la formulación del problema dentro de un contexto de conocimiento más amplio. Esta alternativa posibilitará siempre la generación de conocimiento y el aprovechamiento de todos los estudios para situaciones posteriores.

El último aspecto a considerar en la formulación de un problema, es el valor de la información. El valor de la información se refiere a la relación que existe entre la inversión de recursos que habrá de hacerse en la investigación y los resultados que aportará. Para realizar un análisis del valor, es conveniente dejar claro los siguientes puntos:

- a) ¿Quiénes serán los usuarios de la información?
- b) ¿Cuáles son las expectativas que existen de la información?

- c) ¿Cuáles son los alcances y limitaciones de la información?
- d) ¿Cuál será el uso que se le dará a la información?
- e) ¿Cuál es la aplicabilidad de la información?

Un análisis detallado de estos aspectos facilitará la formulación del problema, así como la consecución de los resultados buscados.

Para facilitar la comunicación de estos aspectos es conveniente elaborar un documento. En la investigación comercial este documento se conoce como *Brief*, que es un resumen de los temas listados que elabora preferentemente quien solicita el estudio. Para ayudar a la elaboración de este documento se recomienda proporcionar un formulario que contemple los temas listados.

Diseño y selección de la muestra

La técnica del muestreo es una herramienta fundamental de la investigación por encuesta. En sentido amplio, el muestreo define a esta técnica. Es un hecho histórico que la encuesta cobró relevancia como herramienta de análisis cuando los estudiosos del campo social trasladaron las técnicas de muestreo de las ciencias exactas al estudio de las relaciones sociales. En ese momento, la investigación por encuesta a partir de muestras se utilizó para conocer el impacto de los nacientes medios de comunicación, así como la fuerza de la propaganda y las campañas políticas. Poco a poco, su uso fue extendiéndose a la publicidad, la mercadotecnia, la demografía y, en general, a todos los campos del conocimiento de la sociedad.

El valor del muestreo radica en la posibilidad de conocer el comportamiento de una población infinita, a partir de un subconjunto. Este procedimiento aporta una valiosa solución: sin necesidad de realizar un censo, es decir la observación o medición de todos los individuos de una población, podemos conocer las características que nos interesan. Bajo este procedimiento es posible ahorrar dinero y tiempo, además, aunque parezca sorprendente, permite obtener mayor precisión en los resultados, que mediante un censo. Las dimensiones de un censo llevan con frecuencia a cometer errores operativos y de medición.

Las muestras utilizadas en las encuestas pueden ser representativas o no de su propio universo. Las muestras representativas reproducen, con cierto grado de error, todas las características de la población o universo estudiado. Las muestras no representativas, conocidas también como no probabilísticas, intencionales o de criterio, no necesariamente guardan las características de la población de donde fueron obtenidas y es imposible calcular cuál es el grado de error que tienen. Este tipo de muestras se utilizan con fines exploratorios, cuando no se pretende inferir los resultados al universo.

Para que una muestra sea representativa del universo es necesario cumplir con dos principios fundamentales:

- a) Todos los elementos del universo deben tener probabilidad de ser incluidos en la muestra.
- b) La probabilidad de cada elemento de ser incluido en la muestra debe ser conocida.

Para que una muestra sea representativa, debemos asegurarnos de que todos los elementos que constituyen a la población tengan posibilidad de ser elegidos. Si por la forma de seleccionar eliminamos la posibilidad de que alguno de los elementos que compone el universo sea elegido, entonces estamos eliminando la posibilidad de representar las características de ese elemento o tipo de elementos.

Cuando se realiza una encuesta para conocer la opinión de la población de una ciudad, eligiendo a las personas en un punto de tráfico intenso, no se está obteniendo una muestra representativa. Esto se debe a que no todas las personas que habitan en la ciudad acostumbran transitar por ese punto de la ciudad. Todas las personas que nunca transitan por ese punto; más aún, las que no pasaron en el momento en que se realizó la encuesta, no tienen oportunidad de ser incluidas en la muestra y, por lo tanto, se reducen las posibilidades de que sus opiniones queden representadas en ese estudio.

En estricto sentido, para que todos los elementos de una población puedan ser incluidos en una muestra, es necesario contar con una lista de todos los elementos, a partir de la cual se pueda realizar ese sorteo. Cuando se realiza este procedimiento, elegimos una muestra en particular, pero todos los elementos tuvieron la posibilidad de ser incluidos. Esto asegura la representatividad.

En la mayoría de los casos de encuesta en ciencias sociales, no se cuenta con esta lista o marco muestral. Ante la ausencia de esta enumeración de elementos, es necesario utilizar procedimientos alternativos de selección que aseguren este principio. En el ejemplo anterior, si bien no se tiene la lista de los habitantes de la ciudad, es posible asegurar la posibilidad de que todos los habitantes puedan ser elegidos a partir de un procedimiento de selección por etapas. El procedimiento típico sería el siguiente:

- a) A partir de un mapa actualizado de la ciudad se seleccionan aleatoriamente algunas manzanas de viviendas. Todas las manzanas de la ciudad tienen posibilidad de ser elegidas.
- b) En cada manzana seleccionada se elige aleatoriamente un número de hogares. Todos los hogares de la manzana tienen posibilidad de ser elegidos.

- c) En cada hogar seleccionado, se escoge aleatoriamente a una persona. Todas las personas del hogar tienen posibilidad de ser elegidas.

Bajo este procedimiento, cualquier habitante de la ciudad tiene la posibilidad de ser elegido. Es factible que sea seleccionada cualquier manzana; dentro de cada manzana, cualquier hogar y dentro de cada hogar, cualquier persona.

El segundo principio para que una muestra sea representativa, se refiere al conocimiento de la probabilidad que tiene cada elemento de ser incluido en esa muestra.

La probabilidad de ser seleccionado se define como las oportunidades que tiene un elemento de ser elegido. Si en una escuela existen 1,000 estudiantes y obtenemos una muestra de 500, la probabilidad de cada elemento de ser sorteado es de 500 entre 1,000. Es decir, un medio.

El inverso de la probabilidad se define como el factor de representatividad de cada elemento; es decir, el número de casos que representa. En el ejemplo de la escuela, si la probabilidad de ser seleccionado de cada estudiante es de $1/2$, el factor de representatividad de cada sujeto es de 2 (inverso de $1/2$). Esto es, la opinión de cada estudiante incluido en la muestra, representa a dos estudiantes de toda la escuela.

En el caso de una encuesta entre los habitantes de una ciudad, el cálculo de la probabilidad o factor de representatividad se obtendría por el siguiente procedimiento:

- a) Si en la ciudad existen 100 manzanas y para la muestra se eligieron 10, entonces la probabilidad de cada manzana de ser elegida en la muestra es de $10/100$. El factor de representación de cada manzana es, por lo tanto, $100/10$ ó 10.
- b) Si en una manzana específica existen 50 hogares y son seleccionados 5, entonces la probabilidad de cada hogar de ser elegido en la muestra de hogares de esa manzana es de $5/50$. El factor de representación de cada hogar en esa manzana es de $50/5$ ó 10.
- c) Si en un hogar elegido dentro de esa manzana existen 5 personas y es seleccionada una persona, entonces la probabilidad de cada persona en ese hogar de ser elegida es de $1/5$. El factor de representación de cada persona en ese hogar es de $5/1$, es decir, 5.
- d) Para conocer la probabilidad de selección de cada persona elegida sobre el total de la población de la ciudad, debemos multiplicar las probabilidades obtenidas en cada etapa de selección; es decir, la probabilidad de selección de la manzana, por la probabilidad de selección del hogar, por la probabilidad de selección de la persona: $10/100$ por $5/50$ por $1/5$. Esto es $1/500$. Significa que cada persona de la ciudad de este ejemplo tiene una

probabilidad entre 500 de ser elegida. También significa que la opinión de esta persona en particular representa la opinión de 500 personas de esa ciudad.

Para diseñar una muestra representativa que cumpla con los criterios señalados, es necesario definir los siguientes tres aspectos:

- a) Definición del universo o población de estudio.
- b) Determinación del tamaño de la muestra.
- c) Forma de selección de la muestra.

Definición del universo

El primer paso para diseñar una muestra consiste en definir la población o universo de estudio. El universo es el conjunto de elementos a los cuales se desea extrapolar los resultados de la muestra. Todos los elementos del universos son homogéneos en cuanto a su definición. De esta forma, podemos definir como universo de estudio a todos los habitantes de una ciudad o a la población en edad de votar de la República mexicana o a las amas de casa usuarias de un determinado producto.

Los criterios para definir al universo pueden ser demográficos, sociales, económicos, políticos, mercadológicos o psicológicos.

La enumeración de todos los elementos de un universo o población se conoce como marco muestral. El marco muestral permite realizar una selección de elementos, asegurando la posibilidad de que todos puedan ser elegidos.

Existen dos tipos de marcos muestrales: los marcos lista y los marcos área.

En los marcos lista los elementos son enumerados uno a uno. El listado de los poseedores de una determinada tarjeta de crédito es un ejemplo.

En los marcos área se representan los territorios geográficos que ocupan los integrantes de una población. El mapa de una ciudad constituye un marco área.

A partir de los dos tipos de marcos, tanto área como lista, es posible realizar un sorteo para seleccionar una muestra.

Prácticamente no existe el marco muestral perfecto. Con frecuencia éstos incluyen elementos que no responden a la definición del universo, excluyen a los elementos que sí deberían de estar o duplican a otros. Estas características se llaman sobrecobertura, subcobertura y duplicidad, respectivamente.

La sobrecobertura, subcobertura y duplicidad de los marcos es producida por la desactualización y falta de verificación de la información. Tomando el caso de la lista de suscriptores de la tarjeta de crédito, la sobrecobertura se daría cuando los suscriptores dejan de ser tarjetahabientes y, sin embargo, no son borrados de

la lista; la subcobertura cuando los nuevos tarjetahabientes no están aún en la lista y la duplicidad cuando un tarjetahabiente aparece dos o más veces. El caso más común de subcobertura se observa en los mapas de las ciudades. Debido a que los levantamientos se realizan con esparciamientos prolongados, con frecuencia encontramos mapas que no incluyen las nuevas urbanizaciones.

Debido al alto costo que significa verificar y actualizar, es difícil encontrar un marco muestral perfecto. Al diseñar una muestra, es necesario partir del hecho de que el marco muestral tiene algún grado de error.

La información referida al universo o población de estudio se conoce como parámetro. La información de la muestra es definida como estimador. De esta manera, los resultados de una muestra son los mejores estimadores de los parámetros. Los parámetros sólo son susceptibles de conocerse a través de un censo.

Para finalizar el apartado de definición del universo, apuntaremos que los parámetros se expresan con letras mayúsculas y los estimadores con letras minúsculas.

Determinación de tamaño de muestra

Después de haber definido el universo o población de estudio, el siguiente paso consiste en determinar el tamaño de la muestra.

Con frecuencia escuchamos referencias al tamaño de la muestra como un porcentaje del universo. Sin embargo, este criterio, en principio lógico y de sentido común, es equivocado. La muestra para ser representativa no tiene que significar una determinada proporción del universo. Sólo imaginemos el caso de una ciudad de varios millones de habitantes; cualquier muestra que fuese obtenida como porcentaje de este universo sería inmanejable y altamente costosa.

De esta manera, el tamaño de una muestra no depende de manera proporcional al tamaño del universo, depende de otras condiciones. Una muestra puede ser pequeña y representativa. A este respecto se puede citar el ejemplo de la cocinera que prueba la sopa, no es necesario que pruebe grandes cantidades de sopa para saber si tiene sal. Con una porción adecuada es suficiente. De qué depende entonces el tamaño de la muestra. El tamaño de la muestra depende de 11 factores que a continuación se listan:

TAMAÑO DEL UNIVERSO

Aunque ya mencionamos que el tamaño de la muestra no es una proporción del tamaño del universo, sí existe una dependencia no lineal entre ambos. Esto es, cuando el tamaño del universo es pequeño se requiere una muestra de tamaño

muy similar. A medida que el tamaño del universo crece, el tamaño de la muestra requerida se incrementa, pero en menor medida; se llega a un punto en que, independientemente de cuánto crezca el universo, el tamaño de la muestra se estabiliza. Para ilustrar este principio, en el siguiente cuadro se presentan los tamaños de muestra requeridos para distintos tamaños de universos, considerando un nivel de confianza de 95 por ciento y un margen de error de más menos 5 por ciento. De estos dos conceptos hablaremos adelante.

Tamaño del universo **Tamaño de la muestra**
Margen de error +/- 5%

10	10
20	19
50	44
100	79
200	131
500	216
1,000	275
2,000	319
5,000	353
10,000	366
100,000	378
1,000,000	380
50,000,000	380

La relación entre el tamaño del universo y el tamaño de la muestra está basado en la evidencia de que las opiniones, conductas, actitudes y en general todos los temas estudiados en una encuesta son finitos y repetitivos. Esto supone que a partir de un tamaño determinado, las observaciones son redundantes y no aportan nueva información para realizar las estimaciones. Para ilustrar este principio, supongamos que llevamos a cabo una encuesta de opinión donde existen solamente dos opiniones; aceptación o rechazo de una medida. Si tomamos una muestra de una persona, asumiremos que la opinión de esa persona representa a todo el universo; si tomamos dos personas, es probable que obtengamos una opinión favorable y una desfavorable, es decir 50/50. A medida que vayamos incrementando el tamaño de la muestra, los resultados obtenidos se acercarán cada vez más a la opinión del universo; supongamos 30 por ciento de aceptación y 70 por ciento de rechazo. Si cada vez que incrementamos el tamaño del universo observa-

mos el comportamiento de los resultados, nos daremos cuenta que a partir de cierto número no importa qué tanto crezca el tamaño de la muestra, los resultados no cambian. Este principio, explicado teóricamente, también es factible de ser probado empíricamente a partir de la delimitación finita y repetida de los eventos a estudiar.

HOMOGENEIDAD–HETEROGENEIDAD DEL FENÓMENO

Si bien los eventos a estudiar se repiten y se encuentran dentro de un territorio finito, existen comportamientos distintivos. Hay fenómenos con mayor homogeneidad, es decir, con menos variación entre los posibles valores que pueden tomar, mientras hay otros con mayor variación o heterogeneidad. El nivel de variación del fenómeno a estudiar o de homogeneidad–heterogeneidad es un determinante del tamaño de la muestra. A mayor variación se requiere mayor tamaño. En el cálculo del tamaño de muestra, la homogeneidad y heterogeneidad se expresan como la probabilidad de ocurrencia o no ocurrencia.

Para explicar los siguientes dos factores que afectan el tamaño de la muestra, utilizaré el tradicional ejemplo del arquero. Imaginemos a un arquero que tira repetidamente hacia una diana. En la medida que las flechas se acercan al centro de la diana, los tiros son precisos; en la medida que las flechas repiten el lugar donde se clavan, los tiros son exactos. Bajo esta concepción, un arquero puede ser exacto, pero no preciso, o al revés; es decir, siempre lanza las flechas al mismo lugar, aunque este lugar esté lejos del centro de la diana.

La imagen del arquero se utiliza para explicar los conceptos de precisión y exactitud de los resultados de una muestra. Imaginemos que la diana es el universo y los distintos tiros del arquero son las posibles muestras. En la medida que los resultados de una muestra se acerquen a los resultados que se obtendrían si se realiza un censo (la flecha cerca del centro), los resultados son precisos. En la medida que los resultados de varias posibles muestras se repitan (las distintas flechas repitiendo el lugar), los resultados serán exactos.

PRECISIÓN O MARGEN DE ERROR

El margen de error o nivel de precisión es el límite de un valor muestra o estimador, dentro del cual se encuentra el verdadero valor en el universo (parámetro). Utilizando el ejemplo del arquero, es la distancia máxima que existe entre el centro de la diana y el lugar donde se clavan las flechas. 5 por ciento de margen de error significa que el valor en el universo o parámetro se encuentra entre los límites comprendidos por ± 5 por ciento con relación al valor muestral.

EXACTITUD O NIVEL DE CONFIANZA

El nivel de confianza o grado de exactitud representa la probabilidad de que el valor del universo (parámetro) se encuentre dentro del margen de error. En términos del ejemplo del arquero, es la probabilidad de que éste repita los disparos dentro de una determinada distancia del centro. 95 por ciento de confianza y 5 por ciento de margen de error significa que existe una probabilidad de 95 por ciento de que el valor del universo (parámetro) se encuentre dentro de un margen de 5 por ciento con respecto al valor muestral (estimador).

NÚMERO DE ESTRATOS

Con frecuencia, además de obtener resultados para el total de una población, se requieren resultados con suficiente representatividad para los grupos que constituyen el universo. Estos subuniversos son los estratos o dominios del universo o población de estudio. Una muestra nacional puede diseñarse considerando estratos de región geográfica o tamaño de localidad. Esto significa que existen submuestras para obtener resultados independientes para las distintas regiones geográficas del país o agrupaciones por tamaño de localidad (metrópolis, zonas urbanas, rurales, etcétera). Para fines de tamaño de muestra, cada estrato se calcula de manera independiente. Esta es una de las razones que explican el por qué, en ocasiones, se diseñan muestras que exceden el tamaño necesario para obtener resultados a nivel total. El tamaño final es la acumulación de los tamaños calculados para cada uno de los estratos.

ETAPAS DE MUESTREO

Cuando no existe un marco muestral de las unidades finales de muestreo, se requiere seleccionar a partir de unidades mayores o conglomerados. Esta selección, que será explicada en el apartado de selección de muestra, se conoce como muestreo polietápico.

Al conglomerar las unidades aumenta la probabilidad de que disminuya la varianza de las unidades observadas. Por esta razón es necesario obtener muestras mayores.

La diferencia entre la variación de los datos observados a partir de un muestreo directo y la variación de los datos obtenidos a partir de un muestreo polietápico, se conoce como efecto del diseño. Esta razón, obtenida a partir de la comparación de las varianzas, determina la proporción en que ha de aumentarse el tamaño de la muestra diseñada para obtenerse directamente, si ésta es obtenida por etapas.

Si la variación de los datos es idéntica en los dos tipos de muestra, la razón será uno. Por tanto el tamaño de muestra permanecerá idéntico. En cambio, si la variación de los datos, obtenidos por una muestra de etapas se reduce a la mitad, entonces la razón será de dos. Eso significa que el tamaño de muestra deberá duplicarse.

CONGLOMERACIÓN DE UNIDADES

Al hacer una selección en etapas, las unidades finales son elegidas en conjuntos o conglomerados. Por otro lado, con objeto de realizar un levantamiento más eficiente, con frecuencia, aunque existe un marco disponible, las unidades son seleccionadas de manera conglomerada.

En la mayor parte de las encuestas se realizan varias mediciones en un mismo punto muestral. De esta forma, es común observar cómo en una manzana se llevan a cabo seis u ocho entrevistas. Este tipo de procedimiento conocido como muestreo por conglomerados, si bien hace más eficiente el levantamiento, reduce la posibilidad de encontrar la variación del fenómeno en su espectro real. En las encuestas donde se realizan varias mediciones dentro de la misma manzana, se ha observado que los comportamientos y opiniones son muy similares entre sí.

Para calcular el tamaño de la muestra, los algoritmos suponen que las mediciones serán obtenidas de todo el universo. Al concentrar en algunos puntos las unidades de la muestra estamos eliminando factibilidad de encontrar todas las posibles variaciones del fenómeno. Por esta razón, cuando se utiliza el procedimiento de conglomeración es necesario aumentar el tamaño de la muestra. El aumento en el tamaño de la muestra debe ser inversamente proporcional a la reducción de la varianza que se realiza al agrupar unidades.

ESTADO DEL MARCO MUESTRAL

La situación del marco muestral es fundamental para calcular el tamaño de la muestra. Las fórmulas suponen que el marco muestral tiene un cierto grado de precisión; es decir que están contemplados todos y sólo los elementos del universo. Las fuentes de imperfección del marco muestral, listadas en el apartado correspondiente, deben ser compensadas aumentando el tamaño de la muestra.

EFFECTIVIDAD DE LA MUESTRA

Una vez seleccionados los elementos que han de constituir una muestra, con frecuencia nos encontramos con la imposibilidad de contactarlos. Técnicamente,

este factor se conoce como no respuesta. Algunas de las formas más comunes de no respuesta son la no localización del elemento (no se encuentra la dirección, el teléfono no contesta o está siempre ocupado, o la persona no es localizada) y la negación o imposibilidad del elemento para dar la información (rechaza la entrevista o no puede contestar). Bajo algunos supuestos de muestreo, es posible sustituir a los elementos que no fueron localizados o que se negaron a contestar.

En la estimación del tamaño de muestra es necesario considerar el nivel de no respuesta esperado, así como los procedimientos de sustitución y reemplazo.

TÉCNICA DE RECOLECCIÓN DE DATOS

La forma de recoger los datos también determina, de manera indirecta, el tamaño de la muestra. El tipo de marco, los niveles de efectividad de la muestra, las etapas y la conglomeración de unidades son, por naturaleza, distintas en las diversas técnicas de levantamiento. En una encuesta telefónica, los niveles de cooperación, las etapas al seleccionar los números, la conglomeración, etcétera, son fundamentalmente diferentes a las posibles características que en este sentido pueden tener la encuesta personal o por correo.

En el apartado de levantamiento de la información revisaremos con mayor detalle las principales técnicas.

RECURSOS DISPONIBLES

Hasta este momento hemos revisado las características técnicas que determinan el tamaño de una muestra. Sin embargo, en la mayoría de los casos, los determinantes más importantes tienen que ver con la disponibilidad de recursos materiales aprovechables, así como la oportunidad con que se requieren los resultados. En este sentido es importante observar que, si bien es posible adaptarse a los recursos disponibles en el tiempo necesario, los resultados obtenidos tienen ciertas restricciones y determinados niveles de precisión.

Dentro de la teoría del muestro y probabilidad existen diversos procedimientos para el cálculo de los tamaños de la muestra: todos ellos consideran los elementos que hemos enumerado. A continuación se presenta una fórmula genérica para el cálculo del tamaño de muestra. Las variables que considera la fórmula son las siguientes:

Variable	Descripción
n	Tamaño de la muestra.
N	Tamaño del universo.
p	Probabilidad de ocurrencia (homogeneidad del fenómeno).
q	Probabilidad de no ocurrencia (1-p).
Me	Margen de error o precisión. Expresado como probabilidad.
Nc	Nivel de confianza o exactitud. Expresado como valor z que determina el área de probabilidad buscada.

La fórmula es la siguiente:

$$n = \frac{Npq}{\left[\frac{ME^2}{NC^2} (N - 1) \right] + PQ}$$

Supongamos que queremos calcular el tamaño de una muestra para el siguiente caso.

Variable	Descripción
N	3,000,000
p	Desconocemos la probabilidad de ocurrencia. Por esta razón asumimos el mayor punto de incertidumbre, que es de 50 por ciento. Ésta debe ser expresada como probabilidad (.5).
q	1-.5= .5
Me	+/- 5 por ciento de margen de error . Expresado como probabilidad (.05).
Nc	95 por ciento de nivel de confianza o exactitud. Expresado como valor z que determina el área de probabilidad buscada (1.96). ⁴

Al sustituir estos valores en la fórmula, quedaría como sigue:

⁴ El valor z se busca en las tablas de distribución de la curva normal. La mayoría de los textos de probabilidad y estadística contienen esta tabla.

$$n = \frac{3,000,000 (.5)(.5)}{\left[\frac{(.05)^2}{(1.96)^2} (3,000,000 - 1) \right] + (.5)(.5)}$$

Al realizar las operaciones indicadas, el valor de n es igual a 384.1

Esta fórmula considera cuatro de los factores listados. En los diseños estratificados, para cada estrato o subuniverso de estudio se aplica esta fórmula.

Existen fórmulas para corregir este tamaño, dependiendo de las etapas de muestreo, la conglomeración de unidades, el estado del marco muestral y la efectividad de la muestra. Estas fórmulas se encuentran en los textos de muestreo presentados en la bibliografía.

Selección de la muestra

Habiendo definido el tamaño de la muestra, el siguiente paso es determinar el procedimiento a seguir para seleccionar a los sujetos o unidades que habrán de formar parte de ésta. Para asegurarse de que los elementos incluidos cumplan con los requisitos señalados en la introducción de este apartado, es necesario revisar los siguientes aspectos:

- Etapas de diseño o estratificación.
- Etapas de muestreo o pasos en la selección.
- Conglomeración.
- Tipo de selección.

ETAPAS DE DISEÑO O ESTRATIFICACIÓN

Para la selección de las unidades de una muestra, es necesario definir si los resultados que buscamos deben representar a:

- a) Un universo o población de manera global.
- b) Si requerimos representación independiente para subuniversos específicos de la población a representar.

En el primer caso, estamos hablando de un diseño de muestra no estratificada. En el segundo, se define una muestra estratificada. En una muestra estratificada el universo de estudio es dividido *a priori* en subuniversos, cada uno de los cuales es tratado de manera independiente.

En una muestra no estratificada no es posible analizar los resultados de los distintos grupos que componen el universo separándolos de éste. En una muestra estratificada sí es posible analizar los resultados de un subuniverso o estrato de manera independiente, prácticamente como si sólo se hubiese realizado la encuesta en este grupo.

Los estratos tienen las siguientes características fundamentales:

- a) Son subuniversos o grupos del universo.
- b) Son definidos *a priori*. Es decir, son considerados como parte del diseño y determinan las características administrativas del estudio.
- c) Son universos independientes entre sí. El tamaño del estrato es calculado con parámetros propios de exactitud y precisión.
- d) Son proporciones del universo. Esta proporcionalidad puede ser natural o ajustada por procedimientos de ponderación.

Es importante distinguir el concepto de segmentación de estratificación. Una muestra no estratificada puede ser segmentada para fines de análisis. Esto significa que podemos dividir al universo de estudio en grupos de interés. Sin embargo, estos segmentos no son definidos *a priori* como parte del diseño, ni sus tamaños son calculados con parámetros independientes de precisión y exactitud.

Si en una encuesta sobre comportamiento electoral deseamos obtener resultados sobre toda la población en edad de votar, podemos diseñar una muestra no estratificada. Desde luego la estructura interna de esta muestra debe guardar las características del universo. Si además de conocer el comportamiento electoral de toda la población, también deseamos conocer los resultados con suficiente confianza y nivel de precisión para las distintas regiones electorales del país, es recomendable diseñar una muestra estratificada. En el diseño de esta muestra se calculan tamaños y métodos de selección de manera independiente para cada una de las regiones en que se decida dividir al territorio nacional.

ETAPAS DE MUESTREO O PASOS EN LA SELECCIÓN

En la mayoría de las encuestas empleadas en ciencias sociales es difícil partir de un marco muestral donde estén listados todos los elementos de universo. Con mucha frecuencia no se dispone de la lista con los nombres de todos los habitantes de una ciudad o país o de los compradores de un producto o servicio en especial. Por esta razón, los sujetos son seleccionados a partir de un procedimiento alternativo de etapas o polietápico. En este procedimiento, debido a que no existe el listado de las unidades finales, se utilizan los listados o representaciones de los conjuntos o conglomerados de unidades.

En una encuesta típica que pretende obtener resultados representativos de los habitantes de una ciudad, se seleccionan de un mapa actualizado un conjunto de manzanas; dentro de éstas se seleccionan viviendas y dentro de ellas se seleccionan personas. Este procedimiento permite, con cierto grado de error, dar oportunidad a todos los habitantes de la ciudad de ser incluidos en la muestra. De esta forma, si bien no existe la lista de los habitantes, sí existe la representación de las áreas donde están asentadas todas las viviendas en que residen los habitantes de la ciudad. Esta representación es un mapa que, como ya apuntamos, se conoce para este fin como un marco área. Como todo marco, los mapas tienen un nivel de imprecisión: zonas nuevas no incluidas, áreas no existentes, etcétera.

Algunas de las etapas más comúnmente usadas en la selección polietápica es la siguiente:

Conglomerado	Unidades seleccionadas
Dentro de territorios amplios o países	Áreas geográficas
Dentro de áreas geográficas	Localidades
Dentro de localidades	AGEBS (áreas geoestadísticas básicas) ⁵
Dentro de AGEBS	SUBAGEBS (divisiones dentro de AGEBS)
Dentro de SUBAGEBS	Manzanas
Dentro de manzanas	Viviendas
Dentro de viviendas	Personas

CONGLOMERACIÓN

Con frecuencia, las poblaciones a estudiar están dispersas geográficamente. Con el propósito de hacer más eficiente el levantamiento de los datos y de construir un marco referencial, las unidades a estudiar son seleccionadas por grupos. Esto permite reducir costos y tiempos.

TIPO DE SELECCIÓN

Una vez identificados los estratos, las etapas de selección, así como los conglomerados, el último paso consiste en decidir el procedimiento específico en que ha-

⁵ El AGEBS es la unidad de división oficial del territorio nacional. Esta división es realizada observando homogeneidad de población.

brán de seleccionarse las unidades en cada etapa. A continuación se enumeran los procedimientos más frecuentes:

Aleatorio simple

En el muestreo aleatorio simple todos los elementos del universo tienen la misma probabilidad de ser incluidos en la muestra. La selección de un boleto sacado de una tómbola donde están incluidos todos los boletos de una rifa, es un ejemplo de este tipo de muestreo. Todos los boletos tienen la misma probabilidad de ser elegidos. En las encuestas dentro de organizaciones o de afiliados a una organización o servicio, donde se cuenta con la lista de nombres, es posible realizar este tipo de muestreo. En general, este tipo de selección es posible cuando existe la lista de los elementos.

Aleatorio PPT

En el muestreo PPT (Proporcional al tamaño) los elementos del universo tienen una probabilidad de ser incluidos en la muestra, proporcional a su tamaño o peso.

Supongamos que queremos obtener una muestra de 100 localidades del país. Podemos seleccionar de manera aleatoria simple. Esto es, tomar la lista de las más de cien mil localidades del país y elegir 100 al azar, por medio del procedimiento de la tómbola o con la ayuda de alguna tabla de números aleatorios. Bajo este procedimiento, todas las localidades, sin importar el tamaño, tienen la misma probabilidad de ser incluidas en la muestra. Ello supone que una localidad de 10 habitantes tiene la misma probabilidad que la Ciudad de México que cuenta con 17 millones de habitantes. Al seleccionar una muestra bajo este procedimiento, es muy probable que las 100 ciudades sean de tamaños pequeños, pues son éstas las que existen en mayor proporción.

Si queremos obtener una muestra de localidades, que a su vez representen a la concentración de habitantes, es recomendable realizar una selección PPT. En este procedimiento cada localidad tiene tantas oportunidades de salir sorteada como habitantes tenga. Si pensamos en la tómbola, esto supondría que el nombre de la Ciudad de México estaría repetido tantas veces como habitantes tiene (17 millones), mientras el nombre de una comunidad con 10 habitantes sólo estaría incluido este número de veces. Existen procedimientos que facilitan esta selección.

Aleatorio sistemático

Con frecuencia, los muestreos aleatorios simples, sobre todo cuando el tamaño es pequeño, no presentan una dispersión de los elementos seleccionados a lo largo de todo el universo. Para contrarrestar este efecto se utiliza el muestreo aleatorio sistemático. En éste la lista de elementos es dividida en tantos grupos iguales como elementos se requieran. Dentro de cada grupo se selecciona un elemento que guarda siempre la misma posición dentro del grupo. Este procedimiento permite asegurar la dispersión de la muestra dentro del universo.

Selección por cuotas

En las muestras por cuotas los elementos son seleccionados de acuerdo a criterios o características conocidas del universo. Para algunos temas esta técnica no probabilística representa una buena opción de bajo costo.

En las muestras por cuotas la muestra adquiere artificialmente, debido a la predeterminación, las características del universo. En las muestras aleatorias la muestra adquiere las características del universo debido al azar. Por esta razón, sólo es posible obtener muestras por cuotas cuando se conoce *a priori* la composición del universo. Por esta misma razón, las muestras por cuota no pueden proporcionar resultados de universos cuyas características definitorias se desconocen.

El procedimiento de selección en una muestra por cuota es el siguiente:

- a) Selección de variables para definir criterios de cuotas.
- b) Determinación de proporciones para cada categoría de las variables o combinación de variables.
- c) Determinación de tamaños de cuotas.
- d) Instrucciones en campo para ubicación y control de cuotas.

Las muestras por cuotas pueden ser proporcionadas o desproporcionadas. Las muestras proporcionadas son aquellas en que las características del universo son guardadas en sus proporciones reales. Por ejemplo, cuando en una muestra para un estudio de mercado se observan las proporciones de nivel socioeconómico: 8 por ciento de nivel alto, 35 por ciento de nivel medio y 57 por ciento de nivel bajo. Si la muestra es de 400 casos, entonces las cuotas serán como sigue: 32 casos de nivel alto, 140 de nivel medio y 228 de nivel bajo.

En las muestras por cuotas no proporcionadas, con objeto de obtener mayor eficiencia, las proporciones de las características del universo pueden ser modificadas.

En el caso de la muestra anterior, supongamos que queremos conocer las tendencias por nivel socioeconómico. Si la muestra es proporcionada sólo tendremos 32 casos de nivel alto, lo cual nos daría una base muy débil. Es posible determinar cuotas desproporcionadas que nos permitan tener base de análisis suficiente para los grupos poblacionalmente pequeños. Una muestra desproporcionada podría ser 133 casos para cada nivel.

Intencional

En las muestras intencionales las unidades son seleccionadas de acuerdo al criterio de la persona que elige. En este sentido, la probabilidad de selección y por lo tanto la representatividad de los sujetos es desconocida.

Este tipo de selección es frecuentemente usado para sondeos o estudios basados en diseño experimental. En este tipo de selecciones, si bien por definición no interviene el azar, es conveniente que se apliquen reglas que eviten que las personas sean elegidas con base en la simpatía que despiertan en el encuestador.

Randomización

El concepto de randomización es propio del diseño experimental. Sin embargo, debido al uso extendido de este método en las ciencias sociales en general y en especial en la mercadotecnia, con frecuencia se ha confundido con las técnicas de selección para encuestas.

En el diseño experimental el objetivo principal es inferir causalidad. Esta inferencia se logra a través de la manipulación de estímulos en los grupos experimentales. La base de esta inferencia se fundamenta en la similitud de los grupos experimentales que se logra gracias a la randomización. Por lo tanto, la randomización es la técnica de asignación de sujetos a grupos experimentales.

En este sentido, la randomización no busca la representatividad de los grupos de un universo mayor, como lo hace el muestreo probabilístico. La randomización busca la comparabilidad de grupos en función de las características fundamentales. De esta manera, las diferencias observadas entre los grupos deben ser interpretadas como causa de las diferencias en los estímulos presentados.

En una prueba de fórmulas, si los grupos son iguales en cuanto a la distribución de edad, sexo o nivel socioeconómico, las diferencias en gusto por las fórmulas podrán inferirse como causa de la formulación y no de las características de los grupos.

En los diseños de muestras polietápicas, es decir, en aquellos donde la unidad final es elegida en etapas a partir de conglomerados mayores, pueden utilizarse

diversas técnicas de selección. Por ejemplo, la selección de localidades puede realizarse bajo probabilidades proporcionales al tamaño, y las manzanas dentro de una ciudad pueden elegirse por muestreo simple aleatorio.

Con frecuencia se define una muestra describiendo sólo una de las técnicas usadas en alguno de los pasos de selección. Por ejemplo, se dice que es una muestra por cuotas porque en el hogar los miembros fueron seleccionados bajo este método. Pero, esta muestra podría ser probabilística en sus etapas anteriores. Es importante tener en mente la posible combinación de métodos al describir un diseño de muestra.

Un ejemplo de técnicas combinadas en un diseño de muestra estratificada se presenta en el siguiente esquema de una encuesta nacional.

Etapa	Tipo de etapa	Unidad	Técnica de selección
Primera	Diseño (estratos)	Regiones geográficas	Definición <i>a priori</i>
Segunda	Muestreo	Localidades	PPT (probabilidades proporcionales al tamaño)
Tercera	Muestreo	Manzanas	Aleatorio simple
Cuarta	Muestreo	Viviendas	Aleatorio sistemático
Quinta	Muestreo	Personas	Cuotas

Evaluación de la muestra

Para finalizar el apartado referente a la muestra se enunciarán algunos criterios que permiten evaluar si la muestra fue diseñada y obtenida adecuadamente. Estos criterios son:

- a) *Comparación con fuentes externas y otras encuestas.* En la mayor parte de las encuestas existen indicadores que pueden ser comparados con la información de otras fuentes, cuyos datos fueron obtenidos a través de otras técnicas distintas a la encuesta. Estas técnicas pueden ser los censos de población, de vivienda y económicos, así como los registros de ventas, padrones electorales, etcétera.
- b) *Proporción de acuerdo a parámetros.* Las muestras probabilísticas deben reproducir las características del universo; es decir, las proporciones de las características fundamentales de la muestra deben ser similares a las proporciones de los parámetros del universo. En el caso de las muestras no probabilísticas, donde se construye artificialmente

la estructura de la muestra, es posible comparar cómo se comportan aquellas características que no fueron controladas.

- c) *Estabilidad de los datos.* De toda muestra es factible obtener submuestras. Si la muestra es representativa del universo los resultados de las distintas submuestras deben ser comparadas entre sí. Si los resultados de las submuestras presentan tendencias completamente distintas, es factible que existan sesgos en el diseño o construcción de ésta.

En los estudios longitudinales, el comportamiento entre las distintas mediciones es un indicador de la confiabilidad de las muestras.

- d) *Estimación empírica del margen de error.* La estimación del error muestral se realiza de manera teórica. Existe el procedimiento para calcular el margen de error de manera empírica; es decir, el conocer el comportamiento real que tuvo. Este cálculo se obtiene a partir de la comparación de los resultados entre posibles submuestras de la muestra (réplicas).

Diseño y construcción de instrumentos

En este apartado revisaremos el concepto de medición y las principales técnicas de construcción de instrumentos en encuestas.

Medición

La medición es el proceso mediante el cual las ideas o conceptos son dimensionados de acuerdo a reglas. A través de este proceso las ideas abstractas son operacionalizadas y cuantificadas. Algunos ejemplos de este proceso se presentan en el siguiente cuadro:

Conceptos	Operacionalizaciones - medición
Rendimiento académico	Número de aciertos en examen
Actitud	Nivel de acuerdo con afirmaciones
Altura	Distancia en centímetros

La dimensionalización de un concepto se realiza a través de indicadores o preguntas. La agregación de estos indicadores al obtener los resultados permite construir índices que representan al concepto.

Para que un concepto sea medido adecuadamente debe cumplir con dos requisitos:

- a) Validez
- b) Confiabilidad

Una medición es válida cuando la operacionalización representa adecuadamente el concepto. En este sentido, la validez asegura la consistencia, exactitud y capacidad de predicción de los resultados. En otras palabras, cuando mide lo que dice medir.

Una medida es confiable cuando la regla de medición es adecuada, libre de errores aleatorios. En otras palabras, cuando mide bien lo que dice medir.

Un ejemplo clásico para ilustrar estos dos principios es la medición de estatura. Estaremos utilizando una medida válida de estatura si hablamos de distancia. No utilizaremos un instrumento válido si utilizamos como medida de estatura una balanza.

Por otra parte, tendremos una medida confiable si utilizamos un metro cuyos centímetros tengan la misma distancia entre ellos y correspondan al sistema de medición estandarizado. Si estos centímetros miden distancias distintas, estaremos haciendo una medición poco confiable.

Otro ejemplo: para medir la exposición a los medios de comunicación, una medida válida sería aquella que registre la conducta real de exposición. Para medir exposición, con frecuencia se pregunta qué tipo de programas o géneros prefiere. Se sabe que en muchos casos las preferencias no corresponden con las conductas reales: los programas culturales y noticiosos adquieren más *rating* en las preguntas de preferencia, que lo que en realidad alcanzan cuando son medidos conductualmente. En este sentido, una pregunta de preferencia para medir exposición sería poco válida.

Siguiendo con el mismo ejemplo. Al registrar la exposición a los medios no es lo mismo ver completo un programa de una hora que sólo unos minutos y cambiar constantemente de canal. Con frecuencia, con el sólo hecho de haber observado unos minutos el programa se asume que la persona vio todo el programa. En este sentido sería una medición poco confiable. Para elevar la confiabilidad de la medida, el *rating* debería ponderarse por la cantidad de exposición al medio en cuestión.

NIVELES DE MEDICIÓN

Existen cuatro niveles de cuantificación o medición de conceptos. Esta clasificación es fundamental para entender el tipo de análisis y resultados que podemos obtener. Los niveles de medición son los siguientes:

- a) *Nominal*. En las mediciones nominales, la dimensionalización de conceptos se hace categorizando las posibles respuestas o comportamientos posibles. Al medir la recordación de un contenido publicitario utilizamos una medida nominal al categorizar las posibles respuestas: las referidas a la imagen o al producto, etcétera. En estas medidas:
- Los números representan nombres de categorías.
 - Los valores no indican magnitud.
 - Los valores no indican relación entre ellos.
 - No es posible realizar operaciones aritméticas y la manipulación estadística es muy limitada.
- b) *Ordinal*. En las mediciones ordinales el sistema de cuantificación del concepto responde a una secuencia u orden. Al preguntar sobre preferencia de marcas u opciones políticas, utilizamos una medida ordinal al registrar el orden de las preferencias. En estas medidas:
- Los números representan orden o secuencia.
 - Los valores no indican magnitud.
 - Los intervalos entre valores no indican distancia.
 - Las relaciones entre los números son de orden o secuencia.
 - La manipulación estadística es posible pero con restricciones.
- c) *Intervalar*. En las mediciones intervalares la cuantificación del concepto guarda un orden numérico que mantiene espacios iguales entre los valores numéricos y que puede tomar la operacionalización del concepto. La edad es una medida intervalar. Las actitudes, medidas con escalas de intervalos aparentemente iguales como *likert*, son también consideradas medidas intervalares. En estas medidas:
- Los valores representan orden y magnitud.
 - El intervalo entre valores es igual.
 - El cero no representa un cero absoluto.
 - Los valores representan relaciones de orden y magnitud.
 - La manipulación estadística es posible.
- d) *De razón*. En las medidas de razón la cuantificación del concepto, además de guardar una relación numérica, parte de un cero absoluto. Estas escalas no se utilizan en ciencias sociales.

TÉCNICAS DE MEDICIÓN

En encuestas, las preguntas estructuradas y las escalas son los instrumentos más usados. Sin embargo, existen otros cada vez más integrados a esta técnica. En este apartado revisaremos las principales herramientas de registro de los temas de interés de las encuestas, así como el proceso de construcción de cuestionarios.

Las principales técnicas de medición en encuesta son las siguientes:

- a) *Preguntas abiertas*. En este tipo de preguntas el encuestado expresa de manera libre su respuesta. No existe preclasificación de respuestas. Las principales ventajas de este tipo de preguntas son la riqueza de información que se obtiene, que se evitan los sesgos planteados por las hipótesis previas y que, en general, son más fáciles de formular. Las principales desventajas de este tipo de entrevistas son la poca comparabilidad de los resultados; respuestas no estructuradas con frecuencia demasiado abiertas; la profundidad de las respuestas depende de la insistencia del encuestador.

En general, las preguntas abiertas corresponden más bien a un estudio de carácter exploratorio.

- b) *Preguntas cerradas*. Éstas presentan a la persona encuestada las alternativas de respuestas. Las principales ventajas de este tipo de preguntas son el nivel de estructuración de las respuestas y la comparabilidad. Las principales desventajas son la dificultad para su formulación y el riesgo de dirigir las respuestas hacia las opciones preclasificadas.

En la formulación de las opciones de respuesta siempre es conveniente partir de un estudio exploratorio. Estas opciones de respuesta deben ser:

- Exhaustivas. Es decir, deben considerar todas las alternativas posibles.
- Excluyentes. Las respuestas posibles deben pertenecer a una sola categoría.
- Relevantes conceptualmente al estudio y al repertorio de respuestas del encuestado.
- Formuladas con el mayor nivel de medición.

- c) *Escalas multidimensionales*. Los conceptos, especialmente las actitudes, están compuestos a su vez por varios factores o dimensiones. En las escalas multidimensionales se miden todos los factores de un concepto a través de varias preguntas. El propósito final de estas escalas es agregar las puntuaciones de las preguntas y dimensiones individuales en un sólo índice que represente la posición del sujeto con respecto al concepto global.

- d) *Registros*. Con frecuencia, como parte del proceso de medición dentro de la encuesta, se utilizan los registros de comportamientos o presencia de ciertos estímulos o características dentro de la unidad estudiada. Este tipo de registro permite obtener información de manera directa de los temas estudiados.
- e) *Reportes*. Cuando la información buscada requiere ser registrada durante periodos amplios, con frecuencia se utiliza la ayuda del encuestado para obtener la información. Para facilitar el registro se le proporcionan formatos adecuados y se le instruye en la tarea.

Redacción de preguntas

Independientemente de que sean abiertas, cerradas o estén incluidas en una escala multidimensional, las preguntas constituyen el principal instrumento de medición de las encuestas. A continuación revisaremos algunos de los principales criterios que es conveniente observar en la redacción de preguntas:

- a) *Focalización*. Las preguntas deben estar formuladas directamente en relación al tema de interés. La falta de dirección o focalización produce respuestas ambiguas y fuera del tema de interés. Un ejemplo: si la intención es conocer la opinión de una comunidad sobre ciertas medidas políticas, la pregunta debe estar formulada en esta dirección. Con frecuencia, esta pregunta es formulada de manera genérica: ¿qué piensa sobre la situación de la comunidad?
- b) *Brevedad*. Las preguntas deben ser cortas. A medida que la pregunta se extiende se pierde la concentración del encuestado, incrementándose la probabilidad de dar respuestas aleatorias. En mercadotecnia, en las pruebas de conceptos, con frecuencia se presentan a los sujetos conceptos demasiado amplios. Generalmente, al terminar de leerlos, el encuestado no tiene una idea clara de qué tiene que contestar y termina por referirse a las partes de la lectura que tuvieron más énfasis.
- c) *Claridad*. Cuando las preguntas no son redactadas de manera clara se incrementa la posibilidad de que cada entrevistado interprete a su manera la misma pregunta. Cuando una pregunta no es clara es muy común que el encuestador tienda a reformularla en sus propias palabras. Para guardar claridad en una pregunta es conveniente cuidar el vocabulario y la estructura gramatical. Una estructura gramatical simple contribuye enormemente a la claridad de una pregunta.

- d) *Simplicidad*. Las preguntas deben estar referidas a un solo tema. Con mucha frecuencia se formulan preguntas de varios temas a la vez o que establecen considerandos dentro de la pregunta.

Una típica pregunta en este sentido sería: Considerando que en esta ciudad estuviera a la venta este producto y que usted lo pudiera comprar a un precio de x , ¿qué opinaría de que se incluyera esta promoción? En este ejemplo, es necesario que el entrevistado guarde varios datos en la memoria y que establezca varias relaciones. La pregunta seguramente será interpretada de manera distinta por cada respondente.

Otra pregunta típica que establece confusión en el entrevistado es cuando se presentan dos temas que nos parecen similares pero que en realidad no lo son: ¿Qué opina de la situación económica y política del país?

- e) *Relevancia*. La redacción de la pregunta y las alternativas de respuesta, además de relevantes conceptualmente, deben ser relevantes para el entrevistado y deben ser localizadas dentro de su marco de referencia.

Un ejemplo de falta de relevancia sería preguntar de manera indiscriminada ¿cómo usa el horno de microondas?

- f) *Criterios de respuesta definidos*. Si una pregunta no presenta criterios claros de respuesta, la información obtenida será ambigua y no comparable. Si preguntamos qué opina sobre un tema y no proporcionamos una escala de respuesta, es muy probable que algunos den su opinión cuantificando su grado de acuerdo, quizá otros más expliquen las causas y otros más describan la situación. Si en cambio preguntamos: ¿De acuerdo a esta escala que contiene las opciones muy bien, bien, regular y mal, cuál es su opinión con relación al tema x ?, obtendremos el tipo de respuesta que esperamos.
- g) *Sin dirección*. Si bien es conveniente presentar los criterios de respuesta de manera clara y exhaustiva, ni éstos, ni la pregunta, deben dirigir la respuesta. Cuando en la pregunta se incluye algún tipo de adjetivación se sugiere la respuesta. Un ejemplo de esta pregunta, recientemente escuchado en un programa de radio que pretendía hacer una encuesta entre sus escuchas fue: ¿Qué opción prefiere usted?, ¿la formación de un cuerpo especial de policía para reducir la violencia en la ciudad o que no se forme?

Identificación de sesgos del respondente

Además de los sesgos causados por la redacción, existen otro tipo de sesgos producidos por el entrevistado, pero que también pueden ser evitados con una construcción adecuada. A continuación se listan estos sesgos:

- a) *Deseabilidad social*. Con mucha frecuencia las respuestas del respondente reflejan, más que una conducta, un deseo o aspiración. Al responder sobre el tipo de revistas que acostumbra leer o la marca del producto que acostumbra consumir, es importante distinguir si esta respuesta es una realidad o refleja una aspiración.
- b) *Sesgo de cortesía*. En ocasiones, los encuestados entienden la situación de encuesta como un compromiso y tratan de responder complaciendo al encuestador. Inclusive al dar su respuesta, ellos mismos interrogan sobre lo adecuado de su respuesta.
- c) *Tendencia a repetir repuestas*. Cuando las preguntas no son claras o el cuestionario es amplio, existe la posibilidad de que el encuestado adquiera un patrón de respuesta sin meditar sus contestaciones. En estas ocasiones, típicamente contesta con la misma frase: de acuerdo, de acuerdo, de acuerdo.
- d) *Prestigio*. Cuando el tema de la pregunta está relacionado con un tema autodescriptivo, con frecuencia el respondente tiende a exagerar y a presumir de su persona. Este tipo de exageraciones son usuales en estudios de estilo de vida.
- e) *Hostilidad/tabú*. Cuando los temas tratados representan angustia o temor para el encuestado, existe el riesgo de dar una respuesta disfrazada o modificada. Las encuestas de salud y hábitos sexuales están altamente expuestas a este riesgo.
- f) *Patrocinio de la encuesta*. Cuando el encuestado identifica quién es el patrocinador o *cliente* puede modificar sus respuestas, dependiendo de su grado de simpatía o antipatía hacia la institución.
- g) *Planteamiento mental*. Frecuentemente, el respondente establece un sistema de pensamiento para contestar todas las preguntas de un cuestionario. Aunque se utilicen preguntas y escalas diversas, el entrevistado mantiene una forma de entender y contestar las preguntas.
- h) *Polarización de respuestas*. Por una tendencia natural, los respondentes tienden a polarizar sus respuestas, inclusive cuando las escalas permiten posiciones intermedias. De esta forma, las respuestas se reducen a opiniones positivas o negativas.

Para evitar los sesgos del respondente, además de observar las recomendaciones de redacción de preguntas, es conveniente realizar los siguientes ejercicios:

- Pilotear el cuestionario.
- Ampliar las instrucciones.
- Modificar la redacción de preguntas.

- Cambiar las escalas de respuesta.
- Modificar la estructura y el orden del cuestionario.

Etapas de construcción del instrumento

En encuestas, el instrumento más usado es el cuestionario. En este apartado se presenta el proceso de construcción, que involucra los siguientes pasos:

- Identificación de objetivos del estudio.* El primer paso para construir un instrumento es identificar y revisar los objetivos específicos del estudio. Éstos determinarán el enfoque, amplitud y tratamiento de todo el instrumento.
- Desglose de objetivos en dimensiones e indicadores.* Una vez identificados cada uno de los objetivos del estudio, el siguiente paso es trabajar por separado las dimensiones e indicadores de cada uno de ellos.
- Elaboración de secuencia.* Habiendo redactado las preguntas y escalas que corresponden a cada uno de los objetivos, el siguiente paso es organizar la información en una secuencia lógica. Es sumamente importante que la información obtenida tenga orden y unidad. Cambiar de tema sin orden produce confusión y enojo en el encuestado, además de reducirse considerablemente la contabilidad del instrumento. Algunos principios a seguir en la secuenciación de un cuestionario son los siguientes:
 - De lo general a lo específico.
 - De lo simple a lo complejo.
 - De lo espontáneo a lo ayudado.
 - De lo público a lo privado.
 - Por tópicos.
 - Por forma de respuesta.

Típicamente un cuestionario tiene la siguiente estructura:

- Introducción y solicitud de colaboración.
- Selección o filtro del respondente.
- Preguntas relacionadas con objetivos.
- Preguntas demográficas.
- Agradecimiento e informe sobre posible supervisión.
- Datos de identificación, como nombre y dirección, para posterior supervisión.

- d) *Redacción de preguntas.* Habiendo definido la secuencia de exploración, el siguiente paso consiste en redactar las preguntas y reactivos de las escalas. En este paso es necesario ensayar varias operacionalizaciones y construcciones, observando siempre las reglas de redacción de preguntas.
- e) *Identificación y redacción de alternativas de respuesta.* Después de redactar las preguntas, el siguiente paso consiste en elaborar las categorías de respuesta. Este paso puede ser realizado de manera simultánea a la redacción de las mismas preguntas. Las categorías de respuesta deben observar las recomendaciones listadas.
- f) *Identificación de pases y dependencias.* Enseguida es necesario identificar las dependencias entre preguntas y explicitarlas a través de instrucciones precisas.
- g) *Redacción de apoyos y tarjetas.* Casi todos los cuestionarios utilizan apoyos visuales como tarjetas con escalas de respuestas o redacción de conceptos. En esta etapa se diseñan y redactan estos estímulos de apoyo.
- h) *Redacción de instrucciones.* El último paso es redactar las instrucciones generales para el encuestador. Éstas deben ser detalladas y amplias.

Recolección de la información

Las técnicas de recolección de información son clasificadas considerando dos criterios: la forma de aproximación a la unidad estudiada y la técnica de registro de la información.

De acuerdo a la forma de aproximarse al respondente, existen tres técnicas:

- a) *Personales.* En las encuestas personales, el encuestador establece comunicación cara a cara con el respondente. Esta técnica puede ser aplicada en distintos ámbitos. Estos son:
 - Casa por casa. Cuando la entrevista es aplicada en el hogar o ámbito natural del respondente.
 - Centros de afluencia. Cuando la entrevista es aplicada donde existen concentraciones de personas que cumplen con las características buscadas. Los centros de afluencia son generalmente los centros comerciales o lugares de alto tráfico, como terminales de transporte o centros públicos.
 - Intercepciones. Cuando se requiere encuestar a una persona que haya realizado una actividad específica como la compra de un producto

o votado. La mejor opción es interceptarla inmediatamente después de haber realizado la actividad. Esta técnica se utiliza con frecuencia para conocer indicadores de consumo y electorales.

- Concentraciones. Este tipo de encuesta es poco frecuente. Se utiliza para medir reacciones frente a actos públicos como conciertos o manifestaciones.

b) *Telefónica*. En este tipo de encuesta, la comunicación se establece a distancia por medio de esta tecnología. Esta técnica de entrevista ha crecido rápidamente.

c) *Correo*. En las encuestas por correo la comunicación se establece de manera escrita y generalmente de manera impersonal.

A continuación se presenta una tabla comparativa de los métodos de recolección de información de acuerdo a la forma de aproximarse al respondente.

Análisis comparativo de los métodos
(calificación: +++ muy alta, --- muy baja)

	Casa por casa	Centro de afluencia	Intercepciones	Teléfono	Correo
Tamaño del cuestionario	++	--	--	+	+
Facilidad de levantamiento	--	++	+	-	+++
Costo de levantamiento	+++	+	+	++	---
Representatividad estadística	+++	-	-	++	---
Representatividad de grupos específicos	+	++	++	++	-
Velocidad de levantamiento	-	++	++	+++	---
Posibilidad de mostrar estímulos/Prod. comerciales	++	+++	++	---	+
Riesgo de sesgo provocado por el encuestador	+++	++	++	-	---
Nivel de cooperación de los entrevistados	-	++	++	+	---
Supervisión-validación de la información	+++	+	+	+++	-
Anonimato	---	++	+	+	--

En cuanto a la forma de registro existen básicamente dos técnicas. La tradicional de lápiz y papel y la automatizada.

El registro a través de lápiz y papel consiste en anotar o registrar las respuestas en un formato estructurado. El registro puede ser hecho por el encuestador o por el mismo respondiente.

El registro automatizado de información tiene dos versiones: CATI (Computer Assisted Telephone Interview) y CAPI (Computer Assisted Personal Interview). En ambos sistemas la información es registrada de manera electrónica, a través de pantallas de captura diseñadas especialmente para esta tarea.

BIBLIOGRAFÍA BÁSICA

ALRECK, P. y R. Settle, *The survey research handbook*, Irwin Professional Publishing, New York, 1995.

CAMPBELLS, D. y J. Stanley, *Diseños experimentales y cuasi experimentales en la investigación social*, Amorrortu Editores, Buenos Aires, 1979.

ESOMAR, *Interviewing for market and opinion research*, ESOMAR, Amsterdam, 1995.

GARZA, A., *Manual de técnicas de investigación*, El Colegio de México, México, 1972.

KALTON, G., *Introduction to survey sampling*, Sage Publication, California, 1983.

KERLINGER, N. F., *Investigación del comportamiento: técnicas y metodología*, Interamericana, México, 1975.

KIMMEAR, T. y James Taylor, *Investigación de mercados*, Mc Graw Hill, México, 1981.

MCGUGAN F. J., *Psicología experimental*, Editorial Trillas, México, 1971.

PADUA, J., *Técnicas de investigación aplicadas a las ciencias sociales*, FCE, México, 1979.

PARDIÑAS, F., *Metodología y técnicas de investigación en ciencias sociales*, Siglo XXI Editores, México, 1979.

RAJ, D., *La estructura de las encuestas por muestreo*, FCE, México, 1972.

REYNOLDS, P., *A primer in theory construction*, The Bobbs Merrill Company, Indianapolis, 1971.

ROJAS Soriano R., *Guía para realizar investigaciones sociales*, UNAM, México, 1982.

_____, *Métodos para la investigación social. Una propuesta dialéctica*, Plaza y Valdés, México, 1992.

SELLITZ, C., M. Jahoda, M. Deutsch y S.W. Cook, *Métodos de investigación en las relaciones sociales*, Ediciones Rialp, Madrid.

Lectura 3

Grupos de discusión De la investigación social a la investigación reflexiva

Bernardo Russi Alzaga

La reflexión es una tarea de vagos y maleantes.
Hay que saber perderse para trazar un mapa,
salir de los caminos trillados, vagar: deambular
por las encrucijadas, abrir senderos a través de
las mieses o el desierto, penetrar en callejuelas
sin salida, asumir que todo camino recorrido
sin mapa es caótico (luego será posible tender
o recoger puentes, bordear pozos o simas,
perforar agujeros o taparlos).

Jesús Ibáñez

Presentación

PRESENTAR UNA TÉCNICA DE INVESTIGACIÓN SOCIAL quiere decir, entre otras cosas, proponerla para su aprendizaje, para su uso y, también, para su crítica. Repasar y reflexionar cómo se usa y cómo se ha aprendido. ¿Desde dónde es que hablo de esta técnica? ¿Desde dónde es que hablo de ella y cómo interfiere este *desde dónde* al tratar de explicarla y revisarla?

Antes que nada prefiero abordar esta reflexión desde una forma amplia, es decir, no sólo es *la técnica*, sino también las diversas formas como se puede entender y usar. Entonces, hay que hablar de *un paquete técnico*. El usar la idea de *paquete técnico* obliga a presentar una revisión sobre los distintos usos –en las diferentes facetas– que tiene una técnica. No sólo es abordar la técnica puesta en escena, con todo lo que ello implica, sino también, tratar de presentar el abanico de posibilidades que pudieran aparecer cuando se le pone en escena.

Es aquí donde aparece el primer problema para una mejor presentación del paquete. Lo que sucede es que cuando se habla de la puesta en escena de una técnica (por lo menos en investigación social) se habla desde la parte que se conoce/desconoce y se usa de ella. Luego, se infiere a lo que otros han dicho, hecho y desecho. La escena se ve afectada por el escenario y viceversa. El observador –investigador– usuario de la técnica se ve afectado por la escena y el escenario

Tomada de: Bernardo Russi, "Grupos de discusión. De la investigación social a la investigación reflexiva", en Luis Jesús Galindo Cáceres (coord.), *Técnicas de investigación en sociedad, cultura y comunicación*, México, Pearson, 1998, pp. 75-115.

donde la utiliza y, a su vez, la escena y el escenario se verán afectados por la forma en como el observador observa y usa la técnica.¹

Dado lo anterior, preferiría nombrar la técnica/paquete técnico como *sesión de grupo*. Lo propongo así porque creo que este nombre se refiere, principalmente, a la variedad que presentan las distintas formas como puede ser usada la técnica, es decir, cómo se configura en paquete. Le podríamos llamar *grupos de discusión* en tanto se pone en escena como técnica de investigación social en la forma como habla, describe y reflexiona Jesús Ibáñez². También se puede usar el nombre de *grupos de enfoque*³ si se tomaran en cuenta las formas y concepción de los investigadores de Estados Unidos. En el campo de estudios de mercado, por lo menos en México, a la primera forma le dicen europea y a la segunda norteamericana. Pero resulta que también cada *forma* de usar esta técnica, tiene sus *características*: en la forma europea el moderador interviene muy poco, usualmente el discurso que produce el grupo no tiene mucha *intervención*; en la forma norteamericana el moderador controla mucho la participación de cada participante, incluso se llega a pedir que conteste cada uno en forma aislada, sin diálogo ni interacción.

Un día me preguntaron que cuál era la *forma* en que yo trabajaba. Yo conteste que no sabía. Resulta que cada moderador puede adecuarse a un estilo, pero la metodología y el diseño que se construya para usar la técnica puede hacer que se varíen las formas de moderación. Realmente creo que hablar de una *forma de uso* es muy peligroso y no se aporta gran cosa.

La variedad que presenta la puesta en escena de las sesiones de grupo incluye aspectos relacionados con el uso de esta técnica en por lo menos dos sociologías: una de Estados Unidos y otra de España (que autodenominan de segundo orden). El nombre *grupo de discusión* es acuñado por esa sociología española, el nombre que le dan los sociólogos que la desarrollan en Estados Unidos es *focus group* o *grupos de enfoque*.

¹ Si consideramos que el que aquí escribe se dedica a ser observador (investigador) usuario de la técnica, entonces habría que aclarar que este *oficio* se lleva a cabo en estudios de diversos tipos: en investigaciones de mercado, estudios de audiencia y comunicación política y diagnósticos de necesidades sociales para proyectos educativos como el diseño y lanzamiento de programas de posgrado. Digamos que estos son los campos de acción donde aplico la técnica y, por lo tanto, la experiencia desde donde se habla.

² Jesús Ibáñez (1928/1992), es uno de los mejores representantes de la sociología crítica española o de *segundo orden*, como el mismo la llamaba. Además, Ibáñez es de los pocos autores que ha escrito y publicado textos en torno a la técnica grupos de discusión. Fue fundador, a lado de Alfonso Ortí y Ángel de Lucas, de diversos espacios para reflexionar temas que han enriquecido el estudio de las ciencias sociales y las diversas metodologías que actualmente se utilizan; finalmente, fue mi profesor y pude aprender de él demasiadas cosas, no sólo de lo académico y la metodología, sino también de la vida.

³ Traducido del inglés *focus group*.

Por lo tanto, se tiene que aclarar que para hablar en términos de paquete técnico, tendrían que contemplarse los elementos que lo conforman desde la dos visiones sociológicas que se conocen y que lo han desarrollado y trabajado más. De antemano, se tendría que plantear que en otros países también se trabaja con este paquete. Pero que tanto los españoles como los estadounidenses han publicado algunos textos que nos permiten acercarnos, en el plano teórico-metodológico, a él. Y creo que las razones de que se conozcan estas ediciones son dos: la cercanía física con los Estados Unidos (y en concreto con su estilo de trabajo en la investigación social en general y de mercados en particular) y la cercanía lingüística con los textos y experiencias españolas (sobre todo en investigación social).

Entonces, cuando hablo de esta técnica tengo que hacerlo de dos maneras: desde el referente más cercano que se tiene por la forma en que se experimentó y conoció (más directa o vivamente, al poder compartir ideas con algunos de sus principales conocedores - Jesús Ibáñez, Alfonso Ortí y Ángel de Lucas-y, por otro lado, desde lo que se ha experimentado y conocido únicamente vía lecturas y algunos trabajos de investigación aplicada en diversas áreas.

Marco teórico/metodológico

Las sesiones de grupo como paquete técnico y la investigación cualitativa

A continuación presento algunas de las principales características de las sesiones de grupo como paquete y como técnica propuesta desde la metodología cualitativa y el paradigma en el que la inserta la sociología crítica española.

Trabajar en el campo de la investigación social implica, entre otras cosas, producir discursos. La técnica es la herramienta de la que se hecha mano para producir el discurso. En este apartado, entonces, hablaremos sobre una técnica que tiene sus peculiaridades al generar su discurso correspondiente. Se trabaja sobre la reducción crítica de los contenidos, mismos que se producen en un discurso grupal, que reproducen y rebordean el sentido, poniendo en juego toda su extensión, el nivel del habla, a fin de permitir que la presión semántica configure el tema del que en cada caso se trate, como campo semántico (como campo, por tanto, de sentido).

El manejo de esta técnica también incluye, si se quiere ver así, el trabajo con información. Información en forma de discurso: flujo de mensajes que circulan entre los miembros de un grupo que pone en forma de habla sus representaciones de la realidad, las exterioriza en un espacio físico-temporal. Esto puede decir muchas cosas a los distintos tipos de investigadores de lo social. Desde luego, a aquellos que trabajan con el discurso podrían apreciar de forma muy particular el valor

de esta técnica, sobre todo en la manera de apropiarsela: antropólogos, lingüistas, sociólogos, historiógrafos, etcétera. Pero me quiero referir en particular a los investigadores, o los que están en proceso de serlo, en el campo de la comunicación. ¿Qué puede significar un paquete técnico como éste, en el que la materia prima es la información, para los estudiosos de la comunicación y de los procesos de producción de información? Así lo dejo, abierto, pues creo que se puede generar un debate para ver cuál sería la mejor forma de apropiarnos de esta técnica, en este campo específico de lo social.

La técnica sesiones de grupo no se propone como una herramienta única, o como dice Ibáñez "...no implica el encontrar una forma canónica de proceder..." (1992, p. 262). Un objeto de estudio social se puede abordar desde distintas formas y procedimientos, diferentes perspectivas, etcétera. Y para acercarse a ese pedazo de realidad y preguntarle cosas, también se puede abordar con un sin fin de herramientas. El investigador debe decidir cuál. Debe de explicar por qué esa o esas técnicas y contar como lo hizo. Lo debe hacer consciente.⁴

¿Cómo hacer consciente una técnica?, ¿cómo sacar de la oscuridad no sólo los datos, sino los mecanismos que configuran metodologías o estrategias para abordar los objetos de estudio? En otras palabras, ¿cómo son los procesos de objetivación y cómo interviene la subjetividad del sujeto al mirar y estudiar el objeto?

También habría que mencionar el carácter cualitativo del paquete técnico. De entrada, aclaro que no es mi intención oponer el concepto cuantitativo contra el cualitativo. Las perspectivas meteorológicas cualitativa y cuantitativa encierran una larga lucha, Fernando Conde⁵ considera *insuficientes* las formas en que se ha abordado su relación y, en particular, creo que habría que apostar por revisiones que ayuden a mejorar las formas de complementariedad de una con la otra (Ibáñez y Beltrán hablan de pluralismo metodológico⁶).

⁴ "La descripción de las técnicas no es un repertorio de prescripciones, no es una descripción normativa. Describir la técnica es estructurar como un sistema de transformaciones el conjunto de elementos operantes en el contexto existencial en la situación de producción. La técnica aparece como un instrumento manipulable a conciencia." (Ibáñez, 1992, p. 262.)

"Las técnicas nos ayudan a formalizar diversos recortes informativos de la realidad, *datos*, sí, pero fatalmente *embarazadas* de interpretación. En la oscuridad todos los datos o son negros o son obvios, sobre todo cuando solemos pensar que nada tienen que ver con la teoría y nuestros propios criterios de descripción." (González, 1993, p. 211.)

⁵ "...los propios términos *cualitativo* y *cuantitativo* tienen una larga y todavía no cerrada historia filosófica, científica, etcétera, ... las relaciones citadas deben enmarcarse de una forma más amplia en los procesos sociales e históricos de construcción de los sujetos/objetos sociales y en la propia historia de los lenguajes/ciencias/metodologías que se gestaron de forma paralela a dicha construcción." (Conde en Delgado y Gutiérrez, 1994, p. 53.) Véase la recomendación de este texto en el apartado correspondiente.

⁶ "...si a la complejidad del objeto corresponde necesariamente un planteamiento epistemológico que he venido calificando de pluralismo cognitivo, ello impone, como correlato necesario, un pluralismo metodológico que permita acceder a la concreta dimensión del objeto a la que en

Las formas que presenta la realidad social, en cuanto a su dinamismo y complejidad, no nos dejan abordarla con una metodología cuantitativa o una cualitativa, por muy complejas que estas sean. Más bien, la realidad requiere de metodólogos que diseñen estrategias que impliquen una *tercer cultura*⁷ de investigación, que desarrolle con imaginación y creatividad el potencial que cada una de las metodologías puede dar si se usan en forma combinada.

Sin embargo, se tendría que precisar el carácter cualitativo de esta técnica. Estaría de acuerdo con Beltrán al señalar que dada la posibilidad de poder trabajar con el lenguaje, implica la necesidad de establecer identidades y diferencias en el mismo. Este es un punto que parece lo aleja de lo cuantitativo y esclarece que la materia prima con la que se trabaja es meramente cualitativa.

El tipo de trabajo que se logra con una sesión de grupo es algo parecido a una simulación. Ibáñez explica cómo, a través de un grupo de discusión, se verifican supuestos básicos, ya que “su realidad está irremediablemente cortada de su posibilidad, no es más que un cebo.”⁸ Asimismo, este autor menciona cómo un grupo, al ser un simulacro de pedazo de realidad, se parece a la forma de consumo en las sociedades de hoy: “El grupo de discusión como objeto de consumo tiene la misma estructura que los productos cuyo consumo se investiga mediante él; es un simulacro, sólo signo de sí mismo.” (Ibáñez, 1992, p. 257.)

Mencionado lo anterior, podemos agregar también que las sesiones de grupo funcionan mejor siendo aplicadas en sociedades de consumo, es decir, espacios sociales que implican un desarrollo bastante fuerte del mercado capitalista: los miembros de ese espacio están acostumbrados a rutinas de consumo que este mercado marca. Así, los patrones de comportamiento también se ven afectados –además de la necesidad que hay que diseminar– por un conjunto de símbolos que represen-

cada caso haya de hacerse frente.” (Beltrán en Ferrando, Ibáñez y Alvira, 1992, p. 44.) En la misma obra, Ibáñez escribe un texto llamado “Perspectivas de investigación social”, en el que señala que las diferentes técnicas (entre ellas los grupos de discusión) pueden cumplir con los requerimientos que tenga el investigador –unas mejor que otras–. Hay que considerar la perspectiva que se está trabajando –distributiva, estructural y dialéctica– para ver donde se está colocado para abordar el objeto y, también, esto permite poder diseñar una *doble pinza* (mínimo dos técnicas) para poder utilizar *dispositivos* que permitan producir información del y sobre el objeto de estudio.

⁷ “Autores como Allan, Maturana, Varela, Von Foerster, Prigogine, Thom, Petitot, MacClintok y un largo etcétera, significan una relectura del viejo paradigma, que contemple e integre lo que el anterior dejaba de lado, es decir, la vida, la complejidad, la organización, la producción, los procesos, lo real y lo social con la necesaria modernización de los fenómenos. (Conde en Delgado y Gutiérrez, 1994, p. 67.)

⁸ Según Ibáñez, se trabaja con el principio de: 1) *dependencia*, que se refiere a la forma de dependencia que tiene el grupo con el moderador y cómo éste trabaja con el discurso; 2) *aparición*, que simula la posibilidad de acción común que nunca se llevará a cabo, a partir de la *cópula predicativa* que realiza el grupo mediante su discurso; 3) *ataque y fuga*, que representa la posibilidad de participar en una sesión donde hay otros como yo, o la no participación, quedando aislado y separado del resto en forma individual. (Ibáñez, 1992, p. 257.)

tan valores adicionales a los cuales el consumidor raramente se resiste (publicidad).⁹

Desde luego que no se está proponiendo que el capitalismo, en su fase consumista, es el que ha creado los grupos (la cita seis nos da un ejemplo). Pero, sí es muy importante entender la propuesta que hace Ibáñez sobre la metamorfosis histórica que han tenido las sociedades capitalistas hasta llegar a lo que hoy llama capitalismo de consumo. En este contexto, podemos ver no sólo la función de los grupos en tanto consumistas, sino también en partidos políticos y grupos religiosos, es decir, todos aquellos espacios a los cuales se puede *pertenecer*.¹⁰

Ante esta posición de Ibáñez, quisiera hacer una propuesta que nos permita ampliar este tema. Para hablar de la configuración del espacio social en México habría que tomarse en cuenta, también, un campo que últimamente ha tenido un fuerte desarrollo en materia de investigación.

Sólo quisiera apuntar lo valioso que se vuelve indagar lo social si tomamos en cuenta la cultura como campo que se fusiona a los objetos que observamos en ciencias sociales. La propuesta que hace la Red FOCYP, constituida por varias universidades, con su eje central en la de Colima, nos aporta elementos para entender la cultura y algunos campos que la configuran y relacionan con el espacio social mexicano, en el que se mueven estos consumidores, miembros de partidos políticos e iglesias a los que nos hemos referido.¹¹

También, desde el estudio de la cultura, Galindo propone la percepción del individuo como eje de entendimiento de los caminos que sigue un sujeto o un grupo

⁹ Ibáñez utiliza la siguiente cita para explicar, todavía más, lo que las sociedades de hoy hacen con el individuo y cómo esto se utiliza para explicar el cómo y por qué funciona la técnica de grupos de discusión: Bradford (1963, p. 12.) "Hemos sido adiestrados para enfrentar nuestros problemas individuales, pero tropezamos con problemas demasiado complejos para algunos de nosotros, como para que podamos resolverlos solos. Quedamos como paralizados frente la avalancha de necesidades crecientes de decisión y acción que hoy en día debe afrontar el adulto, y carecemos de las habilidades que se requieren para llegar a decisiones conjuntas, tal como éstas deberían tomarse." (Ibáñez, 1992, p. 218.)

¹⁰ "La sociedad de consumo no ha inventado los grupos: el acoplamiento de los sujetos individuales al orden social, *el componente libidinal del vínculo social* siempre ha sido grupal, siempre ha consistido en la captación de los fantasmas individuales por un aparato grupal; pero, por primera vez, este acoplamiento grupal es manejado a conciencia. Lo que se manifiesta en la infraestructura y supraestructura: en la infraestructura, pues las situaciones objetivas de acoplamiento al orden de la producción y el consumo son cada vez en mayor medida manifiestamente grupales; en la supraestructura, pues la publicidad y la propaganda ofrecen cada vez más manifiestamente como objeto para el deseo los grupos (el objeto de amor es cada vez más la marca, el emblema, la enseña de pertenencia a grupos)." (Ibáñez, 1992, p. 251.)

¹¹ FOCYP significa Formación de Ofertas Culturales y sus Públicos y es coordinado por el doctor Jorge González, en el Programa de Estudios de la Cultura del Centro Universitario de Ciencias Sociales (CUCS) de la Universidad de Colima; FOCYP propone ocho campos para entender una posibilidad de configuración cultural en nueve ciudades de México, a lo largo del siglo XX: educación, cultura legítima, salud, diversión, educación, medios de comunicación, religión y dinero.

de individuos dentro de su propio espacio social. Hace una propuesta de categorías al que llama “un primer esquema de hipótesis configuradoras del México de hoy.”¹²

Con esta idea, el autor plantea cómo es que se podría hablar de una sociedad de información y una de comunicación. Poner a los individuos y a los grupos bajo esta mirada, implica desarrollar nuevas formas de proceder con respecto a la técnica que aquí nos ocupa, o de cualquier otra. ¿De qué manera se consume, se milita y se reza en una sociedad que incluye tecnología de alto desarrollo?, ¿cómo repercute esto en los objetos de estudio y en la forma de percibir del investigador/observador?

Sesiones de grupo como dispositivo conversacional

Un sujeto individual (no social) no sometido a una situación discursiva, tan sólo ofrecerá enunciados en los que las relaciones semánticas se expresan de modo fragmentario –hablar en grupo (conversar) o hablar con uno mismo–; lo que se pretende es que emerja un campo semántico en toda su extensión.¹³

Las hablas individuales tratan –entre sí– de acoplarse al sentido (social). Es así que se puede hablar del consenso, de la búsqueda de éste en la dinámica grupal, después de la transferencia que hace el moderador hacia el grupo mismo, que se pone a trabajar sobre lo que el grupo dice, y no le dice lo que deben decir (no pregunta). Hay que colocarse en una posición estratégica, en forma de encuentro, no de búsqueda. Así, el grupo es una fábrica de discursos que hacen uno solo, el del mismo grupo. Este discurso se produce a través de discursos individuales que *chocan* y se escuchan, y a su vez, son usados por los mismos participantes en forma cruzada, contrastada y enfrentada.¹⁴

¹² Galindo pone al descubierto cómo miró y, por lo tanto, cómo construyó su propuesta; “La percepción del país se ordenó en diez ámbitos de configuración de la acción y del sentido: economía, política, religión, cultura, información, historia y vida cotidiana, institución y emergencia, identidad y alteridad, género y lo público y lo privado. (Galindo, 1995)

¹³ “Se llama *situación de discurso* el conjunto de las circunstancias, en medio de las cuales se desarrolla un acto de enunciación (escrito u oral). Tales circunstancias comprenden un entorno físico y social en el que se realiza ese acto, la imagen que tienen de él los interlocutores, la identidad de estos últimos, la idea que cada uno se hace del otro (e incluso la representación que cada uno posee de lo que el otro piensa de él), los acontecimientos que han precedido el acto de enunciación (sobre todo las relaciones que han tenido hasta entonces los interlocutores y los intercambios de palabras donde se inserta la enunciación).” (Ducrot, “Diccionario enciclopédico de las ciencias del lenguaje”, en Ibáñez, 1992.)

¹⁴ “Se trabaja con la *enunciación*, se asume como punto crítico, en el que lo social se reproduce y cambia, como el objeto, en suma, de las ciencias sociales. En toda habla se articula el orden social y la subjetividad.” (Ibáñez, 1992, p. 289.)

La primera vez que un grupo de alumnos escuchó que debían realizar una breve investigación para comenzar a aprender la técnica, sintieron tal inseguridad que tuvo que comparar este trabajo con el de un artesano que comienza desde lo que conoce en teoría, desde lo que le han contado (o leído), desde lo que ha visto (práctica del otro). La posibilidad de dominar una técnica, es la posibilidad de dominar un oficio artesanal que, de entrada, jamás será dominada en su total complejidad. Más bien, lo que puede suceder es que el propio aprendiz/investigador vaya configurando el propio conocimiento de la técnica y el uso de ella a través de lecturas y prácticas.

En general, cuando se tiene que abordar la didáctica de la investigación, uno se enfrenta a prácticas demasiado cerradas.

La investigación que se hace con el paquete técnico de sesiones de grupo es compleja. El mismo paquete plantea una gran variedad de posibilidades de trabajo, como el pintor que tiene enfrente el lienzo y lo puede abordar con distintos tipos de técnicas. El investigador elige, decide y tendrá que justificar por qué y cómo lo hizo, reconocerá los alcances, pero también las limitantes:

La técnica tiene una dimensión grupal. Si el discurso social se haya diseminado en lo social mismo, el grupo de discusión equivaldrá a una situación discursiva, en cuyo proceso este discurso diseminado se reordena para el grupo: situación de grupo igual a situación discursiva.

Si el universo del sentido es grupal (social), parece obvio que la forma del grupo de discusión habrá de adaptarse mejor a él, que la entrevista individual, por abierta (o profunda) que sea. La reordenación del sentido social requiere de la interacción discursiva, comunicacional.

El moderador hace una propuesta, el grupo la fija y la ordena, según criterios de pertinencia, para poder así constituir el sentido social correspondiente al campo semántico concreto en el que se inscribe dicha propuesta.

¿Por qué se habla de un moderador?, ¿por qué y cómo *fija y ordena* (la retina) el grupo?, ¿qué son los criterios de pertinencia? Se trata de búsqueda de consensos, de discensos, se revelan deseos y necesidades, creencias, información preconstituida en la mente del individuo que contrasta con la de los demás, todo lo anterior proviene del aprendizaje cotidiano, de las vivencias, de lo percibido en la experiencia. Dentro del juego discursivo e interactivo se contrastan representaciones sociales e individuales, que encuentran su referente en la realidad social.¹⁵

El sujeto parlante es dueño de sus opiniones, pero no de la estructura que las genera. Por eso, el orden social no es consciente o se conoce fragmentariamente algunas de sus partes (requisito para su funcionamiento, como es requisito que

¹⁵ Manuel Martín Serrano explica muy bien el cómo se construyen las representaciones sociales de lo que acontece en el entorno, lo explica enfocando, principalmente, el papel que juegan los medios de comunicación de masas en las sociedades contemporáneas. (Martín S., 1986.)

desconozcamos lo que decimos para que el lenguaje pueda seguir funcionando en nosotros).

El grupo de discusión es un artificio metodológico que reúne diversas modalidades de grupos, en una articulación específica, es un dispositivo conversacional.

¿A qué se parece un grupo de discusión? Conversación grupal natural, grupo de aprendizaje como terapia psicológica, un foro público. No es ninguno de los anteriores, pero simula todas estas formas a la vez.

De hecho, hablar de un paquete técnico, es hablar de la gran posibilidad de variantes en el uso de la técnica. Esta es otra razón para hablar de una técnica en *paquete*, cada observador/investigador lo abordará de manera diferente, respetando las reglas meteorológicas mínimas, pero siempre estará en una situación abierta y estratégica para cambiar y dar giros cuando sea necesario.

Las perspectivas de investigación, un acercamiento al paradigma de investigación social de la sociología crítica española ¹⁶

Bajo esta mirada, se plantean tres niveles que diferencian los tipos de investigación en cuanto a su diseño. Generalmente la palabra *diseño* remite a una operación tecnológica¹⁷ y se cuestiona el por qué un investigador selecciona, en su diseño, tal o cual técnica de investigación o método de trabajo.

El investigador social suele elegir, sin pensar demasiado en la elección, la técnica que tiene más a mano: bien por razones personales (uno es experto en esa técnica), bien por razones organizativas (uno trabaja en una organización constituida para trabajar con esa técnica), bien por razones institucionales (uno pertenece a una institución interesada en vender esa técnica). (Ibáñez, 1992, p. 49.)

La tecnología nos da razón de cómo se hace, este sería el nivel tecnológico que plantea este paradigma, y aquí se sitúa la técnica con la que se trabaja y produce información. Pero para llegar a este nivel, primero hay que plantear los problemas del por qué se hace así, nivel metodológico; y, también, habría un nivel anterior a éste, que es el epistemológico, es decir, para qué o para quién se hace.

¿Dónde están situadas las diferentes técnicas y cómo las concibe el investigador? La pregunta sugiere relacionar cada técnica con la posibilidad metodológica

¹⁶ La propuesta tecnológica (así la denomina Ibáñez), la cual enmarca una perspectiva de trabajo en la investigación, está inscrita en la propuesta paradigmática que hacen algunos integrantes de la sociología crítica española. (Alvira, Ibáñez, García, 1992.)

¹⁷ Del latín *designare* = indicar. Se refiere a alguien que *traza* y *diseñare* = dibujar. (Gómez, 1995, p. 228.)

con que se trabaja, y el sistema de conceptos que se manejan (desde un punto de vista epistemológico) para interpretar la información que se produce con la operación tecnológica (y, por lo tanto, con el uso de la técnica seleccionada).

Veremos que cada uno de estos tres niveles pueden ser *puntuados* de diferente manera al relacionarlos con las tres perspectivas que se plantean en este paradigma: la distributiva, la estructural y la dialéctica.

Se plantea que la perspectiva distributiva trabaja sobre todo en el nivel tecnológico (es empirista), la perspectiva estructural sobre el nivel metodológico (articula empirismo y formalismo) y la perspectiva dialéctica puntúa sobre todo en el nivel epistemológico (articula empirismo, formalismo e intuicionismo).

Hay una inversión al llegar al tercer nivel, la tecnología y la metodología se atienen a lo dicho y lo sabido –son positivas–, la epistemología persigue lo no dicho y no sabido –es negativa– (como la perspectiva dialéctica de la investigación social es negativa, frente a la positividad de la perspectiva distributiva y estructural). (Ibáñez 1992, p. 49.)

Ibáñez se remite a Bourdieu cuando éste señala las “tres operaciones necesarias para el dominio científico de los hechos sociales”: una conquista contra la ilusión del saber inmediato (epistemológica), una construcción teórica (metodológica) y una comprobación empírica (tecnológica). Ibáñez dice que Bourdieu se inspira en Bachelard (1949) “para quien el hecho científico se conquista, se construye y se comprueba.” (Ibáñez, 1992.)

Lo que se quiere plantear bajo este paradigma, es el cómo se diseña bajo la posibilidad de integrar los niveles (epistemológico, metodológico y tecnológico) con las tres operaciones (conquista, construcción y comprobación) y las tres modalidades (cómo, por qué y para qué o para quién).

Niveles	Operaciones	Modalidades
Epistemológico	Conquista	¿Para quién/para qué?
Metodológico	Construcción	¿Por qué...?
Tecnológico	Comprobación	¿Cómo?

EL NIVEL TECNOLÓGICO

Este nivel opera con la perspectiva distributiva, es decir, se trabaja con elementos. En el caso de la investigación social nos estaríamos refiriendo a individuos. La técnica que mejor representa a este nivel y a esta perspectiva es la encuesta estadística, ya que su operación básica está dentro de la perspectiva distributiva.

EL NIVEL METODOLÓGICO

A este nivel se le ubica más que nada en la perspectiva estructural. En este sentido, lo que se observan son las relaciones que se dan entre los elementos, o relaciones entre individuos o grupos de individuos que nos reflejan las estructuras sociales. La técnica que mejor representa esta perspectiva, es la de grupos de discusión o, en su caso, el paquete técnico que trata este apartado.

Las relaciones entre los sujetos de la sociedad se da a través de la comunicación, y ésta, a su vez, se puede encontrar en la conversación. La técnica de grupos de discusión es un dispositivo conversacional, se trata de llevar a una fragmento —un grupo de personas— de la sociedad a conversar en una situación de laboratorio —espacio cerrado al exterior (cámara de Hessell, salón de un hotel o agencia de investigación).

EL NIVEL EPISTEMOLÓGICO

Este nivel se centra, básicamente, en la perspectiva dialéctica. Opera bajo esquemas que describen las relaciones entre las estructuras o cambios del sistema: cambios sociales. La técnica que mejor reproduce este esquema es la del socioanálisis, ya que contempla las diferencias que se producen desde el orden del decir al orden del hacer (una asamblea sindical).

LOS EJEMPLOS

Se proponen tres ejemplos para observar el isomorfismo que hay entre la articulación de lo social, con la posibilidad del diseño, y el uso de la técnica: Ibáñez (1992) propone el voto, la opinión pública y la asamblea para poder hablar de los modos de participación política donde se podrían dar las articulaciones señaladas anteriormente.

El voto

Si se quisiera observar la tendencia del voto en un periodo de campañas electorales, la entrevista es lo que mejor simula el juego de pregunta-respuesta, cuando se da a elegir una respuesta entre un conjunto (como cuando se encuentra una papeleta con los nueve partidos que contienden a una elección presidencial). Así, la encuesta estadística es el instrumento ideal para saber el comportamiento electoral, ya que el entrevistador y el entrevistado se oponen cuando uno pregunta y el

otro responde, sin alterar los papeles, el segundo solamente tendrá la posibilidad de contestar dentro del marco de opciones que le presentan (como en las votaciones).

La opinión pública

La opinión pública se construye a partir del diálogo. Los grupos de discusión son un dispositivo conversacional, es la herramienta ideal para construir discursos a través del diálogo entre un grupo de sujetos. De esta forma podemos saber sobre los procesos de construcción de la opinión pública, a partir de contestar los porqués de las diferentes posiciones que se generan. En este sentido, las sesiones de grupo se pueden utilizar también para diversos estudios de opinión, no sólo en lo referido a temas políticos o determinados productos comunicativos de los medios masivos, sino también, sobre aspectos que están relacionados con productos y servicios que usualmente exploran los mercadólogos.

La acción de masas

El socioanálisis es análisis institucional en situación o en vivo, una asamblea es el ejemplo. En este espacio operan dos órdenes, el del decir y el de la acción, el plano semántico y el plano pragmático: primero se trabaja con el orden del decir y luego se pasa al orden del hacer.

El socioanálisis trabaja también como un dispositivo conversacional, como los grupos de discusión, pero, a diferencia de éstos, el socioanálisis pasa también al plano pragmático, al orden del hacer. Desde una sesión de grupo no se puede esperar que se realice algún cambio más allá del propio grupo, en el socioanálisis sí.

LA TÉCNICA Y SU ALCANCE SOCIAL

Se parte de que la técnica alcance la acción social. Los ejemplos anteriores tratan de mostrar cómo los juegos de lenguaje, en la realidad, pueden ser reproducidos en diferentes situaciones con técnicas correspondientes y, sobre todo, en qué momentos se pasa del orden del decir al del hacer.

Juegos de lenguajes similares: el juego de pregunta-respuesta lo reproduce la encuesta y por ello es factible el sondeo electoral, su tendencia como un modo de participación política; la conversación es puesta en el espacio de los grupos de discusión, para poder trabajar en la forma en cómo se construye la opinión pública o formación de expresión pública; el socioanálisis alcanza a estudiar la acción de

masas al poder observar el cambio del plano semántico al pragmático, a través de una asamblea y las acciones que de ahí se derivan.

Los tres niveles

Tecnológico	¿Cómo se hace? Perspectiva distributiva... Técnica: Encuesta estadística	Trabaja con elementos.
Metodológico	¿Por qué se hace así? Perspectiva estructural... Técnica: Grupos de discusión	Trabaja con la relación entre los elementos.
Epistemológico	¿Para qué y quién se hace así? Perspectiva dialéctica... Técnica: Socioanálisis	Trabaja con las relaciones entre los elementos y el sistema.

Las sesiones de grupo en el proceso de investigación

ESTRUCTURA ABIERTA Y LA MIRADA DEL OBSERVADOR

La investigación mediante grupos de discusión está sostenida por la estrategia de un sujeto. Éste observa e investiga un objeto que a su vez puede constituir un sistema.¹⁸ Hay dos medios de informar (darle forma a –o introducir– una forma = informar¹⁹) un sistema: inyectándole información desde fuera (programándole), o construirle de modo que pueda producir información (producir orden en el sistema a partir del ruido –azar– del ecosistema).

Se parte de la idea de que los objetos (sociales) están organizados de alguna manera; tienen, por naturaleza, una forma de funcionar y comportarse; son en sí mis-

¹⁸ Manuel Martín Serrano, al explicar los antecedentes cognitivos de la cibernética, describe cómo desde el Renacimiento se configuraban los objetos de estudio de aquel entonces: "Se creía que, por muchas que fuesen las transmutaciones que se pudiesen llevar a cabo entre las formas que adopta la metría y las formas que adopta la energía, el orden de los objetos de la naturaleza se mantiene constante a lo largo del tiempo. En realidad, no habría creación ni destrucción, así pues, los renacentistas creían que la ciencia podía dar cuenta del funcionamiento de los objetos y de cómo los objetos eran. Lo que ni podía ni debía plantearse era la finalidad de los objetos, ni como se podían crear nuevos objetos. (Martín S., Apuntes de clases de doctorado, 1989.)"

¹⁹ "Conocer cómo están informadas las materias es conocer cómo son los organismos." (Martín S., 1989.)

mos entrópicos.²⁰ Para poder *manipularles*, es necesario controlarles a través de un mecanismo (tecnológico) que trabaje con información, una herramienta que la produzca y poder usarla (neguentropía).

El principio en el uso de la técnica para investigación social, es que el sujeto tiene enfrente objetos complejos que, a su vez, están en sistemas y ecosistemas más complejos. En la medida en que la sociedad avanza y crece, se le percibe, no sólo como más vieja, sino también más difícil de entender. Por lo tanto, el investigador que quiera hacer uso de esta técnica tiene que verse como un observador-sujeto en proceso, es decir, en cambio constante.

Es importante destacar lo anterior, ya que la naturaleza de la técnica que aquí nos ocupa implica un diseño que puede cambiar sobre la marcha. Nos podemos dar cuenta de que el diseño no es algo prefijado como en otras técnicas, ya que la manera en que se produce la información implica el conocimiento *a priori* de entropía en los grupos sociales con los que se trabaja. A continuación veremos cómo es que se plantea el diseño de los grupos de discusión.

EL DISEÑO

Distributivo

En el diseño que corresponde a la perspectiva distributiva, la encuesta, está separado de su ejecución. La parte fundamental de la ejecución se basa en una estructura de datos tridimensional: individuos, preguntas y respuestas.

Antes de empezar la ejecución, el diseño debe contemplar la tridimensionalidad anteriormente señalada: muestra, cuestionario y cuadro de resultados. Con la muestra selecciono al grupo de individuos que extraigo del universo que pretendo observar. Con el cuestionario planteo las preguntas y respuestas a elegir cuando los individuos tengan que dar la información. El cuadro de resultados me da la posibilidad de relacionar la posible respuesta del individuo con las hipótesis que de antemano pretendo afirmar (correlación de variables), así se tiene el programa diseñado.

El diseño distributivo comienza un proceso de búsqueda y debe limitar lo que se busca en las líneas de la tridimensionalidad de sus elementos. Nadie puede contestar algo que este más allá de las opciones que se dan para contestar y únicamente se contestará lo que se pregunte en el cuestionario.

²⁰ "Entropía es lo mismo que desorden, equiprobabilidad, aleatoriedad. Con el paso del tiempo, las cosas tienden a adquirir un estado aleatorio, a perder el orden interno. Neguentropía es la introducción de orden en una organización, si se quiere, información."

Actualmente la encuesta se encuentra muy desarrollada, y se ha trabajado (como podrá verse en el capítulo dedicado a esta técnica) de tal manera que pueda aprovechar las diversas opciones que presentan los objetos sociales en sí mismos. Para ello se utiliza, por ejemplo, la técnica de sesiones de grupo para comenzar a construir y diseñar el cuestionario que se aplicará.

Estructural

A diferencia del diseño distributivo, el estructural implica uno que atraviesa toda la investigación, por lo tanto, no está separado de la ejecución. Es decir, no es la parte fundamental de la investigación ya que irá cambiando dependiendo de las necesidades estratégicas del sujeto-investigador. Así, el proceso está regulado por la estrategia del sujeto.

Esto indica, a diferencia de la encuesta y la perspectiva que mejor representa que, en el caso del diseño estructural, el proceso de investigación implica un proceso de encuentro y hay que tener abierto el dispositivo; el diseño no se puede cerrar a elementos que de antemano controlo, al contrario, lo tengo que dejar abierto a lo que pueda ir sucediendo o encontrando para rediseñar la investigación, o por lo menos su parte inicial.

Ibáñez utiliza algunas imágenes para mostrar mejor las diferencias en cuanto al tipo, cantidad y calidad de información que se recoge en las dos técnicas mencionadas. La encuesta refleja una imagen fotogramática, luz incoherente donde los rayos no están en fase. En esta imagen, cada punto contiene información sobre un punto del objeto. Si se partiera dicho objeto-imagen por la mitad, se tendría *toda la información de la mitad del objeto*. Imaginar una fotografía como objeto de estudio; sabemos que la imagen está compuesta de muchos puntos pequeñísimos, de diferentes tonalidades, que logran configurar toda la imagen. Si observo una parte de la imagen, o mejor, si la separo o rompo, tendré toda la información de ese pedazo de imagen, pero pierdo el resto.

El grupo de discusión representa una imagen hologramática, en la cual la luz es coherente y los rayos están en fase. Cada punto de la imagen-objeto contiene información de todo el objeto. Si se partiera la imagen-objeto por la mitad, se tendría la mitad de la información de toda la imagen-objeto.

TAMAÑO DEL GRUPO

Usualmente, la sesión se empieza indicando que es importante escuchar lo que dicen los demás. Esto es muy importante ya que una persona no puede escuchar varios discursos a la vez.²¹

Hay que tener cuidado en no saturar los canales de comunicación que puede tener cada miembro del grupo, no sólo pidiendo que no hablen varios a la vez, sino también controlando el número de participantes. Asimismo, hay que cuidar el mínimo de canales de comunicación y, por lo tanto, de actuantes. Ibáñez menciona cuatro participantes como mínimo y diez como máximo. La lógica que utiliza es la siguiente: debe haber relación equilibrada entre el número de actuantes y de canales posibles de comunicación; a medida que el tamaño del grupo crece, también lo hace el número posible de canales de comunicación, pero en forma algebraica.²²

Cada participante articula el o los discursos que escucha con el de sí mismo. Así, se vuelve importante que no tenga ni muchos canales de comunicación posibles ni pocos. Cuando hay más de siete, debe haber una sección que no hable.²³

EL TIEMPO

Se recomienda no más de una hora y media. Es más difícil puntuar el comienzo que el final.

No se quiere decir que debe durar 90 minutos exactos, pero se ha percibido que después de este tiempo, el grupo sufre cierto desgaste energético y comienza a tener un rendimiento disparejo. Tómese en cuenta el desgaste que implica el desplazarse al lugar donde se convoca. Sin embargo, se dan grupos y diseños que requieren de más tiempo de trabajo y no presentan ningún problema.

El comienzo está puntuado por la convocatoria. Se solicita a un grupo de personas que asista a un lugar para charlar o dialogar, se introduce al grupo en un labo-

²¹ Bion dice que "sólo si los individuos se acercan suficientemente unos a otros es posible dar una interpretación sin necesidad de gritar: de la misma manera es necesario que todos los miembros del grupo puedan comprobar elementos en los que se fundamentan las interpretaciones. Por esta razón, el número y el grado de dispersión deben ser limitados." (1974, p. 136.)

²² "Entre dos (A, B) cabe una relación (AB). Entre tres (ABC) caben tres relaciones (AB, AC, BC). Entre cuatro (A, B, C, D) caben seis relaciones (AB, AC, AD, BC, BD, CD). En general, entre n caben $(n(n-1)/2)$ relaciones." (Ibáñez, 1992, p. 273.)

²³ "Un razonamiento convergente es proporcionado por Brondal... se refiere a la capacidad de sincronización, en el momento de la recepción de una cadena sintáctica; un receptor (o destinatario) sólo puede captar simultáneamente un máximo de seis términos; la discusión de un grupo exige para cada actuante la sincronización de las recepciones de las actuaciones de los otros, y es imposible atender a más de seis hablando. Un grupo de más de siete sólo funciona (sin esquizofrenizarse) si una parte se calla". (Ibáñez, 1992, p. 273.)

ratorio (cámara de Hessell) y quien convoca decide el aquí y el ahora. Usualmente se cuenta con personal especializado que *recluta* a las personas que conformarán el grupo de trabajo.

El final puede ser puntuado por el moderador/preceptor o por el grupo, ya que a éste último se le da la posibilidad de buscar el consenso. Si lo puntea el moderador/preceptor, el consenso parecerá decidido por el mismo. Si lo puntea el grupo, tardarán mucho en ponerse de acuerdo. Se vuelve necesario el equilibrio y, por ello, es más difícil que el inicio de la sesión. Parece ser que lo mejor es advertir el tiempo de duración: esto precipita al grupo a la urgencia de consenso cuando se está llegando a la hora mencionada.

Configuración del grupo de discusión

El grupo está configurado básicamente por dos partes: el moderador/preceptor y el grupo.

EL MODERADOR/PRECEPTOR

En México, se denomina moderador a la persona que va a dirigir la sesión. Usualmente es el encargado de participar en otras etapas de la investigación, sobre todo en la del diseño y en la del análisis del discurso producido por el grupo. Yo prefiero llamarle investigador u observador, ya que cumple esta doble función, no sólo la de dirigir la sesión (hay investigaciones, especialmente cuando son encargadas desde otro país, en las cuales se pide sólo la parte de moderación y el resto de las etapas de la investigación se hace en el país de origen).

Ibáñez utiliza el nombre de *preceptor*²⁴ dada la función que cumple en la sesión de grupo. Principalmente, el preceptor no controla la palabra (es lo ideal), sólo propone los temas que va a discutir el grupo. Tiene los hilos del títere, pero realmente devora al grupo a través del análisis posterior de su discurso.

El observador-investigador está en una situación de dominio, ya que convoca, provoca, el tema de discusión, asigna el espacio y limita el tiempo.

²⁴ "Por eso hemos sustituido la denominación usual de *monitor* (de *monere* = avisar) quien designa el papel manifiesto: *avisa* indica lo que hay que hacer, pero ¿por qué y para qué?, ¿en nombre de quién? *Preceptor* (de *pre* + *capere* = adelantarse a asir -a agarrar los hilos-) designa el sentido de ese papel: la relación asimétrica, pues representa el poder instituido, la jerarquía fundada en un adelantamiento (connotaciones presentes en *preceptor pedagógico* y en *precepto legal*); un *precepto* es una *prescripción* (= *pre* + *escritura*), el *preceptor* se adelanta a agarrar a los otros mediante la *escritura* (los tiene cogidos, cogiendo, adelantándose a coger, los hilos de la escritura)." (Ibáñez, 1992, p. 271.)

Convoca al decir quienes deben hablar del tema que el mismo investigador decide. Todos los individuos llegan a un mismo lugar, que por lo regular es una Cámara de Hessel en algún centro de investigación u otro lugar especialmente designado para estas actividades. Finalmente, el que convoca decide a qué hora empieza la sesión y cuánto debe durar.

SELECCIÓN DE PARTICIPANTES

En la selección se dan dos tipos de operaciones, una algebraica y otra topológica. Determinación algebraica (abstracta) de las clases de participantes. Contacto concreto, mediante redes topológicas, con los participantes, su origen y las relaciones de cada uno en ese origen, en ese contexto.

Debe recordarse que al trabajarse con grupos de discusión se esta operando con las relaciones entre los sujetos; sujetos en diálogo durante una sesión. La observación se centra en los huecos formados por los silencios o límites espaciales y por las fronteras o límites temporales. El espacio está demarcado por zonas polarizadas que separa los extremos y los frentes. Y el tiempo está demarcado por las fases transicionales o los momentos de cambio.

Supóngase que se quiere remodelar la imagen de marca de una cerveza que ha sido generada como una cerveza de imagen regional en un estado fronterizo del país. Las zonas polarizadas estarían representadas de un lado (polo regresivo) por señores de edad de clase media tradicional, en una ciudad del centro del estado, y del otro lado (polo progresivo), jóvenes *in*, de profesión emergente, en una ciudad del mismo estado pero que hace frontera con Estados Unidos. Los primeros pueden presentar una fase transicional, en cuanto que muchos de ellos pudieran ser campesinos emigrantes del campo a la ciudad, personas que han cambiado de contexto. Los segundos, los jóvenes, pudieran presentar un estilo de vida que se ha visto alterado por la excesiva abundancia de productos norteamericanos después de la firma del TLC.

Cada grupo también tiene fronteras interiores y exteriores. Ibáñez utiliza las imágenes del tonto y el loco para describir los dos límites de la comunicación: cuando se opera sobre relaciones de inclusión que tienen los participantes, todo es común y por tanto, comunicado, repetitivo (tonto); cuando se opera sobre relaciones de exclusión nada es comunicable, por tanto, no hay posibilidad de relación (loco). Sólo hay comunicación si las relaciones son de intersección: hay parte común, pero también hay parte no común.

Piénsese en estos tipos de relación en la sociedad de hoy en día. A nivel micro no se pueden mezclar personas de generaciones distintas, es muy difícil hacerlas platicar. A nivel macro, es muy duro poner en relación a un patrón con sus trabajadores.

Para que la comunicación sea posible, un grupo concreto de investigación debe estar incluido por fronteras excluyentes y debe incluir fronteras inclusivas (pero no fronteras excluyentes). (Ibáñez, 1992, p. 491.)

El contacto topológico para traer a los participantes al grupo implica la utilización de redes de relaciones, de comunicación. Hay dos tipos de redes, las naturales (primarias) y las artificiales (secundarias, que son creadas en el mismo proceso de investigación).

Las redes primarias son redes preexistentes, de parentesco, de vecindad, de trabajo, de amistad, etcétera. Es necesario hacer creer a los participantes que juegan su propio juego y la relación preexistente es muy importante. La red *polímera* encadena relaciones de distintos tipos y parece la más conveniente siempre y cuando cumpla con dos condiciones: que los participantes no se conozcan entre sí ni sepan nada del preceptor-observador; y que la red se rompa por la introducción de un profesional que no dejaría cualquier flujo de duda.

Si la red implica una relación asimétrica se impondrá orden al grupo (por ejemplo empleados por su jefe). Si la red tiene relaciones simétricas (coherentes) el grupo es neutral (por ejemplo, vecinos). Si la red tiene relaciones simétricas (incoherentes) el grupo entra en desorden ya que, a pesar de que cualquier contenido es válido, hay que tomar en cuenta los efectos de la red (por ejemplo lo que digan participantes que provienen de ONG's sobre un tema como el EZLN, sería significativo el que éstos hablaran mal del Ejército Zapatista).

TERRITORIO

El local donde se realiza la sesión del grupo es, por lo regular, un recinto cerrado hacia afuera y abierto hacia adentro. Es cerrado hacia afuera pues se trata, mientras dura la sesión, de que nadie moleste y de que los participantes se sientan lo más confortable posible para poder dialogar con los demás; es abierto hacia adentro pues la discusión irá poniendo nuevas rutas o temas a tratar por los miembros participantes y por el preceptor-observador. El grupo tiene la posibilidad de crear consensos y nueva información.

El grupo es un espacio de trabajo y su espacio está determinado: la mesa contiene contra la tendencia de fusión imaginaria en el grupo, no permite o no se permite que las alianzas subsistan mucho tiempo para buscar los consensos; las sillas contra la tendencia de permutabilidad. El grupo tiende a trabajar y sus miembros a asumir papeles diferentes.

Los locales naturales son mejores. Esta naturalidad radica principalmente en su neutralidad: abiertos para todos. Por ejemplo, locales ideales para representar

esta neutralidad son los hoteles (cualquiera puede entrar) y los centros de investigación y universidades (representan el saber o los que saben).

Funcionamiento del grupo de discusión

El grupo, a lo largo de su sesión de trabajo, vive un momento interior y otro exterior. El primero, es la historia que crea el propio grupo durante la sesión, es su historia (ficción); el segundo, es el que mete al grupo en la historia o la historia del grupo, es un proceso que nunca termina.

EL ESCENARIO

El grupo entra a un escenario que consta de personajes y relaciones. Los personajes son el moderador/preceptor, los miembros del grupo y el propio grupo: los dos primeros son reales, el último es imaginario en tanto que aparece y desaparece cuando la sesión de trabajo inicia y termina, respectivamente.

Las relaciones son asimétricas (transferencia y comunicación) y simétricas (de fusión y degradación del grupo): al comienzo, el silencio del conductor pesa, el grupo lo interpela o espera a que les diga qué es lo que debe suceder o decir; la transferencia vira del conductor al grupo y el primero, aunque no deja de perder su condición de sujeto-supuesto-saber en su totalidad, debe tratar de participar lo menos posible en la discusión; al transcurrir la discusión, la transferencia vira al grupo y tratan de ponerse de acuerdo entre ellos: –la discusión sale a la luz–.

Cada participante tratará que su posición se refleje en el grupo, al tratar de buscar el consenso en el grupo o en parte de él, fluirán mensajes a través de canales comunicativos bajo un esquema de juegos de dominación, una comunicación dual y circular (en momentos estarán hablando dos miembros del grupo, pero luego llegará un tercero al cual se le tendrá que contestar, así, cada miembro del grupo, estará dualizando y circulando los canales de comunicación a lo largo de la sesión).

Lo que más dinamiza los flujos de comunicación son los terceros que entran a romper los consensos logrados en distintos momentos de la sesión. La fusión nunca llegará a realizarse, el grupo siempre será un objeto.

LA ESCENA

Es la actuación de los personajes en la obra, es decir, actúan el conductor-observador y los miembros del grupo.

El moderador actúa a través de él mismo y sus dobles, es decir, a través de algunos de los miembros del grupo. Su actuación incluye una provocación inicial

y una continuada; se propone de que hablar en un principio y, después, se proponen o retoman temas e ideas que interesan explorarse. Se inicia con un componente racional, el tema a discutir, y sigue con uno emocional, el deseo de discutir.

La propuesta inicial puede ser denotada o connotada: en la primera aparece el nivel de tema (cerveza), nivel superior (bebidas), apareciendo el contexto; en la segunda se propone un tema que encadena a otro tema por condensación metafórica (los indígenas en México—el EZLN), o por desplazamiento metanímico (los errores de diciembre-devaluación-necesidad de conocer las expectativas sobre la devaluación del peso frente al dólar).

El deseo de los miembros del grupo es manejado por el moderador/preceptor, esto hace que se den complicidades y enfrentamientos. El moderador/preceptor no participa en la discusión, trabaja sobre ella. El conductor adopta diferentes posiciones ante el discurso: represiva, cómplice y de pantalla. La última es la más correcta, en la cual debe mostrar que ve y escucha todo.

Interpretación y análisis del discurso

Algunos alumnos de licenciatura y posgrado o investigadores ya con experiencia, se han llevado cierta sorpresa al no encontrar una especie de *guía* o algo parecido, para poder hacer el análisis del discurso producido en una sesión de grupo. Y es que al análisis del discurso, en este caso por lo menos, no se le puede hacer un solo tipo de análisis. Parece ser que la propuesta es que se abran las posibilidades de análisis, recuérdese que el observador/investigador es un *sujeto en estrategia* y cada discusión, en una sesión de grupo, refleja y refracta (a nivel micro) una sociedad y una historia.

Ibáñez²⁵ presenta tres niveles de análisis: el *synnomo*, autónomo y nuclear.

Empezando por el nuclear, se diría que es la *captación de los elementos de verosimilitud*, y ante esto, Ibáñez menciona que la verosimilitud es una simulación de la verdad, que la verdad definitiva no existe como tal. Clasifica cuatro tipos de verosimilitud: referencial, poética, lógica y tópica. Me centraré sólo en la referencial y la tópica.

La referencial está definida por un ejercicio que *produce el efecto de realidad* mediante un orden metafórico/semántico; es dar orden y una clasificación a un re-

²⁵ “El orden social es el orden del decir: está hecho de dictados e interdicciones. Hay sistemas dinámicos, en los que sólo hay intercambio de energía, y sistemas lingüísticos en los que hay también intercambio de información (éstos están regulados por códigos, por códigos genéticos los sistemas biológicos, por códigos lingüísticos los sistemas neológicos). Los cursos de nuestra existencia están regulados por discursos. Hay una doble red lingüística, de términos (descritos por sintagmas nominales) y de caminos (prescritos por sintagmas verbales).” (Ibáñez, 1992, p. 312.)

ferente que aparentemente está desordenado y es caótico: es una forma de ordenar la realidad mediante un discurso.

La verosimilitud tópica produce el efecto de sociedad. El sentido que provoca este efecto explica los valores que la sociedad acepta, es en lo que descansa la sociabilidad. Los sistemas sociales crean convenciones dado que son aceptadas por sus integrantes; así como van cambiando normas y costumbres, van alterándose los convencionalismos a través de distintos juegos de sociabilidad: qué se vale y qué no. Ahora, habría que imaginarse los ejemplos de esto a través de discursos que se refieren a la realidad política, religiosa o social.

Lo anterior, nos remite a plantear la faceta hermenéutica de técnica que tratamos.²⁶

Dilthey considera que el papel de las ciencias sociales es interpretar el contenido mental o las experiencias de vida a partir de sus expresiones. Existen dos tipos de expresiones de vida:

1. Los conceptos que son pensamientos estructurados, se expresan en forma de texto y tienen la intención explícita de comunicar.
2. Las acciones que no se proponen comunicar algo, aunque en su constitución existe un contenido espiritual individual y un contexto de ese contenido.

La comprensión de las acciones requiere tanto de reconstruir el contexto como de revivir ese contenido espiritual. Para Dilthey existen formas elementales y formas superiores de comprensión. Las formas elementales son las que se dan en la

²⁶ Von Wright, en el primer capítulo de su libro *Explicación y comprensión*, habla de dos tradiciones que en la historia de las ideas difieren en el planteamiento de las condiciones que debería satisfacer una *explicación científicamente respetable*. Las llama la *aristotélica* y la *galileana*. El contraste entre ambas tradiciones radica en el punto de vista de cada una de ellas sobre lo que es la explicación científica. La tradición aristotélica se caracteriza por las explicaciones teleológicas o finalistas mientras que la galileana, por la explicación causal o también llamada mecanicista, busca explicar y predecir fenómenos.

A partir del siglo XIX, cuando las ciencias naturales habían sentado ya sus bases intelectuales, el amplio desarrollo de los estudios sobre el hombre, su historia, su lenguaje y sus instituciones, encuentra correspondencias en ambas tradiciones.

Comte y Stuart Mill representarían la tradición galileana replantada como *positivismo*. Von Wright habla de tres principios que caracterizan al positivismo: 1. El monismo metodológico que sostiene la idea de la unidad del método científico a pesar de la diversidad de objetos temáticos de la investigación científica. 2. La consideración de que las ciencias naturales exactas (física, matemática) son el canon o ideal metodológico que mide el grado de desarrollo y perfección de todas las demás ciencias, incluidas las humanidades. 3. La idea de las explicaciones causales en un sentido amplio, es decir la subsunción de casos individuales bajo leyes generales hipotéticas de la naturaleza, incluida la naturaleza humana.

La otra posición surge como reacción contra el positivismo y alcanza un lugar importante a finales del siglo XIX, es más diversificada y heterogénea; podrían entrar en la clasificación diversos nombres pero Von Wright prefiere llamarla *hermenéutica*. Entre los autores que repre-

vida diaria, en las relaciones interpersonales en que los sujetos comparten un mismo mundo espiritual y mental; no hay necesidad de procesos de inferencia porque existe un mundo común (*objective mind*).

En las formas superiores de comprensión es donde surge la necesidad de la hermenéutica, ya que es necesario reconstruir contextos y revivir situaciones desconocidas para el intérprete. Esta labor de inferencia es la que constituye el método empático rescatado por Dilthey.

De esto se desprende que para Dilthey el significado no es solamente una cuestión de lenguaje, sino que se encuentra también en las acciones. El significado en la experiencia de vida es variable y contextual; y el método para conocerlo es empático.

Finalmente, no creo que estén agotados, en el marco aquí presentado, todos los elementos para seguir explicando conceptos de diferentes autores, para ver todas las posibilidades de análisis del discurso producido en una sesión de grupo. Sólo agregaría que hay que abrir mundos posibles, posibilidades que configuren nuevas formas, pero aceptar que éstas luego se deshacen y se crean otras nuevas. Esto es seguir en una situación estratégica, en una posición abierta de querer seguir observando y construyendo, hasta que en un momento dado, también yo, defino en dónde y cuándo paro.

sentan este pensamiento están Droysen, Dilthey, Simmel, Max Weber, Windelband, Rickert, Croce y Collingwood. Defienden una postura antipositivista en metodología, es decir que rechazan los tres puntos del párrafo anterior y acentúan las diferencias entre las ciencias que aspiran a generalizaciones sobre fenómenos reproducibles y predecibles –leyes–, (llamadas nomotéticas por Windelband), y las que buscan comprender las peculiaridades individuales y únicas de sus objetos –estudios descriptivos de lo individual–, (idiográficas).

Droysen, el filósofo e historiador alemán, establece una dicotomía metodológica que tuvo gran trascendencia en la diferenciación de las dos posturas y es la que existe entre *explicación* y *comprensión*. Las ciencias naturales tienen, según Droysen, como objetivo explicar; la historia tiene como objetivo comprender los fenómenos. Más adelante, Dilthey amplió y sistematizó estas ideas.

Aunque aparentemente no hay una gran diferencia entre ambas palabras, ya que cualquier explicación (causal o teleológica) llevaría en principio a la comprensión de las cosas, para estos autores la palabra comprensión incluye una dimensión *psicológica* que Simmel llama, a nivel de método, *empatía* y que describe como “recreación en la mente del estudioso de la atmósfera espiritual, pensamientos, sentimientos y motivos, de sus objetos de estudio.” El término comprensión está también vinculado con la intencionalidad, es decir que se comprenden “los objetivos y propósitos de un agente, el significado de un signo o de un símbolo, el sentido de una institución social o de un rito religioso.”

En el siglo xx hay una tercera reformatión de estas ideas, el positivismo surgió como neopositivismo o positivismo lógico y en la hermenéutica se desarrolló una vertiente de la filosofía analítica que, con Wittgenstein, influyó en la filosofía lingüística o del lenguaje ordinario.

El paquete técnico y sus posibilidades

Hay varios tipos de sesiones de grupo. Decíamos al principio que Ibáñez le llama a esta técnica *grupos de discusión*, que en Estados Unidos le dicen *focus group* o *grupos de enfoque*. Pero, en cada modalidad, se pueden abrir muchas formas más que ya no nos dejarían poner estos nombres. Desde luego que no se propone que se invente algún nuevo nombre si alguien utiliza la técnica de tal o cual manera. Simplemente, y respetando a todos aquellos que han fundamentado el nombre que han utilizado para llamar como le llaman a la técnica, tan sólo quiero referirme a las diversas formas en que se usa.

Los requerimientos nos marcan

Las estrategias de configuración del paquete técnico, ordinariamente, están supeditadas a las necesidades del o los que financian la investigación y, por lo regular, cuando piden que la técnica a usarse sea la de sesiones de grupo, ya saben de que manera y estilo se habrá de efectuar. No hay que olvidar que las técnicas son la *pinza*: con la que el cazador atrapará su presa, y quien paga esta cacería es quien decide muchas cosas.

No quiero decir o sugerir que quien paga decide todo, incluso sin importar los objetivos, sin embargo, algunas veces quien conoce y trabaja cotidianamente con la técnica se ve limitado por otro tipo de criterios que no necesariamente surgen de la experiencia de campo o de diseño, sino de parte de quien financia.

Desde luego, no hay un solo tipo de financiadores o clientes. Las instituciones que están detrás de quien contrata o pide la investigación también marcan el abanico de posibilidades que se van a tener o que se pueden presentar.

Por otro lado, pero muy de cerca, están los recursos financieros con los que se cuenta para la investigación. ¿Es una investigación para realizar una tesis de grado?, ¿el cliente es una transnacional, una ONG (organización no gubernamental) o una universidad pública? (en Chile, por ejemplo, a partir del cambio que se dio con el *no* a Pinochet, el gobierno civil empieza a contratar investigación para poder mejorar la gestión política).

En algunos casos, el moderador/preceptor también es seleccionado *a priori* cuando se elige la técnica. Hay gustos y formas de manejar las sesiones, hay quien hace mejor las cosas cuando se lo piden, o hay quien ya conoce el producto o el tema del cual se tratará en las sesiones, o incluso, hay quién ya lo ha trabajado tanto, que se encuentra *viciado* y no logrará lo que se quiere.

Posibilidades de combinación: en busca de metodologías plurales

Habría que aceptar la posición de sujeto/observador/reflexivo. Reconocer un distanciamiento entre sujeto y objeto no alcanza a explicarnos muchas cosas que pasan a lo largo de nuestras observaciones; una manera de poder explorar estas relaciones es a través de la reflexividad: preguntarse por qué y cómo es que observo, y de qué manera es que en este proceso el sujeto afecta y altera el objeto, y a su vez, el objeto altera al sujeto y, por lo tanto, altera su mirada, su manera de observar.

La creatividad y los juegos metodológicos no deberían marcar pautas inflexibles en la manera de abordar los objetos de estudio. Uno se debería de alejar de estas propuestas; lo construido para determinado tipo de estudio ha sido diseñado para determinado objeto, en situaciones determinadas: mundos posibles. La configuración de éstos cambia en tanto se decide la forma de percibir y observar. La cuestión sería cómo poder aprovechar lo ya hecho para abrir nuevos caminos, y aprovechar la experiencia para contraponer miradas y posibles interpretaciones.

La posibilidad de combinar técnicas para obtener información, no sólo responde al criterio de cantidad –tener más información sobre el objeto–. Se propone una cultura de investigación que trate de complejizar las metodologías: ¿si los objetos de estudio en ciencias sociales son complejos, por qué las metodologías no lo son también? La pedrada depende del tamaño del sapo.

El principal problema es el financiamiento, pero no es el único. La cultura de investigación en muchas instituciones de nuestro país no permite desarrollar nuevas propuestas. Es como picar piedra.

La idea sería lograr que se permitieran posibilidades para abrir mundos posibles: ¿qué pasa si quiero explorar nuevas posibilidades de lectura sobre un objeto con más de una técnica? Se pueden diseñar metodologías que nos dejen hacerlo.

TIPOS DE RELACIÓN TECNOLÓGICA

Pueden existir distintos tipos de relaciones entre las técnicas que se conozcan y se puedan utilizar. a) La de fase explorativa para el diseño, b) la de complementariedad y c) la de profundización.

El primer caso, se refiere a la utilización de la información que produce una técnica para el diseño de otra. Esto implica dos fases, en la primera se trabaja la técnica que produce información, para que luego se diseñe y aplique la otra.

En el caso *b*, la cosa es muy sencilla: observo con dos o más técnicas. No se contraponen con *a*. Se busca la construcción de hipótesis a través de distintas formas de mirar el objeto. Se rompe la idea de que la información que produjo una técnica es mejor que la(s) otra(s). Cada una implica distintas maneras de mirar, cada una tiene sus procedimientos, sus implicaciones y maneras de reconstruir la

realidad. Se requiere de imaginación metodológica para diseñar, estar abiertos y decidir cuándo cerrar. Y desde luego, después, contar lo que pasó, cómo se tomaron las decisiones y por qué se consideró esa trayectoria y no otra.

En el caso *c*, el de profundización, se pretende que a partir de lo que no puedo encontrar con una técnica sobre mi objeto de estudio, sí lo consiga con otra.

No se propone la promiscuidad tecnológica. Más bien, se parte de la idea de que un sujeto en estrategia, además de estar abierto a posibilidades, desarrolle con imaginación y creatividad diversas formas de entendimiento del objeto de estudio. Y, desde luego, que no se pierda el rigor metodológico.

También, habría que tener mucho cuidado en la relación que guardan las técnicas con el diseño metodológico y el marco teórico que pudiera imperar en el diseño de la investigación. Creo que lo mejor es que haya un marco teórico o metodológico que trate de contemplar los aspectos que se contemplan con cada técnica.

EJEMPLOS DE SITUACIONES DE COMBINACIÓN

Las sesiones de grupo y la encuesta

En el caso de sesiones de grupo ya se realizan este tipo de prácticas metodológicas. Usualmente, en el campo de la mercadotecnia y estudios de audiencia, se propone que para diseñar una encuesta se parta del discurso generado en las sesiones de grupo. Aprovechar el fragmento de realidad traído a este espacio para poder utilizar ese lenguaje, sus estructuras y, sobre todo, las representaciones exteriorizadas durante la discusión.

En este tipo de metodologías, no creo que la función de las sesiones de grupo signifique una fase exploratoria, para que luego se aplique la encuesta, aunque en ocasiones se considera así. Sin embargo, no se desperdician recursos para que la etapa de la investigación que se realiza con las sesiones de grupo, tenga sólo esta función. Se puede observar que cada una de las técnicas produce un tipo de información. Aquí, la propuesta es que el discurso generado en una sesión se aproveche para el diseño de la encuesta; esto ha provocado el mejoramiento de la construcción de reactivos en la encuesta, y también da la posibilidad de plantearse nuevos objetivos y preguntas sobre una realidad que, usualmente, el técnico de las agencias de investigación conoce parcialmente. Imaginar ¿en qué parte de la complejidad de los objetos de estudio en ciencias sociales es su constante cambio y movimiento?

Las sesiones de grupo y la entrevista a profundidad

¿Complementariedad o profundización? Todo depende. Hay ocasiones en que los actuantes de un grupo platican cosas que si lográramos contextualizarlas con ele-

mentos más personales, se podría entender por qué se habló lo que se habló. Sobre todo si esto es necesario para lograr un avance sustancial en la investigación y en lo que se está explorando.

Un caso típico es la investigación de audiencias. La complejidad de los públicos de los medios de comunicación está atravesada por una dinámica tal que se vuelve imposible describirla: cultura y vida cotidiana, niveles educativos, contextos familiares y sociales, nivel socioeconómico, etcétera. ¿Cómo configurar el perfil de una audiencia? Una posibilidad es combinar la manera de estudiarlas, que las técnicas produzcan información de varias de sus partes.

Recuérdese que la información que produce una sesión está dada en situación grupal, que cada actuante pone en discusión lo que cree, sabe y percibe de lo que se habla, y así, entre todos, construyen el discurso grupal.

Pero, ¿qué pasa si tengo la posibilidad de conocer más elementos de alguno o algunos de los actuantes, en cuanto a este contexto de vida que utilizan y los marca para decir lo que dicen en situación grupal? Indudablemente, tendré muchos elementos que entrarían en juego para comprender lo que los sujetos me dicen de manera individual y aislada, y lo que dicen al enfrentar sus representaciones con los otros sujetos en un grupo.

La interpretación de la información que produce cada técnica cruzará la mirada del observador, y le permitirá crear un mayor rango de relaciones que producen nuevas configuraciones para seguir explicando al objeto.

Estrategias de configuración de sesiones de grupo

LOS OBJETIVOS

Ordinariamente, los objetivos del estudio y el tema que trataran éstos, determinan, en gran medida, el tipo y procedimiento de configuración que tendrá el grupo o los grupos que se vayan a realizar.

Usualmente, se buscan las distintas formas de relación que hay entre el tema/objeto que se investiga y las partes o sectores de lo social que están más cerca en tiempo y espacio de éste.

Un ejemplo: si se quiere saber qué líneas de investigación, contenidos, duración, necesidades a satisfacer, etcétera, es decir, todo lo que pudiera tener que ver con el diagnóstico que se requiere para lanzar un posgrado de alguna escuela o facultad de una universidad, nos marca de entrada una limitante: ¿a quién le pregunto? Esto conlleva a pensar en la parte del diseño que tiene que ver con los elementos que se deben tener en cuenta para contemplar el campo de acción que tocaría un posgrado de determinada disciplina.

Esta es una primera fuente de información. Lo que se sepa de antemano, más lo que se genere para explicar el campo donde se trata de incidir, me va a dar las primeras pistas para la configuración.

Las fuentes

Además de la importancia que generan los objetivos para la configuración de los grupos, hay algunas fuentes básicas que también se toman en cuenta. La manera que se tenga para *clasificar* lo social con marcos teóricos previamente construidos, por lo regular, tiene mucho peso.

Por ejemplo, la Asociación Mexicana de Agencias de Investigación de Mercados y Opinión Pública (AMAI) desarrolló un estudio donde trató de obtener la clasificación de los niveles socioeconómicos (NSE), con el objeto de ubicar a qué nivel podían pertenecer los diferentes sectores de consumidores de los productos que ofertan las empresas que contratan investigación con las agencias que estudian el mercado mexicano.²⁷ Los NSE contruidos sirven para poder diseñar estudios con encuestas, sesiones de grupos, entrevistas a profundidad, etcétera, porque aportan información básica sobre la posibilidad de consumo de los sujetos. A partir de esta premisa, se agregan otro tipo de variables (ya no socioeconómicas) que terminan por definir el perfil que tendrán los grupos que se necesite formar.

²⁷ La AMAI se formó en 1992 y reúne a las principales agencias de investigación de México. En los antecedentes del estudio sobre NSE, realizado a lo largo de 1993 y 1994, menciona “que la distribución del ingreso, de los servicios educativos y culturales” son variables fundamentales en la investigación publicitaria y de mercadotecnia, “que era necesario unificar criterios de medición, formas de segmentación y tamaños de mercado” para poder mejorar el nivel profesional de la investigación en nuestro país. El proyecto se limitó a estudiar las zonas urbanas con poblaciones mayores a 50,000 habitantes, y mediante técnicas multivariadas se construyeron 13 categorías (por ejemplo: último año de estudios del jefe de familia, nivel de mando en el trabajo del jefe de familia, características de la vivienda, posesión de electrodomésticos, etcétera) que permitieran ubicar a la población en los distintos NSE. Los resultados arrojaron no sólo los distintos tipos de niveles, sino también, el porcentaje de población que incluía cada uno en aquellos años. Así, podemos ubicar que: el nivel A/B (nivel *alto*) contiene a 8 por ciento de la Ciudad de México y 5 por ciento que vive fuera de esta ciudad; los niveles C+, C y D+ (que constituyen el nivel *medio*) contienen 11, 21 y 26 por ciento, respectivamente, con un total de 32 por ciento de la población que vive en la Ciudad de México, y 36 por ciento de la población que vive fuera de la capital: 7, 26 y 25 por ciento, respectivamente; el nivel *bajo*, constituido por los niveles D y E, reúne a 60 por ciento de la población que vive en la gran ciudad y a 62 por ciento que radica fuera de ésta, quedando la distribución para cada uno de los dos niveles como a continuación se indica: 23 y 11 por ciento para la metrópoli y 25 y 12 por ciento para provincia.

EL TEMA/OBJETO DE ESTUDIO

El otro tipo de variables que pudieran incidir en la descripción del perfil de los grupos está marcado por el tipo de contexto en el que se inserta el tema/objeto de estudio que se investiga. Si bien podemos definir fácilmente el NSE, los objetivos marcan algunas limitaciones, el tema/objeto de estudio, por lo regular, marca el resto de lo que se va a tomar en cuenta.

Las variables que se conocen *a priori*, y que describen la relación ya existente (esto se conoce sobre todo en estudios previos o por experiencia de contacto con el tema/objeto de estudio) entre los distintos elementos, pueden incidir para la investigación del tema/objeto. Estas variables pueden describir comportamientos de grupos sociales o individuos en torno al objeto de estudio: no olvidar que el objeto de estudio en ciencias sociales es la misma sociedad.

Los contextos que encierran a los grupos a tomar en cuenta son diversos, pero lo típico sería lo siguiente: momentos de transición o tensión económica (piénsese en el TLC entre 1992 y 1993), políticas educativas *modernas* (un tema que ahora está de moda es cómo las universidades crecen en el área de posgrados a través de mecanismos rápidos y baratos); procesos de democratización, como el que trata de vivir México desde hace algunos años (imaginarse una competencia real y justa entre los partidos políticos para llegar a gobernar en congresos locales, municipios, estados, Congreso y Gobierno Federal), la crisis ambiental, que los ecologistas y diversas organizaciones quieren ayudar a resolver, y, por citar un último ejemplo, la monumental crisis económica que vivimos a partir de *los errores de diciembre*, a finales de 1994.

Un ejemplo sobre estudios de diagnóstico para universidades públicas

El caso que a continuación se explica consta de una metodología dividida en dos partes. La primera se desarrolló en la Escuela de Artes Visuales de la Universidad Autónoma de Nuevo León, y la segunda en la Escuela de Ciencias de la Comunicación de la Universidad de Sonora.

Este caso se desarrolla a lo largo del trabajo en dos universidades, porque ambas instituciones tuvieron una necesidad muy parecida y coincidió que me dieran la oportunidad de proponerles un manera de resolver algo que, de entrada, tenía muchas posibilidades de ser abordado.

Cabe mencionar que la idea de invitarme, estaba contaminada por el querer probar la técnica de sesiones de grupo para este tipo de estudios.

Tal vez, en este caso, no sea muy ilustrativo la faceta administrativa de la investigación con sesiones de grupo, dado que se trabajó con muy pocos recursos, económicos y humanos, pues no se contó previamente con equipos de trabajo

capacitados para cooperar en las distintas etapas de la investigación. Hubo un trabajo previo con la Escuela de Artes Visuales de la UANL, donde se pudo capacitar a un grupo de profesores en el manejo de la técnica, tratando de que la comprendieran desde sus dimensiones más amplias para poder llegar al día de la reunión nacional, con un equipo que participara en todas las etapas de la investigación. Esta experiencia fue muy gratificante, pues, aunque este equipo se formó para trabajar en la logística y coordinación de la reunión, se pudo compartir y reflexionar con ellos un par de días y, estoy seguro, pudieron ampliar sus conocimientos de metodología y aplicarlos, con alumnos de las diversas clases que se imparten en la Escuela de Artes Visuales.

Y, una advertencia más, quiero decir que se pretende presentar un resumen con los resultados más importantes de ambas etapas de trabajo. Este resumen, debido a falta de presupuesto, aún no está concluido pero que ya se tiene muy avanzados los resultados de una de las escuelas y, en todo caso, ninguna de las dos instituciones se opone a difundir comentarios y apreciaciones de las investigaciones que he realizado. Destacar esto es muy importante, ya que las investigaciones que se hacen para instituciones privadas son difíciles de mencionar, y más difícil aún publicar resultados. Así, a lo largo de la explicación, me referiré menos al trabajo hecho con la UNISON y las escuelas de Ciencias de la Comunicación, y más a la experiencia de investigación que se realizó con las escuelas de arte que se reunieron en la ciudad de Monterrey, para discutir sobre educación, cultura y creatividad.

Las necesidades

Actualmente, las instituciones mexicanas de educación superior han tratado de impulsar políticas de posgrado en diversas áreas. En este caso, hablamos de posgrados en el área de Artes Visuales y Comunicación.

En ambas escuelas se tenía la necesidad de crear una maestría que permitiera abrir una línea de trabajo innovadora para los cuerpos directivos y funcionarios de dichos centros escolares.

Dado que las instituciones tienen sus reglamentos internos para aprobar estas líneas de trabajo, la necesidad incluía el poder justificar una serie de requisitos que se piden para poner en marcha una maestría.

También había problemas de presupuesto, nada nuevo. Sin embargo, había que tomar en cuenta los pocos recursos con los que se me había dicho que se contaba para el trabajo. Esto es lo que marcó, fundamentalmente, el que se tuviera que trabajar en dos tiempos y en dos instituciones. Cada organización trabaja a su ritmo y con sus posibilidades, también hay que adaptarse a las circunstancias que este contexto genera.

Los propósitos

Mi propósito principal fue el comenzar a implementar una metodología que permitiera a instituciones de este tipo realizar diagnósticos que les aporten información para la toma de decisiones, de manera más eficiente, con más contacto con la realidad o, por lo menos, con aquella que se afecta por lo que hacen.

En las dos escuelas se realizó un encuentro nacional para los profesores y alumnos de arte, y uno regional (zona noroeste) para los educandos, profesores, investigadores y directivos de las escuelas de comunicación de la zona. Así, la idea era aprovechar a la gente que llegaría a estas reuniones para poder realizar mi trabajo.

Los objetivos

En el caso de la Escuela de Artes Visuales, se trataba de conocer las principales necesidades que hay en las escuelas de arte del país, en cuanto:

- a) Instalaciones, formación docente, planes de estudio y estructuras administrativas.
- b) Saber la percepción que se tenía sobre el campo laboral en cuanto a su situación y perspectivas a corto y a mediano plazo.
- c) Saber sobre el conocimiento de ofertas de posgrado en el país y su opinión sobre éstas.

Metodología para las escuelas de arte

- Primero se tuvo que construir una base de datos para poder tener el tipo de estudiantes, egresados, representantes de organismos públicos o privados, profesores con licenciatura o posgrado, directivos e investigadores que habían llegado al congreso.²⁸

²⁸ Ordinariamente, se tiene información previa que nos da indicadores en cuanto a esto, sin embargo, queríamos saber más precisamente quién había llegado el mismo día de inicio de la reunión; lo que se diseñó para construir este universo fue lo siguiente:

Nombre:

Edad:

Institución o nombre de la escuela donde trabaja o estudia:

Semestre que cursa (para estudiante):

Puesto que ocupa en la institución:

Nivel académico:

Institución privada o pública:

Ciudad:

- Segundo, para conocer cada segmento, se revisó la base de datos con toda la información que contenía, para poder saber realmente qué tipo de gente había llegado y poder pasar al diseño.
- Tercero, se pidió ayuda a un grupo de estudiantes de la escuela anfitriona para que la gente seleccionada de cada segmento llegara al punto de reunión donde se realizaría la sesión.
- Cuarto, se preparó un pequeño salón con una mesa, 11 sillas y una cámara de video que controlaba un camarógrafo para grabar la sesión.

DISEÑO

- Se detectó la llegada de estudiantes de más de 30 escuelas de diversas ciudades del país, directores de seis escuelas (dos públicas y cuatro privadas) y 15 profesores de escuelas invitadas (más los anfitriones, que no se registraron pues participaban en la organización del evento).
- Se decidió formar dos grupos: uno de estudiantes y otro de funcionarios de escuelas de arte, que en su mayoría eran directores. Debido a las actividades paralelas al congreso (en su mayoría talleres de trabajo para estudiantes y profesores) no se pudo concretar la reunión de profesores.
- Se diseñó un filtro/cuestionario para ubicar a cada participante que estuviera en las sesiones, la información se pedía antes de entrar para poder realizar una selección previa a la sesión.
- En este punto, se trataba de equilibrar los grupos, sobre todo el de estudiantes, pues no se sabía, por ejemplo, cuántos hombres o mujeres y de qué semestre serían localizados por el equipo de reclutamiento que cooperó en esta actividad.

REALIZACIÓN DE SESIONES

Se realizaron tres sesiones: una con directivos de las diversas escuelas (U de M, CEDIM, Arte A.C., UANL, CEDART y ENAP/UNAM), otra con estudiantes de la UANL y otra con estudiantes invitados de las otras universidades.

FASE DE ANÁLISIS

Para realizar el análisis de las sesiones, se diseñó una matriz²⁹ que permitió clasificar los temas que se tocaron, relacionar las verbalizaciones, tal cual fueron mencionadas por los actuantes, con las oraciones que interpretan dichas verbalizaciones y, finalmente, ésta también permitió relacionar oraciones conclusivas para cada tema.

A continuación, se muestra la guía de análisis que se construyó para la sesión de directivos de escuelas de arte:

- a) Percepción de las necesidades de educación superior.
- b) Necesidades básicas de la educación artística.
- c) Opinión y evaluación de la planta docente.
- d) Planes de acción.
- e) Ofertas actuales de posgrado.
- f) Factibilidad de un posgrado en artes.
- g) Percepción del campo profesional.
- h) Nuevas propuestas de posgrado.

Ahora, mostraré ejemplos de oraciones que interpretan lo discutido en la sesión de directivos, después de la oración se verán ejemplos de verbalizaciones para observar cómo se puede reforzar lo interpretado.

Ejemplo sobre el tema b

- Debe existir una constante revisión del plan de estudios y ésta debe efectuarse por personal capacitado.
 - Estamos viviendo un momento de evolución y la implementación de un nuevo plan de estudios que abre una nueva licenciatura en artes plásticas.
 - En relación a los planes de estudio y programas de revisión, debe haber personal capacitado para esto.
 - En el caso del plan de estudios, no basta con la sola renovación, sino además necesitan saber dónde adaptarlos.
 - Es necesario la revisión de planes curriculares para adaptarlos, primero, a las necesidades propias de cada ciudad, en este caso Guanajuato, y después contactar con otras universidades.
 - Hay cosas previas a la revisión del plan de estudios. Yo no estoy seguro de que la sola reformulación de éste resuelva problemas que son serios.

²⁹ Esta matriz fue elaborada con la ayuda de los investigadores: Laura González, Rodrigo Pérez, Marisol García, Eduardo Escamilla y Enrique Pimentel.

- El no contar con niveles de bachillerato propios de cada institución, disminuye el nivel de preparación al momento de ingresar a los estudios superiores.
 - El problema de la Universidad Veracruzana es que como no se contempla el nivel bachillerato, los alumnos que quieren cursar carreras artísticas llegan con una formación nula.
- La difusión de lo que se produce en las instituciones es una necesidad urgente, debido a que es importante que las escuelas difundan lo que ahí se hace y que se ve disminuido por no encontrarse en el centro de la ciudad.
 - Nuestro problema en la Universidad de Montemorelos es la difusión del arte, la difusión para nosotros es muy importante, estamos un poco alejados del centro.
 - No se da la difusión apropiada de lo que se está haciendo, nosotros pensamos que los estudiantes necesitan que, como escuela, los promovamos.

Ejemplos del tema c

- La especialización y actualización son importantes dentro de las carreras en artes: se necesitan especialistas en historia del arte y grabado, ya que la experiencia demuestra la insuficiencia en estas áreas.
 - Por parte de la facultad de Artes Plásticas de la Universidad Veracruzana, los problemas son la actualización y especialización del sector académico.
 - En Guanajuato, nos enfrentamos al problema de, a pesar de contar con profesores, no tener especialistas en historia del arte, historia del grabado, ese tipo de maestros hacen falta.
 - Mis maestros todos tienen título, pero algunos tienen problemas con materias de historia, por ejemplo, les cuesta un poco de trabajo entender la lógica del posgrado.
- Aunque en las materias teóricas existe una carencia de profesores, se percibe que debe haber en capacitación una mejor preparación pedagógica y didáctica, en actualización: estar cotidianamente enterado de lo que se produce en el arte y una producción propositiva y constante.
 - En materias teóricas tenemos una carencia muy grande de maestros.
 - En la preparación de los profesores se requiere de una capacitación pedagógica más amplia, una actualización permanente con relación a la produc-

ción plástica, una información al día, una producción constante, que se mantengan siempre propositivos y una formación didáctica que les permita transmitir esas experiencias a los alumnos.

- Los profesores con algo de *currículum*, o sin él, pero con mucha experiencia, tienen muchos problemas al intentar colocarse en otras universidades que no sean las locales. Se percibe esto como un problema nacional.
 - A nivel interno o local se resuelven los problemas de algunos docentes con cierto *currículum*, pero si este maestro quiere ir a otras universidades está desarmado.
 - No tenemos profesores con grado de licenciatura, pero sí preparados: creo que esto es un problema a nivel nacional.
- El contar con maestros extranjeros en la enseñanza de las artes, aunque aporta rendimientos, involucra un costo mayor, que tiene como consecuencia que se trabaje apenas con lo disponible.
 - En nuestra institución, la Montemorelos, tenemos maestros que cuentan con maestría e incluso con doctorado, tenemos de Francia, tres de Estados Unidos, que ayudan a renovar el *currículum*.
 - Lo que pasa es que ella los importa.

Ejemplos del tema d

El enriquecimiento, desarrollo de los alumnos y su integración mutua, crecerá si se invita a artistas de buen nivel a dar cursos, cátedras o seminarios.

- Que ciertos artistas colaboren con la institución y que de ahí se enriquezcan los alumnos, que se desarrollen para no dejarlos ir a otras partes.
- Hay posibilidad de invitar a algunos artistas extranjeros de gran nivel que puedan impartir cursos, cátedras o seminarios.

Los estudios dentro y fuera del país son algunos de los planes que se contemplan para el avance en el conocimiento de las artes.

- Una de las soluciones fue presentar proyectos en el Fondes, Fondo Nacional para la Difusión de la Educación Superior, con becas dentro o fuera del país.

- En Guanajuato ya han habido algunos avances para que los profesores vayan a otros lugares.
- El intercambio de profesores a nivel escuelas es importante y significa un avance, sólo que se exige que el intercambio sea mutuo.
 - Un avance sería fortalecer el intercambio a nivel escuelas en el país.
 - Estamos en posibilidad de aceptar maestros de otras universidades, siempre y cuando allá ellos también nos acepten a nosotros.
 - Dentro de la escuela se ha dado la posibilidad de salir.
- El tiempo completo se está ampliando para profesores con doctorado, maestría o materias de talleres, además de que se les estimula con cursos.
 - Los maestros que realizan sus maestrías o doctorados se les manda a cursos y la mayoría tiene su tiempo completo.
 - Se busca que se tengan maestros de tiempo completo en talleres y de medio tiempo en materias teóricas.
- Otro avance en la solución del problema académico de las artes son ciertas publicaciones que sacan las escuelas, tales como las de Guadalajara y La Esmeralda.
 - En Guadalajara se acaba de publicar el Plan Institucional de Desarrollo, el cual contempla las áreas de investigación, extensión y docencia.

Ejemplos del tema g

- En México, el aprendizaje de las artes debe ser por medio de la práctica, pues ésta y la teoría están desligadas, lo que provoca ciertos problemas en el campo laboral.
 - En México sólo hay tres licenciados en museografía y muy pocos en musicología, en el Claustro de Sor Juana tenemos que hacer trueques para que éstos se relacionen con el medio y se preparen mejor.
 - A los alumnos les enseñamos en clase lo que se debe hacer y cuando vamos a los talleres les enseñamos lo que no se debe hacer, pero así es como se hace. Mi idea es que la gente se vaya formando en el campo real.
- El posgrado en el extranjero no importa tanto, no es tan solicitado el título.

- En las instituciones de Estados Unidos no es tan necesario el título.
- En Europa, se permiten ciertas situaciones donde no intervienen los grados.
- En el campo laboral, sigue siendo un problema la situación de los maestros al no haber una preparación docente, o por lo menos con nivel de licenciatura.
 - Los maestros pueden tener nivel académico muy alto, pero no estar capacitados en el área docente, esto se está resolviendo con diplomados.
 - Es difícil que se tengan licenciados en arte.
- Se percibe que es necesario agrupar las actividades de práctica y de enseñanza de las artes en un centro particular, que permita mejorar estos aspectos de optimación de tiempos, recursos y crecimiento como institución dedicada al arte.
 - En Guadalajara se está estructurando todo, englobando todo lo que es arte, todo estaba disperso y se conjuntó en un centro universitario. La mayoría de los que están proyectando tienen posgrado y son arquitectos.
 - Se optimaron las horas y los recursos. Hay una nueva nivelación académica y se aumentan las horas teóricas.

Ejemplos del tema h

- Los requisitos para un posgrado, ya sea en producción o en enseñanza, deben tenerse muy en cuenta al momento de pretender crear uno, y saber qué orientación va a tener.
 - Para una maestría en artes se necesita, en primer lugar, que exista una biblioteca, profesores con maestrías y un doctor en historia y filosofía. Si no tienen eso mejor que ni se metan.
 - Hay un problema de definición muy grave de hacia a dónde van a orientar su posgrado.
- Todavía se está proyectando la creación de una maestría de enseñanza en artes, pero se hace notar que es urgente la licenciatura en esta cuestión, determinando sus alcances y límites.

- Que yo sepa la Universidad Veracruzana está estudiando un proyecto de docencia en arte para maestría y diplomados, por ejemplo en el diseño de la fotografía.
- Aquí se piensa que las escuelas de enseñanza del arte no tienen una determinación, no hay un límite de enseñanza de licenciatura; una de las preocupaciones que se manifiesta en la facultad, es el pensar en una licenciatura en docencia.
- La maestría en docencia se percibe como viable porque sustenta lo que después se va a crear, es la que determinará si el arte sobresale o no, y es la que mayores requisitos tiene.
 - El posgrado en educación es el que tiene más requisitos de todos, porque abarca mucho, y más si es en artes visuales.
 - Necesitamos artistas que sepan enseñar. Ahora tenemos una tarea más difícil, pues tenemos que sacar adelante el arte.
- Para la implementación de alguna maestría de docencia en artes, el principal problema sería la demanda.
 - Yo creo que el posgrado orientado a la educación sería el más importante, pero qué demanda van a tener, ése sería el cuestionamiento.
 - En la maestría teórica no habría alumnos.
- El posgrado en enseñanza involucraría aspectos teóricos, los cuales se perciben como problemas reales en la enseñanza de las cuestiones artísticas.
 - El gran problema que tenemos es que los alumnos se pongan a leer o a estudiar, ellos quieren producir.
 - Un artista no puede platicar, por eso un posgrado teórico no tendría respuesta.
- Se percibe que el posgrado en producción no sería viable, debido a que esto se aprende por otros medios: experiencia y práctica.
 - No veo tan necesario un posgrado en producción, creo que esto es más asimilado a través de diplomados y talleres.
 - Los alumnos están buscando un posgrado como productores, difícilmente se van a meter a un posgrado para maestros, ellos no están pensando en ser maestros.

- Interesa a los alumnos la tendencia a las nuevas tecnologías, que se pueden utilizar para realizar con ellas creaciones artísticas.
 - Nos interesarían los medios electrónicos como medios de producción. Llama mucho la atención una maestría en medios electrónicos contemporáneos.
 - Están interesados por la multimedia, pero la desconocen.
 - Sería el manejo de la fotografía y su relación con la computación.

- El panorama de posibilidades respecto a los posgrados que podrían implementarse es extenso y abarca distintas disciplinas artísticas pasando por materiales como la madera, la piedra y la pintura, pero siempre conjuntadas con las nuevas tecnologías.
 - Se podría hacer con talleres montados con supertecnologías para trabajar la madera, la escultura.

- Existen otras variantes que se perciben como opciones dentro de la preferencia de maestrías, cuyo perfil es la combinación de lo tradicional con lo nuevo.
 - La maestría en comunicación y arte, donde se utilizan la retórica y la hermenéutica.
 - Para saber qué está pasando con el pensamiento posmoderno, de toda la problemática tecnológica y, por otro lado, el pensamiento del arte en el quehacer contemporáneo, combinar lo tradicional con la tecnología, recuperar conceptos y técnicas.

- La creación de una nueva maestría tendría que igualar la práctica y la teoría, rescatando lo conceptual y haciendo de ésta especialización un verdadero compromiso de conocimiento.
 - Habría que buscar el desarrollo intelectual y conceptual, que debería de ser lo más importante.
 - Una maestría tiene que aparejar el estudio y el análisis de los conceptos consecuentes con el nivel de tecnología. Ahí es donde debe irse aparejando el posgrado, tanto en el aspecto teórico como en el práctico de talleres, también del uso de la tecnología mucho más complicada pero, sobre todo, con un compromiso.

CONCLUSIONES

A continuación se presentan las conclusiones de la sesión de directivos de escuelas de arte:

- a) La infraestructura, la revisión constante del plan de estudios por parte de personal capacitado, la difusión de lo que se produce al interior de las universidades y la cuestión económica, son los aspectos fundamentales que sobresalen de las necesidades básicas existentes en la educación artística.
- b) Las artes en nuestro país se ven desplazadas tanto por las autoridades gubernamentales como por las universitarias, las cuales siguen el proyecto neoliberal que hace que la cultura se concentre en grandes ciudades; además se percibe que es urgente una planeación general.
- c) La situación de los cuadros docentes en la enseñanza artística deben adecuarse y mejorarse cuanto antes: mediante la especialización, la actualización y la capacitación para contar con suficiente grado académico y poder enfrentar las responsabilidades que tienen frente a la educación artística.
- d) Las principales acciones para solucionar las necesidades de las cuestiones del arte, deben regirse bajo la organización y la motivación hacia los maestros, creación de espacios y optimización de recursos, intercambio tanto nacional como extranjero de profesores, así como publicaciones que los motiven.
- e) Un posgrado en las condiciones actuales tendría que aportar una verdadera especialización a quienes lo tomen, no debe quedarse en conocimientos de bajo nivel pues un estudio como éste implica poder contribuir a la misma sociedad, en busca de mejorar día con día.
- f) Los conocimientos que se adquieren en San Carlos no contribuyen a una especialización más avanzada, pues su *currículum* no la puede aportar, pero sí funciona como un puente para obtener ya sea el grado de maestría o experiencia en el medio.
- g) La maestría en educación artística es bien recibida, aunque hay dudas respecto a cuál será su matrícula, adelantando la poca afluencia; la maestría en producción es más oportuna y se centra principalmente sobre cuestiones tecnológicas, destacando la computación y la multimedia.
- h) Se percibe el campo profesional como un espacio donde se aprende haciéndolo; donde lo importante es la organización y la superación de los profesores, así como la integración de un centro donde se agrupan las actividades y enseñanzas artísticas.

PROPUESTA

A manera de propuesta, se sugirió avanzar en una propuesta de posgrado, tomando en cuenta lo encontrado en las sesiones realizadas. Sin embargo, también se sugirió realizar más sesiones con los siguientes segmentos: egresados, profesores, estudiantes que cursen o hayan cursado un posgrado en esta área y representantes del campo artístico, tanto del sector privado como del público.

Si se pudieran realizar las sesiones mencionadas, se podría tener información de más sectores que también están muy cerca a la toma de decisiones de las escuelas y facultades.

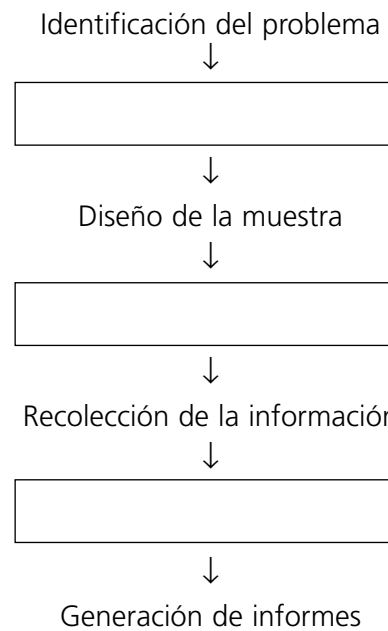
Desde luego que habría que tomar en cuenta que el estudio se tiene que comprender en su dimensión local (de ciudad) o regional (estatal), sin embargo, lo que pudiera encontrarse en otros lugares, con otras escuelas, sería bueno, sobre todo con aquellas que ya cuentan con un posgrado.

BIBLIOGRAFÍA

- DELGADO, Juan Manuel y Juan Gutiérrez, (eds.), *Métodos y técnicas cualitativas de investigación en ciencias sociales*, Editorial Síntesis, Madrid, 1994.
- GALINDO Cáceres, Jesús, *Política, cultura y comunicación Para una percepción de mundos posibles en el espacio social mexicano*, Universidad Iberoamericana-León, Cuadernos de Mass Culturas 1, México, 1995.
- GARCÍA Ferrando, Manuel; Jesús Ibáñez y Francisco Alvira, *El análisis de la realidad social. Métodos y técnicas de investigación*, Alianza Editorial, 3a. reimpresión, España, 1992.
- GONZÁLEZ, Jorge A., "Metodología y sociología reflexivas: navegar porcelosos mares del placer", en *Estudios sobre las culturas contemporáneas*, Universidad de Colima, volumen v, núm. 15, México, 1993.
- IBÁÑEZ, Jesús, *Del algoritmo al sujeto. Perspectivas de la investigación social*, Siglo XXI, España, 1985.
- _____, *Más allá de la sociología. El grupo de discusión: técnica y crítica*, Siglo XXI, 3a. edición, España, 1992.
- _____, *Por una sociología de la vida cotidiana*, Siglo XXI, España, 1994.
- ROGERS, Carl, *Grupos de encuentro*, Amorrortu Editores, 8a. reimpresión, Buenos Aires, 1993.
- SERRANO, Manuel Martín, *Métodos actuales de investigación social*, Editorial Akal, España, 1978.
- _____, *La producción social de comunicación*, Alianza Editorial, España, 1986.
- VERJAT, Alain (ed.), *El retorno de Hermes*, Editorial Anthropos, España, 1989.

C. EJERCICIOS

1. Complete el siguiente diagrama de las etapas de la encuesta



2. ¿Cuáles son las características metodológicas de la encuesta?

3. ¿Cuál es el tamaño de muestra requerido para una población de 8 000 individuos? Considere un nivel de confianza de 95% y un margen de error +/-5% y desarrolle la fórmula.

4. Con base en las lecturas, complete el siguiente cuadro:

	Enfoque cuantitativo	Enfoque cualitativo
Estrategia de investigación		
Fundamentos		
Técnicas de investigación		
Tipo de datos		
Papel de la teoría		

5. Relacione las siguientes columnas:

- () Nominal 1. En este nivel de medición la cuantificación del concepto parte de un cero absoluto y guarda una relación numérica.
- () Intervalar 2. En este tipo de medición los números representan nombre de categorías, lo valores no indican magnitud y no es posible realizar operaciones aritméticas.
- () De razón 3. En esta medición el sistema de cuantificación responde a una secuencia u orden.
- () Ordinal 4. En este tipo de medición los valores representan orden y magnitud, el intervalo entre valores es igual y los valores representan relaciones de orden y magnitud.

6. ¿En qué consiste el proceso de selección de la muestra llamado aleatorio sistemático?

7. De acuerdo con Roberto Hernández, Carlos Fernández y Pilar Baptista, los enfoques cuantitativo y cualitativo utilizan cinco fases similares, ¿cuáles son?

8. Operacionalice los siguientes conceptos: (Puede tomar como ejemplo el cuadro de la página 63 de la lectura de López Romo).

<i>Derechos económicos</i>	
<i>Derechos políticos</i>	
<i>Derechos civiles</i>	

9. López Romo menciona varios criterios para la redacción de preguntas de una encuesta, ¿cuáles son y en qué consisten?

10. ¿Cuál es la utilidad de los grupos de discusión en la investigación sobre derechos humanos?

E. AUTOEVALUACIÓN

1. Según Roberto Hernández, Carlos Fernández y Pilar Baptista, los enfoques cualitativos y cuantitativos comparten algunas fases del proceso de investigación, ¿cuáles son?

- a) Observar y evaluar fenómenos, así como comprender el sentido que los actores le atribuyen a sus acciones.
- b) Demostrar el grado en que las suposiciones tienen fundamento y utilizar instrumentos estandarizados.
- c) Observar y evaluar fenómenos con fundamentos positivistas.
- d) Observar y evaluar fenómenos, así como establecer suposiciones como consecuencia de la observación y evaluación realizadas.

2. Con base en el texto de Roberto Hernández, Carlos Fernández y Pilar Baptista, complete el siguiente enunciado: Un _____ elige una idea que transforma en preguntas, de ahí deriva hipótesis y variables, desarrolla un plan para probarlas, mide las variables, analiza las mediciones obtenidas y establece conclusiones.

- a) estudio cualitativo
- b) estudio cuantitativo
- c) modelo mixto
- d) modelo de dos etapas

3. De acuerdo con el texto de Heriberto López Romo, la precisión o margen de error es:

- a) El límite de un valor muestra o estimador, dentro del cual se encuentra el verdadero valor en el universo (parámetro).
- b) El proceso mediante el cual las ideas o conceptos son dimensionados de acuerdo a reglas.
- c) Una escala de medición de una variable aleatoria respecto a su esperanza.
- d) Una medida de dispersión para variables de medición y de intervalo.

4. Con base en la lectura de Heriberto López Romo, señale dos cualidades de la escala de medición ordinal.

- a) Los números representan nombres de categorías y los valores no indican magnitud.
- b) El cero no representa un cero absoluto; los valores representan relaciones de orden y magnitud.
- c) Los intervalos entre valores no indican distancia; las relaciones entre los números son de orden y secuencia.
- d) La cuantificación del concepto guarda una relación numérica; esta escala de medición parte de un cero absoluto.

5. Según el texto de Heriberto López Romo, la exactitud o nivel de confianza es:

- a) La probabilidad de que el valor del universo se encuentre dentro del margen de error.
- b) Una característica que puede ser medida adoptando diferentes valores en cada uno de los casos de un estudio.
- c) Un indicador de lo que estamos tratando de probar y se define como una explicación tentativa del fenómeno investigado.
- d) Una propiedad que puede variar y cuya variación puede medirse u observarse.

6. De acuerdo con la lectura de Roberto Hernández, Carlos Fernández y Pilar Baptista, el modelo multimodal es:

- a) El procedimiento de recolección de datos para probar hipótesis y establecer patrones de comportamiento.
- b) El procedimiento de recolección de datos sin medición numérica para descubrir y afinar preguntas de medición.
- c) La convergencia o fusión de los enfoques de investigación cuantitativo y cualitativo.
- d) El fenómeno que ocurre cuando la metodología se lleva a cabo bajo la perspectiva de alguno de los enfoques.

7. Según Roberto Hernández, Carlos Fernández y Pilar Baptista, ¿cuáles de las siguientes características son propias de los métodos cualitativos?

- a) Ofrecen la posibilidad de generalizar los resultados y otorgan control sobre los fenómenos.
- b) Dan profundidad a los datos, la dispersión, la contextualización del ambiente, los detalles y las experiencias únicas.
- c) Pretenden medir, confirmar datos, explicar relaciones causales.
- d) Permiten conocer con precisión los resultados y el riesgo implícito de las decisiones que se toman con base en esos resultados.

8. Con base en la lectura de Russi Alzaga, señale ¿cuáles son los tres niveles que diferencian los tipos de investigación en cuanto a su diseño?

- a) Teórico, metodológico, tecnológico.
- b) Epistemológico, empírico, metodológico.
- c) Teórico, metodológico, epistemológico.
- d) Epistemológico, metodológico, tecnológico.

9. De acuerdo con Russi Alzaga, ¿en qué perspectiva se ubica el nivel metodológico?

- a) En la perspectiva estructural.
- b) En la perspectiva dialéctica.

- c) En la perspectiva distributiva.
- d) En la perspectiva técnica.

10. Según el texto de Luis Jesús Galindo, ¿cuál es la relación entre el tamaño de la muestra y el tamaño del universo?

- a) El tamaño de la muestra es un porcentaje del universo.
- b) El tamaño de la muestra depende de manera proporcional al tamaño del universo.
- c) Existe una dependencia lineal entre el tamaño de la muestra y el tamaño del universo.
- d) Existe una dependencia no lineal entre el tamaño de la muestra y el tamaño del universo.

Módulo III.
**Metodología de la investigación
en derechos humanos**

OBJETIVOS ESPECÍFICOS

- Conocer las diversas técnicas de investigación social.
- Ubicar el trabajo que realiza el área especializada de estudio de la CDHDF en el contexto de las investigaciones en derechos humanos.
- Identificar y aprender de las fortalezas y debilidades de las investigaciones en derechos humanos.

A. SÍNTESIS

I. El estado de las investigaciones en derechos humanos

Los estudios sobre los derechos humanos no han sido numerosos ni sistemáticos, pues si bien se han desarrollado trabajos sociológicos sobre procesos de democratización, sobre la percepción de la igualdad de derechos entre la población o sobre violencia contra minorías (por citar algunos ejemplos) sus autores no suelen tener conciencia de estar trabajando en un ámbito que pueda denominarse “sociología de los derechos humanos”, sino que se interesan exclusivamente por el tema particular que investigan.¹

Tal dispersión se deja notar especialmente en la investigación empírica, donde se echa de menos un esfuerzo de conceptualización sociológica de los procesos típicos que conducen a sociedades donde el respeto a los derechos fundamentales es habitual.²

Ahora bien, hay un reconocimiento de que se ha realizado un esfuerzo en la investigación empírica. Específicamente, en los últimos tiempos se han desarrollado investigaciones cuantitativas de gran alcance.

II. Antecedentes de las investigaciones cuantitativas sobre derechos humanos

El origen de estas investigaciones se remonta al *Social Indicators Movement* de los años setenta, a lo que se añadió el interés del gobierno del presidente Carter en Estados Unidos sobre esta materia. La Foreign Assistance Act, en su artículo 502b, prohíbe la ayuda externa estadounidense a países con un patrón continuado de violaciones de derechos humanos. Pero para poder informar a la Cámara de Representantes acerca de si se cumple o no tal condición, el Departamento de Estado tenía que contar con datos de la situación en los diversos países del mundo.³

De este modo, se originó un informe anual que, al menos nominalmente, tenía como meta reunir información objetiva sobre el nivel de respeto internacional de los derechos humanos. A este informe se han ido uniendo otros, algunos de ellos conocidos por la di-

¹ Aymerich, Ignacio, “Introducción”, en *Sociología de los derechos humanos. Un modelo weberiano contrastado con investigaciones empíricas*, Valencia, Tirant lo Blanch/Universitat de Valencia, 2002, p. 14.

² *Ibid.*, p. 15.

³ *Ibid.*, p. 16.

fusión periodística que reciben, como el de Amnistía Internacional; o el de Charles Humana, que fue asumido como propio por Naciones Unidas.⁴

A estos informes generales se añaden estudios centrados en aspectos específicos de los derechos humanos, como el elaborado por el equipo de Ted Gurr y Barbaba Harff sobre genocidios y politicidios o el de Zehra Arat sobre la correlación entre desarrollo económico y derechos humanos.⁵

III. Utilidad de las investigaciones cuantitativas sobre derechos humanos

Hay dos razones específicas por las cuales serán útiles las investigaciones cuantitativas en la materia. Estas razones son:

- Permitirían contar con un patrón de la eficacia de los derechos humanos.
- Permitirían realizar comparaciones y correlaciones a lo largo del tiempo y con otras variables.

En este sentido, es necesario contar con un patrón común de valoración de la eficacia de la Declaración de Derechos Humanos de 1948, máxime considerando que cada Estado sigue su propio proceso de positivización. Además, un sistema de evaluación cuantitativa, es decir, un sistema de indicadores, permitiría comparaciones no sólo sincrónicas sino diacrónicas, que se podrían correlacionar con la evolución de otros indicadores (sobre distribución de la renta, escolarización, urbanización, etcétera). El primer paso para llevarlo a la práctica debería ser el establecimiento de una premisa básica: que los derechos humanos no son una realidad simple y que por tanto es necesario un indicador agregado (Aymerich, p. 355).

En este punto se introduce un problema fundamental de las investigaciones empíricas sobre los derechos humanos, el de la conceptualización de los derechos humanos, que se abordará a continuación.

IV. Conceptualización de los derechos humanos

Este tema no se refiere a una definición filosófica o jurídica de los derechos humanos, sino a una definición operativa que permita contar con un sistema de evaluación cuantitativa como del que se ha estado hablando. Por ejemplo, a diferencia del índice de inflación que mide una sola realidad (la variación temporal de los precios), un indicador de derechos humanos valora realidades muy distintas entre sí (las garantías procesales de los derechos fundamentales junto con el derecho a vacaciones pagadas, por ejemplo) (Aymerich, p. 355).

Por ello, sería necesario construir indicadores independientes para cada uno de los derechos contenidos en la Declaración Universal de 1948 y después definir los criterios que permitiese sintetizarlos en un indicador global. Pero esto es imposible si antes no se ha

⁴ *Idem.*

⁵ *Ibid.*, p. 17.

establecido conceptualmente qué son los derechos humanos. No se trata de un problema teórico; sino que es preciso saber si estamos midiendo valores, normas o hechos, porque evidentemente los tres no pueden ser sumados como si fuesen una misma realidad (Aymerich, p. 355).

Así, por ejemplo, para evaluar cuantitativamente la realización de los derechos económicos y sociales hay que examinar hechos (como la disponibilidad presupuestaria para garantizar la escolarización universal o el grado de desigualdad en la distribución de la renta), mientras que el derecho al honor tendría que ser valorado en función de la existencia de normas que lo garanticen, y en función de las normas procesales que establezcan qué tipo de acciones pueden ser ejercidas para evitar atentados impunes contra este derecho. Pero agregar normas y hechos plantea problemas conceptuales premilitares sin cuya solución no puede avanzarse (Aymerich, pp. 355-356).

Para resolver este problema es necesario entender a los derechos humanos como hechos sociales concretos, y no como valores, normas o principios, porque de ese modo sería posible establecer criterios comunes para medir su eficacia.

Una vez aclarados estos problemas iniciales, se puede seguir con la exposición de las principales características de las investigaciones actuales en materia de derechos humanos.

v. Aproximación a las fuentes cuantitativas sobre derechos humanos

a. La situación actual de las fuentes de información

La materia prima disponible como base para una sociología de los derechos humanos es amplia y creciente, aunque enfrenta diversos problemas de dispersión y fiabilidad de los datos, además de otros problemas de tipo cualitativo que se puntualizan a continuación.

- *Información limitada.* Muchas fuentes de información se restringen a información muy limitada en el tiempo o en el espacio (ejemplo de ello son los informes sobre violaciones de derechos humanos en un determinado país y en un periodo breve).
- *Imposibilidad de generalización.* Si la información con la que se cuenta está limitada a contextos y momentos específicos no es posible extraer conclusiones generales por contraste con otros lugares o hacer un seguimiento temporal de la evolución en el propio país.
- *Conclusiones no confiables.* Otro problema se da cuando hay cruces de datos entre diversas fuentes, de modo que las conclusiones que pueden extraerse por comparación entre algunas de ellas pueden resultar engañosas. Esto se debe a que la aparente coincidencia de resultados no se debe a la calidad y fiabilidad de los procedimientos estadísticos empleados, sino a que hay fuentes que toman algunos datos de otras, por lo que evidentemente coinciden.
- *Sesgo ideológico.* Un principio fundamental de cualquier investigación social es orientar sus resultados de acuerdo con objetivos de veracidad y no con fines políticos. Sin embargo, pueden hacerse objeciones a la fiabilidad de algunas fuentes orientadas ideológicamente, ya sea porque son informes gubernamentales sesgados por moti-

vos de política exterior o porque son informes de organizaciones con fines de denuncia (Aymerich, p. 358).

Una vez aclaradas las principales deficiencias metodológicas de las fuentes de información, se pueden describir las fuentes que podrían ser útiles y fiables. Estas pueden ser clasificadas de la siguiente manera:

- Publicaciones periódicas con tratamiento particular de la información.
- Estudios monográficos.
- Fuentes con caracteres de periodicidad y generalidad en la información.
- Los informes de Freedom House y Charles Humana.

b. Publicaciones periódicas con tratamiento particular de la información

En el primer grupo se ubican las publicaciones periódicas que contienen información puntual sobre cuestiones relacionadas con los derechos humanos. Se trata de una muestra, no exhaustiva, de las fuentes primarias para la obtención de datos. Sólo con fines de información, se presenta una lista de algunas de estas fuentes, citadas por autor.

- *International Civil Liberties Report*
- *Amnesty International Newsletter*
- *Bulletin of Human Rights*
- *CIL Bulletin*
- *Freedom Review*
- *Human Rights Watch Quarterly*
- *Persecution of Human Rights Monitors*
- *Freedom in National Press Systems*
- *Journalism Morbidity Survey*
- *International Newsletter on Treatment and Rehabilitation of Torture Victims*
- *Report of the Task Force on Detainees of the Philippines* (Aymerich, pp. 359-360).

La lista podría continuar, pero basta como muestra del campo de publicaciones de ámbito particular, tanto en lo geográfico como en el objeto. La información contenida en ellas es muy extensa y útil con fines de denuncia, pero no es suficiente para un análisis riguroso por las siguientes razones:

- *Falta de comparación.* Entre las publicaciones periódicas que se mencionan en la lista hay muchas que pertenecen a organizaciones ideológicas o de grupos de activistas, y que cumplen una función de difusión de información sobre casos concretos, merecedores de atención y solución. El problema es que no llevan su análisis a un nivel de comparación.
- *Falta de sistematicidad.* Otro problema es que la mayoría de las publicaciones citadas no tienen el propósito de sistematicidad en cuanto a que la información cubra todo el espectro de derechos fundamentales reconocidos en la Declaración Universal de los Derechos Humanos. Así, por ejemplo, algunas publicaciones se centran en problemas relativos a detenciones arbitrarias y falta de garantías procesales, mientras otras se

ocupan de violaciones por discriminación racial o la persecución de periodistas, lo que también impide hacer un seguimiento comparativo entre diversos Estados (Aymerich, pp. 362-363).

De lo expuesto se concluye que sólo podría obtenerse un rendimiento práctico de estas fuentes de datos con miras a análisis particulares que no implicasen estudios comparativos de largo alcance. Así, estas publicaciones podrían servir de base para fijar los patrones tipo de violaciones de derechos humanos (dado que muchas de ellas se centran precisamente en documentar los casos más graves de estas violaciones), a partir de la síntesis de los elementos coincidentes en los casos documentados en correlación con factores económicos, políticos, jurídicos y otros (Aymerich, p. 363).

c. Estudios monográficos

En un nivel más avanzado de generalidad en el tratamiento de la información, se ubica una serie de estudios con los que se intenta describir en términos globales la situación de los derechos humanos. Estos estudios fueron las primeras tentativas que se generaron al respecto. Su valor reside en haber sido los precursores de la investigación cuantitativa contemporánea sobre los derechos humanos (Aymerich, p. 364).

Estos estudios tienen en común el hecho de que trataron de describir la situación general de los derechos humanos a nivel internacional, más que fijarse en violaciones localizadas en situaciones particulares (Aymerich, p. 365). En cuanto a sus debilidades cabe señalar las siguientes:

- *Falta de amplitud.* En general, se trata de tentativas no demasiado ambiciosas de hacer una tabulación de Estados en función de su nivel de respeto por los derechos humanos, pero empleando medios insuficientes en la mayor parte de los casos.
- *Insuficiencia de información propia.* Entre otras cosas, casi ninguno opera con su propio sistema de información, por lo que su fiabilidad depende de terceras fuentes de datos.
- *Carencia de seguimiento.* Por otra parte, emplean una metodología relativamente sencilla, por lo que hacen segmentaciones en el tiempo tan amplias que las diferencias en intervalos menores desaparecen, impidiendo hacer seguimientos más pormenorizados.
- *Falta de comparación.* Seleccionan los Estados objeto de análisis en función de variables como los niveles de desarrollo, que impiden comparaciones con los Estados excluidos de tales categorías.
- *Definiciones ambiguas.* Emplean definiciones operativas demasiado ambiguas de los objetos de análisis. Como se verá, éste es un punto crucial de la metodología de un análisis sociológico de los derechos humanos basado en información cuantitativa y técnicas estadísticas (Aymerich, pp. 365-366).

d. Fuentes con caracteres de periodicidad y generalidad en la información

En este tercer nivel, se colocan las fuentes de información más operativas, es decir, fuentes documentales que presentan características de generalidad, periodicidad y en algunos ca-

tos tratamiento sistemático de la información (Aymerich, p. 379). Estas fuentes son: *Amnesty International Reports*, *Human Rights in Developing Countries*, *Country Reports on Human Rights Practices*.

Amnesty International Reports. Este informe es publicado por Amnistía Internacional con periodicidad anual. Cubre prácticamente la totalidad de los países. Sin embargo, no cumple exigencia alguna de tratamiento sistemático de la información. Este hecho impide emplear los informes de Amnistía Internacional como un indicador suficiente de la situación de los derechos humanos a nivel internacional (Aymerich, pp. 379-380).

Human Rights in Developing Countries. Se trata de una publicación anual editada por los centros para los derechos humanos de Noruega, Dinamarca, Holanda, Finlandia y Canadá. El propósito de esta publicación es orientar la adjudicación de fondos de ayuda al desarrollo en función del respeto a los derechos humanos en los países receptores. Sin embargo, esto limita el ámbito de análisis a trece países donde tienen influencia política y económica (Aymerich, p. 381).

Country Reports on Human Rights Practices. Es una publicación del Departamento de Estado de Estados Unidos, que también nació con la intención de vincular la ayuda externa con el grado de respeto de los derechos humanos de los Estados receptores. Se trata de un anuario que se publica desde 1976. Incluye información sobre respeto a los derechos humanos y sobre la situación económica, social y cultural de cada Estado. Sin embargo, no hay un patrón sistemático para la evaluación de los derechos humanos, de forma que aspectos analizados en un país no son ni siquiera mencionados en otros (Aymerich, p. 382).

e. Los informes de Freedom House y Charles Humana

Estos informes cumplen con los requisitos de periodicidad, tratamiento sistemático de la información y generalidad, tanto geográfica como en la unidad de análisis.

The Annual Survey of Political Rights and Civil Liberties. Este informe es editado por Freedom House de Nueva York/Washington, es un informe anual que cubre todos los Estados independientes del mundo y los territorios dependientes. Incluye informes particulares sobre la situación de los derechos civiles y políticos de cada Estado (Aymerich, p. 384).

World Rights Guide. Es un informe periódico elaborado por Charles Humana, un investigador académico cuyo trabajo tiene la ventaja de que no está a cargo de alguna institución política o asociación de activistas, lo que contaría como indicio de sesgo político (Aymerich, pp. 390).

vi. Objeciones a las fuentes cuantitativas y a las conclusiones teóricas

Pese a que estos informes son los más rigurosos en términos metodológicos, se reconocen tres objeciones que de manera general pueden imputarse a dichas fuentes:

a. Objeciones al sistema de obtención de datos

Se trata de la fiabilidad que puede tener la información cuando se consideran las dificultades de acceso a ella, lo que permite sospechar que buena parte de los datos que actual-

mente se manejan pueden estar distorsionados por la existencia de importantes lagunas (Aymerich, p. 401).

Otro problema se presenta cuando se emplea un sistema comparativo de la situación de los derechos humanos que resulte en una clasificación de los diversos Estados en una tabla por puntos (como el caso de Charles Humana o Freedom House). Esta metodología trata a los Estados como unidades indivisibles, pero a veces las violaciones no son emprendidas por el Estado en su conjunto, sino por determinados órganos (Aymerich, p. 402).

b. Objeciones desde los criterios estadísticos de fiabilidad

Uno de los problemas más detectados en los sistemas desarrollados hasta el momento para evaluar comparativa y cuantitativamente la situación internacional de los derechos humanos es el de su escasa fiabilidad atendiendo a los criterios habituales requeridos por las técnicas estadísticas. Esta circunstancia se hace tanto más grave cuantas más conclusiones teóricas se extraigan de informes plagados por deficiencias de este tipo en el procesamiento de la información (Aymerich, p. 405).

c. Objeciones desde el multiculturalismo

Finalmente, una de las objeciones más decisivas con respecto a los sistemas de evaluación de la eficacia de las normas internacionales sobre los derechos humanos es el problema de la diversidad cultural frente a la univocidad de tales normas (Aymerich, p. 415).

B. MATERIAL DE ESTUDIO

Bibliografía obligatoria

Aymerich, Ignacio, "La investigación cuantitativa y estadística sobre los derechos humanos", en *Sociología de los derechos humanos. Un modelo weberiano contrastado con investigaciones periódicas*, Valencia, Tirant Lo Blanch/Universitat de Valencia, 2002, pp. 363-567 **[Lectura]**.

La investigación cuantitativa y estadística sobre los Derechos Humanos

4.1. POSIBILIDADES Y UTILIDAD DE LA INVESTIGACIÓN CUANTITATIVA SOBRE LOS DERECHOS HUMANOS

Con el análisis de la investigación empírica en sociología de los derechos humanos nos introducimos en un campo netamente diferenciado del de la teoría sociológica. En realidad, no tendría por qué haber necesariamente una distancia tan marcada, y las investigaciones de ambos campos deberían beneficiarse del conocimiento mutuo de sus resultados. Si no ha sido así se debe, en buena medida, a razones meramente coyunturales, como por ejemplo que en los trabajos empíricos se han ocupado sobre todo investigadores norteamericanos entre los que tiene un peso importante la formación estadística, mientras la teoría sociológica es obra más bien de juristas o sociólogos europeos, y no ha habido hasta ahora demasiado contacto entre los dos grupos. Sin embargo, no se trata de deficiencias insalvables, y la necesidad de la coordinación entre la investigación empírica y el enfoque jurídico y de la teoría sociológica sobre los derechos humanos se hace cada vez más patente.

Renato Treves, por ejemplo, en relación con el proceso de positivación e internacionalización de los derechos humanos, esperaba de la sociología del derecho el desarrollo de investigaciones encaminadas a «verificar y a medir la eficacia de las normas

Tomada de: Ignacio Aynerich, "La investigación cuantitativa y estadística sobre los derechos humanos", en *Sociología de los derechos humanos. Un modelo weberiano contrastado con investigaciones periódicas*, Valencia, Tirant Lo Blanch/Universitat de Valencia, 2002, pp. 363-567.

estatales e internacionales relativas a los derechos [humanos]»¹. Igualmente, James R. Scarritt señala que si los responsables políticos de diversos países aspiran a la mejora de los derechos humanos de sus conciudadanos y de los de otras naciones, «los investigadores pueden ayudarlos (al tiempo que aportan una crítica útil de sus esfuerzos) a base de desarrollar un esquema para medir adecuadamente los cambios en el concepto, los patrones de mejora y la extensión de la aplicación de los derechos humanos»². También Richard P. Claude y Thomas B. Jabine insisten en que un análisis global de los derechos humanos asistido por técnicas estadísticas puede ser útil no sólo «para propósitos académicos o científicos; también puede cooperar a fortalecer la definición de las políticas públicas, la abogacía y la educación»³. El interés que estos —y otros— autores manifiestan por el proyecto de un sistema de verificación de la eficacia de las normas relativas a los derechos humanos responde, además, a una necesidad que se ha ido manifestando paulatinamente y que tiene causas concretas.

Como ha señalado Antonio Cassese⁴, los abusos contra la dignidad de las personas no son nada nuevo en la historia, pero desde 1948 se ha producido un cambio decisivo en la escena internacional. Cuando, con anterioridad a esa fecha, se producía un atentado contra las personas en la forma de torturas, discriminaciones o desapariciones violentas en un Estado, la opinión pública de las demás naciones podía formular juicios o condenas, que sin

¹ TREVES, Renato. *Diritti umani e sociologia del diritto*. En «Sociologia del diritto». XVI/1989/1. p. 10

² SCARRITT, James R. *Definitions, dimensions, data and designs*. En NANDA, P., SCARRITT, James R. y SHEPHERD, George W (eds). *Global Human Rights: Public policies, comparative measures and NGO strategies*. Westview Press. Boulder (Colorado), 1981. p. 119. (La traducción es mía).

³ CLAUDE, Richard P y JABINE, Thomas B. *Exploring human rights issues with statistics*. En CLAUDE, Richard P y JABINE, Thomas B. (eds.). *Human Rights and Statistics. Getting the record straight*. University of Pennsylvania Press. Philadelphia, 1992. p. 5. (La traducción es mía)

⁴ CASSESE, Antonio. *Los derechos humanos en el mundo contemporáneo*. Ed. Ariel. Barcelona, 1991. p. 7.

embargo no pasaban nunca del ámbito bilateral, pues no había instancias internacionales desde las que establecer los hechos, calificarlos y, en su caso, emprender acciones. El patrón de medida para tales condenas no podía ser otro que los valores occidentales, reflejados en las constituciones de los Estados que eventualmente realizaban tales juicios. Sin embargo, desde 1948, año de la Declaración Universal de los Derechos Humanos, todos los países cuentan con un código internacional con respecto al cual es posible decidir cuándo, en qué medida y con qué responsabilidad ha habido atentados contra la dignidad de las personas. Lo que antes no pasaba de ser una condena moral puede ahora ser calificado con mucha mayor precisión como «genocidio», «tratos o penas inhumanos, crueles o degradantes», «discriminación por razón de raza», etc. Por tanto, como señala Cassese, las normas internacionales sobre los derechos humanos de 1948-1966 representan un progreso en muchos sentidos, pero particularmente (aún cuando éste no sea tal vez el aspecto más importante pero sí el que en este momento interesa) significan un avance en relación a la objetividad de un patrón de medida internacionalmente reconocido.

Ahora bien, media todavía un paso entre contar con normas internacionalmente reconocidas de derechos humanos en las que se tipifique con precisión las conductas indebidas, y la eficacia plena de tal sistema en cuanto a establecer inequívocamente la información decisiva sobre estos derechos, lo que sólo será posible cuando la calificación jurídica de los hechos no pueda ser impedida a instancia de parte. William M. Evan, tras analizar la importancia de los Pactos Internacionales de Derechos Civiles y Políticos y de Derechos Sociales, Económicos y Culturales en orden a establecer mecanismos jurídicamente vinculantes en relación con los derechos humanos, se pregunta:

«¿Hasta qué punto cumplen los Estados con lo establecido en estos dos Convenios sobre derechos humanos? No puede darse una respuesta definitiva por la ausencia de un servicio comprensivo, internacional e imparcial de observación y documentación que periódica y

sistemáticamente recoja datos comparables sobre el cumplimiento por parte de los Estados de la gama de normas sobre derechos humanos»⁵.

El único medio previsto en tales normas internacionales para obtener información acerca del respeto de los derechos humanos se inició con la resolución 624 B, de 1 de Agosto de 1956, del Consejo Económico y Social, adoptada a instancias de la Comisión de Derechos Humanos, resolución en la que se instaba a los Estados a elaborar informes periódicos sobre su situación interna. Este mismo procedimiento se incluiría más adelante en la cuarta parte del Pacto Internacional de Derechos Civiles y Políticos de 1966, que arbitra medios para obtener información sobre la situación internacional de los derechos humanos. Sin embargo, la operatividad de esta vía depende de nuevo de los informes que cada Estado queda obligado a remitir al Comité de Derechos Humanos y en los que da cuenta de las decisiones que ha adoptado para dar efecto a los derechos reconocidos en el Pacto⁶. Ahora bien, como expone Bertie Ramcharan, el cumplimiento de esta obligaciones está lejos de ser pleno. Así, por ejemplo, el artículo 9 de la Convención internacional para la eliminación de todas las formas de discriminación racial de 1965 obliga a las partes a presentar informes periódicos detallando las medidas adoptadas para dar cumplimiento a las disposiciones de la Convención. Sin embargo, en su sesión anual de 1990, el Comité para la eliminación de todas las formas de discriminación racial constató que...

«...desde el 15 de Marzo de 1982 estaba pendiente la entrega de no menos de 89 informes referentes a 62 Estados, de los cuales 42 correspondían a 15 Estados, cada uno de ellos con 2 ó 3 informes pendientes, y que

⁵ EVAN, William M. «Human Rights, the nation-state and transnational law». En FERRARI, Vincenzo (ed.). *Laws and Rights. Proceedings of the International Congress of Sociology of Law for the Ninth Centenary of the University of Bologna* (May 30-June 3, 1988). Dott. A. Giuffrè Editore. Milano, 1991. p. 332.

⁶ Cfr. artículo 40.1 del Pacto Internacional de Derechos Civiles y Políticos.

cuatro informes iniciales que debían haberse entregado entre 1973 y 1978 todavía no se habían recibido»⁷.

También el artículo 16 del Pacto internacional de derechos económicos, sociales y culturales obliga a los Estados que lo hayan ratificado a presentar informes sobre las medidas adoptadas para la plena realización de estos derechos. Ramcharan comenta al respecto que, de manera semejante al caso anterior, en la sesión anual del Comité de Derechos Económicos, Sociales y Culturales de 1991 se puso de manifiesto que 31 Estados no habían llegado a enviar informes de ningún tipo, a pesar de las obligaciones contraídas como signatarios del Pacto⁸. Además del problema del incumplimiento de esta obligación, es difícil que algún Estado se autoinculpe voluntariamente en tales informes. Por otra parte, se contempla también la posibilidad de que un Estado signatario del Pacto, en caso de que considere que otro Estado no cumple las disposiciones del mismo, se lo podrá señalar mediante comunicación escrita, y si en el plazo de tres meses no recibe una explicación satisfactoria podrá ponerlo en conocimiento del Comité⁹, ante lo cual éste iniciará un procedimiento tendente a la mediación y conciliación entre el Estado denunciante y el denunciado¹⁰.

En cualquier caso, el Comité no ejerce papel alguno en el establecimiento de los hechos que sea equivalente a la fase de instrucción en un proceso judicial, ni cuenta con capacidad coactiva para imponer estas normas, y en el procedimiento descrito sólo puede ejercer funciones de arbitraje entre intereses de partes que no quedan obligadas a someterse a autoridad superior. Lo que

⁷ RAMCHARAN, Bertie G. *A reporting system for international humanitarian law?* En BELLO, Emmanuel G and Prince Bola A. Ajibola, SAN. *Essays in Honour of Judge Taslim Olawale Elias*. Vol. I: *Contemporary International Law and Human Rights*. Martinus Nijhoff, Dordrecht/Boston/London, 1992. p. 131.

⁸ Idem. nota 5, p. 131

⁹ Artículo 41.1 del Pacto Internacional de Derechos Civiles y Políticos.

¹⁰ Artículos 41.1 y 42 del Pacto Internacional de Derechos Civiles y Políticos.

queda garantizado en el propio artículo 41.1, que considera dispositivo de cada Estado el reconocimiento de la competencia del Comité para recibir y examinar las comunicaciones en que un Estado alegue que otro no cumple con sus obligaciones derivadas del Pacto. Tampoco se asignan competencias al Comité para sancionar a los Estados que incurran en falsedades en la información que remiten, ni capacidad para obligar a los Estados a que en sus informes se ciñan necesariamente a un conjunto previamente determinado de cuestiones a evaluar. Sólo a partir de 1992 comenzó el Comité a solicitar informes específicos por propia iniciativa cuando se daban situaciones excepcionales respecto a los derechos humanos. De lo expuesto se deduce que, en cuanto a la obtención de información sobre la eficacia de las normas internacionales sobre los derechos humanos, no existe un procedimiento suficientemente fiable y cuyos datos puedan ser establecidos de manera incuestionable.

No obstante, como señala Ineke Boerefijn, el Comité de Derechos Humanos planteó en 1978 la posibilidad de emplear informes no oficiales (es decir, distintos de sus propios informes) en sus sesiones, a lo que se opusieron algunos Estados, prevaleciendo el criterio de que no podía hacerse uso de ellos. No obstante, «miembros del Comité han usado ya desde bien pronto información de la Comisión Internacional de Juristas, Amnistía Internacional y otras organizaciones no gubernamentales de derechos humanos»¹¹. El uso de esta información se ha generalizado recientemente, y hoy en día ya no es objeto de discusión en el Comité. Así, por ejemplo —como menciona Boerefijn—, en su segundo informe periódico sobre la situación de los derechos humanos en Egipto, el Comité se refirió explícitamente a «numerosas alegaciones documentadas por fuentes fiables no gubernamentales de

¹¹ BOEREFIJN, Ineke. «Toward a strong system of supervision: the Human Rights Committee's role in reforming the reporting procedure under article 40 of the Covenant on Civil and Political Rights». En *Human Rights Quarterly*. Vol 17. n° 4. Noviembre de 1995. p. 784.

información»¹². Lo que avala la opinión de que el desarrollo de sistemas de verificación de la situación de los derechos humanos con suficientes garantías y que aporten información relevante tiene utilidad inmediata, aun cuando en el caso del Comité de Derechos Humanos todavía no puede hablarse más que de tentativas que están en sus comienzos y que el propio artículo 40 del Pacto Internacional de Derechos Civiles y Políticos impide o, al menos, obstaculiza desarrollar más adecuadamente.

En Julio de 1998 se acordó en Roma la creación de un Tribunal Penal Internacional con competencia en ciertos delitos contra los derechos humanos, lo que puede llegar a implicar un nivel de eficacia significativamente mayor que las del existente Comité nacido del Pacto Internacional de Derechos Civiles y Políticos o instancias similares, fundamentalmente porque para la fase de instrucción habrá que arbitrar mecanismos probatorios más consistentes que concluyesen en un establecimiento de los hechos con garantías. Tanto en relación con este tribunal como con el Comité de Derechos Humanos habría diversas consideraciones importantes que hacer, como las relativas a la jurisdicción, la imputabilidad o la ejecución de las sentencias, pero lo único que en este momento se considera es el aspecto de la obtención de información fiable, que se concretaría en una instancia superior a las partes que, en relación con los derechos humanos, cuente con medios probatorios para establecer objetiva y fiablemente los hechos y con criterios procesales para que su calificación jurídica quede al margen de intereses particulares. En ausencia, todavía, de jurisprudencia de este tribunal, la disponibilidad de esta fuente de datos no ha pasado de ser un proyecto, y esto dejando aparte la cuestión de que los hechos probados abarcarían un conjunto de delitos que sólo se refieren a una pequeña parte de las normas internacionales sobre derechos humanos. Habría que completarlos con otras instancias

¹² Comments of the Human Rights Committee, Egypt. U.N. Document CCPR/C79/Add. 23 § 10 (1993). Adoptado en la sesión 48ª del Comité, de 26 de Julio de 1993. Citado por BOEREFIJN, Op. Cit. p. 785.

que evaluarán las normas que amparan la participación política, o la libertad de prensa, o los derechos económicos y sociales.

De todo ello se deduce que, en ausencia de esta instancia superior, la información sobre los derechos humanos está supeditada en cierta medida a los intereses políticos de los Estados. Cada uno de ellos adapta con frecuencia su evaluación del nivel de respeto de los derechos humanos en otros Estados en función de las directrices de su política exterior o de sus intereses comerciales. Pero, a su vez, la crítica a la posible falta de objetividad de estos sistemas de evaluación suele hacerse desde los intereses contrapuestos de otros Estados, o bien desde organizaciones políticas o marcadamente ideológicas. Si hubiese una metodología diseñada con suficientes garantías y mediante la cual pudiese obtenerse información verificable, se contaría con una tercera instancia que podría arbitrar estos intereses encontrados, y buena parte de la información que circula sobre los derechos humanos podría ser contrastada con el mínimo exigible de rigor.

Es precisamente en este punto donde se inserta la posibilidad de un sistema de investigación sobre la eficacia de las normas internacionales sobre derechos humanos que pueda servir de referencia y cuyos datos sean menos cuestionables (en la medida de lo posible), por su objetividad y por su independencia respecto a los intereses políticos de los Estados. El Ministerio de Asuntos Exteriores de Holanda, por ejemplo, ha hecho público su compromiso de supeditar la ayuda externa a Estados del tercer mundo a las violaciones de los derechos humanos documentadas en esos países por los informes de Amnistía Internacional¹³. El procedimiento de obtención y difusión de la información cumpliría, de darse la

¹³ Cfr. BAEHR, Peter R. «Concern for development aid in fundamental rights: the dilemma as faced by The Netherlands». *Human Rights Quarterly*. n.º 4/1 (1982). pp. 39-52. Por otra parte —y como se verá más adelante— Holanda, junto con otros estados europeos, ha explicitado más esta vinculación entre ayuda externa y respeto de los derechos humanos evaluado con patrones que tratan de ser objetivos en el informe *Human Rights in developing countries*.

posibilidad de un informe como el aludido, unos requisitos mínimos de cientificidad, ya que los datos obtenidos podrían ser sometidos a pruebas de contrastación (de falsabilidad, en términos popperianos). Ciertamente ya hay progresos en cuanto a fuentes de información no lastradas por intereses políticos de los Estados gracias a la labor de diversas ONGs, pero la expectativa a la que aluden los autores arriba citados se refiere más bien a fuentes de información estadística, o a estudios cuantitativos y basados en una comparación sistemática de la situación de los derechos humanos en el escenario internacional. Porque, en cuanto fuente de información contrastable, estos datos sólo serían realmente útiles si permitiesen análisis diacrónicos a medio y largo plazo, a la vez que análisis sincrónicos que abarcasen la totalidad de Estados del mundo. Un sistema de información como este no puede manejarse eficazmente si el enorme volumen de datos que implica no se sintetiza de alguna manera a través de algún medio de cuantificación. Este es el tipo de información a que aludían Treves, Scarritt o Claude, arriba citados, y el que motiva que la convicción sobre la utilidad de investigaciones estadísticas relativas a los derechos humanos se haya ido consolidando en los últimos tiempos, y como consecuencia se haya producido un incremento de la información documental disponible, como habrá de verse. No obstante, cabe ir señalando ya que las tentativas realizadas hasta ahora no han conseguido una metodología suficientemente consolidada, y que el momento presente está más bien marcado por el debate acerca de las posibles direcciones a emprender en orden a un sistema aceptable para la comunidad científica internacional.

Si bien la utilidad de un sistema como el aludido no es cuestionada, no puede ocultarse, sin embargo, que las tentativas llevadas a cabo están orientadas a menudo por una excesiva fe en las virtualidades de la investigación cuantitativa, que parece manifestar la permanencia de ideas positivistas en la metodología científica. Me refiero a que la comprensión de los derechos humanos no tiene mayores garantías de objetividad o de rigor científico por el simple hecho de que sus resultados puedan

expresarse en cifras. A veces la aplicación de técnicas estadísticas parece poder bastarse para construir una argumentación, como si todo fuese cuestión de hacer bien las cuentas. Por ello habrá de hacerse una evaluación crítica de los resultados obtenidos y habrá que examinar también los intentos de definir sistemas de verificación de la vigencia de los derechos humanos orientados sobre otras opciones metodológicas.

De entrada debe quedar claro que no existe ningún sistema capaz de *medir* los derechos humanos; pretenderlo sería poco menos que un engaño. La utilidad de los sistemas de análisis cuantitativo está en que pueden ofrecer un punto de contraste, un referente más o menos indirecto sobre el que poner a prueba las hipótesis sobre el grado de desarrollo de los derechos humanos. Pero si se toman las cifras desnudas, al margen de todo modelo teórico de desarrollo, e interpretando cada dato individual como un reflejo exacto de la realidad, entonces la metodología cuantitativa puede volverse tan injustificable como su contrario; la suposición gratuito de hipótesis sin ningún esfuerzo por contrastarlas.

En primer lugar, la elección de los criterios de cuantificación no puede resolverse exclusivamente en términos de adecuación entre la metodología y el objeto. Tomando como ejemplo el caso de la temperatura, la elección de un sistema de cuantificación Fahrenheit o por grados centígrados no puede decidirse en términos de cuál es más *verdadero*, ni cabe decir que el estado de temperatura de un cuerpo esté expresado con más *exactitud* eligiendo uno u otro de los sistemas. Los criterios para resolver esta opción son contingentes, y la validez del sistema adoptado descansa únicamente en el consenso de la comunidad científica. Ahora bien, debe separarse esta validez consensual del sistema de cuantificación de la validez de cada medición particular, que no queda relativizada por la primera. Que existan diversos sistemas de *cuantificación* no significa que no existan sistemas *fiabiles* de *medición* de la temperatura.

Pasando a un grado mayor de complejidad, algo semejante ocurre con los indicadores económicos. El índice de inflación, por ejemplo, responde a un criterio de cuantificación de las variaciones

de precios en que se sintetiza de forma ponderada un amplio conjunto de variables. La metodología empleada para su elaboración puede debatirse y concluir en diversas respuestas posibles. De hecho, periódicamente es revisada. Pero si ha de tener alguna utilidad, ésta descansa en que el criterio que se haya adoptado se mantenga durante un período relativamente largo. Porque sería absurdo cambiar el criterio de cuantificación cada mes bajo el pretexto de un mayor afinamiento metodológico. En tal caso, el dato obtenido en cada ocasión no sería comparable a los restantes, y un mero dato aislado no tiene utilidad alguna. Estos indicadores ofrecen verdaderamente información en la medida en que permitan comparaciones a largo y medio plazo y en que puedan observarse las tendencias evolutivas en períodos prolongados de tiempo. Por eso resulta preferible la posibilidad de la comparación diacrónica aunque sea a costa de una pequeña merma en las posibilidades de revisión de la metodología para hacerla más adecuada¹⁴. Y queda también claro que no tiene demasiado sentido debatir si la medición hecha en un momento puntual es un reflejo verdadero de la realidad, sino más bien si en las comparaciones de ciclo largo pueden detectarse tendencias evolutivas que reflejen más o menos directamente cambios reales. La conclusión que se extrae de todo ello es que debe tomarse la información cuantitativa con un cierto margen de relatividad, para no caer en el error de suponer que cada dato es una trasposición directa de la realidad.

¹⁴ Por ejemplo, Naciones Unidas ha corregido la metodología empleada anteriormente para calcular el PIB per cápita por varias razones, entre las cuales se cuenta que el PNB se expresa en dólares USA, de forma que la medida del PNB de un determinado Estado puede verse alterada por una revalorización del dólar en los mercados financieros internacionales. En tal caso, la conversión de la moneda local provoca un índice de PNB menor sin que realmente haya descendido la tasa de crecimiento. Para valorar más adecuadamente la capacidad adquisitiva de cada moneda, Naciones Unidas ha desarrollado un indicador llamado PPA (paridades de poder adquisitivo). Las tasas de desarrollo medidas con el PPA varían respecto a las medidas con el PNB empleado anteriormente. Pero esto no significa que los datos obtenidos hasta ahora fuesen *falsos*, y mucho menos que fuesen inútiles.

Sin embargo, con todas las cautelas que sean necesarias, un sistema cuantitativo de evaluación de las normas internacionales sobre derechos humanos no carece de utilidad. En primer lugar, si se trata de verificar la eficacia de la Declaración de Derechos Humanos de 1948, teniendo en cuenta que cada Estado sigue su propio sistema de positivación constitucional en forma de derechos fundamentales (lo que a menudo no permite comparaciones lineales), y que la vigencia formal de las normas no implica necesariamente su eficacia, se hace necesario diseñar un patrón común de valoración de dicha eficacia. Y segundo, un sistema de evaluación cuantitativa, es decir, un sistema de indicadores, permitiría comparaciones no sólo sincrónicas sino diacrónicas, que se pueden correlacionar con la evolución de otros indicadores (sobre distribución de la renta, escolarización, urbanización, etc.). El primer paso para llevarlo a la práctica debería ser el establecimiento de una premisa básica: que los derechos humanos no son una realidad simple, y que por tanto es necesario un indicador agregado. A diferencia del índice de inflación, ya citado, que mide *una* sólo realidad (la variación temporal de los precios), un indicador de derechos humanos valora realidades muy distintas entre sí (las garantías procesales de los derechos fundamentales junto con el derecho a las vacaciones pagadas, por ejemplo). Sería necesario construir indicadores independientes para cada uno de los derechos contenidos en la Declaración de 1948 y después definir los criterios de agregación que permitiesen sintetizarlos en un indicador global. Pero esto es imposible si antes no se ha establecido conceptualmente qué son los derechos humanos. No se trata de un problema teórico; es que es preciso saber si estamos midiendo valores, normas o hechos, porque evidentemente los tres no pueden ser sumados como si fuesen una misma realidad. Así, por ejemplo, para evaluar cuantitativamente la realización de los derechos económicos y sociales hay que examinar hechos (como la disponibilidad presupuestaria para garantizar la escolarización universal o el grado de desigualdad en la distribución de la renta), mientras que el derecho al honor tendrá que ser valorado en función de la existencia de normas que lo garanticen, y además en función de las normas procesales que establezcan qué tipo de

acciones pueden ser ejercidas para evitar atentados impunes contra este derecho. Pero cómo agregar normas y hechos plantea problemas conceptuales preliminares sin cuya solución no puede avanzarse.

Como he discutido en el capítulo 1, existen diversos modelos de conceptualización, algunos de los cuales conducen a contradicciones y problemas. La definición conceptual idónea es la que examina los derechos humanos como pretensiones típicas de legitimidad de un sistema político con un derecho racionalizado, porque de esta manera las proposiciones normativas en que se formulan los derechos fundamentales pueden ser evaluadas fácticamente, en la medida en que los ciudadanos acepten de hecho la legitimidad del sistema, es decir, en la medida en que obedezcan estas normas, y esto puede comprobarse empíricamente. A partir de aquí queda claro que evaluamos siempre hechos, y que por tanto no estamos agregando realidades heterogéneas. Igualmente, a partir de esta conceptualización podemos describir el proceso típico por el que las comunidades políticas llegan a formular por medio de los derechos humanos sus pretensiones de legitimidad: monopolio de la violencia legítima, burocratización del cuadro administrativo, formalización y racionalización del derecho, etc. Lo que mediría el indicador es el grado de avance en una escala unidireccional prefijada: el proceso de implementación de los derechos humanos, conceptualmente definido en sus principales pasos. La propuesta del modelo de base weberiana que he hecho en la primera parte va encaminada precisamente en esta dirección. Pero debo ir anticipando ya que la mayor parte de la investigación empírica ha descuidado esta premisa metodológica, pasando a analizar cuantitativamente los derechos humanos sin preguntarse qué tipo de realidad se estaba evaluando. Subyace a todas estas investigaciones el modelo normativista (expuesto en el primer capítulo) pero sin que se haya llegado a tomar conciencia de la opción por un modelo entre otros posibles. Pero las implicaciones metodológicas y de conceptualización serán examinadas una vez que se hayan expuesto las principales investigaciones empíricas.

El objeto de este capítulo es, por tanto, presentar el estado actual de tales investigaciones sobre la eficacia de las normas de

derechos humanos, investigaciones que tienen una base fundamentalmente cuantitativa y en que se emplean técnicas estadísticas. El examen de estas investigaciones requiere, sin embargo, distinguir los distintos niveles en que se estructura este campo, la tipología de las investigaciones llevadas a cabo hasta ahora o las críticas básicas a diseños metodológicos empleados. Debe advertirse, además, que se trata de un campo bastante heterogéneo, en el que buena parte de las investigaciones no se realizan bajo el epígrafe «sociología empírica de los derechos humanos», y en que los investigadores son politólogos, juristas, sociólogos, expertos en estadística, en relaciones internacionales, etc. La heterogeneidad de los trabajos requiere que para presentarlos con un mínimo de orden se estructuren de acuerdo con algún criterio clasificatorio. Con ese propósito, examinaré el conjunto de fuentes actualmente disponibles siguiendo un esquema de niveles de profundidad, comenzando por aquellas que ofrecen datos sobre situaciones particulares en relación con los derechos humanos, para ir avanzando hacia las que abarcan datos sobre períodos de tiempo más largos (que permitan comparaciones diacrónicas) y sobre más un espectro geográfico más amplio (siendo la situación óptima la de las fuentes que ofrecen datos sobre todos los Estados del mundo). Evidentemente, prestaré atención preferente a este último tipo de datos. Examinaré también en este capítulo las principales objeciones metodológicas hechas a las principales fuentes de datos y sus procedimientos de selección y tratamiento de la información cuantitativa.

4.2. PRIMERA APROXIMACIÓN A LAS FUENTES CUANTITATIVAS SOBRE DERECHOS HUMANOS

4.2.1. La situación actual de las fuentes de información

El material «bruto» disponible como base para una sociología de los derechos humanos es amplio y creciente, aunque cuenta con problemas de dispersión y fiabilidad de los datos, además de otros

problemas de tipo cualitativo. Entre estos problemas, cabe destacar que muchas fuentes se restringen a información muy limitada en el tiempo o en el espacio (por ejemplo, informes sobre violaciones de derechos humanos en un determinado país y en un período breve), lo que impide tanto extraer conclusiones generales por contraste con otros lugares, como hacer un seguimiento temporal de la evolución en el propio país. Otras fuentes ofrecen información cuantitativa, pero en el contexto de un estilo narrativo que impide un tratamiento sistemático de dicha información —y, desde luego, impide también cualquier análisis comparativo en el tiempo o en el espacio—. Otro problema lo representa el que se den casos de cruce de datos entre diversas fuentes, de modo que las conclusiones que puedan extraerse por comparación entre algunas de ellas pueden resultar engañosas, en el sentido de que la aparente coincidencia de resultados no se debe a la calidad y fiabilidad de los procedimientos estadísticos empleados, que por distintas vías hubiesen llegado a idénticos resultados, sino sencillamente a que hay fuentes que toman algunos datos de otras, por lo que evidentemente coinciden. Este problema ha dado lugar a interpretaciones erróneas precisamente por no haber tenido en cuenta este factor. Por último, pueden hacerse objeciones a la fiabilidad de algunas fuentes orientadas ideológicamente, o por ser informes gubernamentales sesgados por motivos de política exterior del país correspondiente, etc.

Por estas razones, es necesario comenzar desbrozando el terreno de las fuentes de información para depurar aquellas que verdaderamente pueden ser útiles y relativamente fiables. Con ese fin, en una primera aproximación, presentaré un panorama de aquellas fuentes disponibles pero de menor relevancia en virtud de alguna de las deficiencias señaladas arriba.

4.2.2. Publicaciones periódicas con tratamiento particular de la información

El primer grupo es un conjunto de publicaciones periódicas que contienen información puntual sobre un abanico amplio de

cuestiones relacionadas con los derechos humanos. Se trata de una muestra, no exhaustiva, de las fuentes primarias para la obtención de datos.

- *International Civil Liberties Report*. Editado por ACLU (American Civil Liberties Unión. International Human Rights Task Force) de Los Angeles. Revista semestral. A pesar del nombre, se limita a los Estados Unidos, ocupándose de cambios legales que afecten a los derechos humanos, ratificaciones de tratados internacionales, litigación en tribunales estadounidenses, etc.
- *Amnesty International Newsletter*. London. Mensual. Contiene artículos puntuales sobre casos de violaciones de derechos humanos y los esfuerzos de Amnesty International al respecto. Esta publicación, sin embargo, resulta mucho más limitada que el informe anual de Amnesty International, que se analizará más adelante.
- *Bulletin of Human Rights*. Ginebra. Revista trimestral. Informa sobre acciones de la ONU relativas a los derechos humanos (ratificaciones de convenios, actividades de comités relacionados, etc.)
- *CIJL Bulletin*. Publicado por el Centre for the independence of judges and lawyers de Ginebra. Aparece con periodicidad semestral desde 1979. Se interesa particularmente por la independencia del poder judicial, y sus informes se limitan a las violaciones de derechos humanos que afecten a jueces y abogados.
- *Freedom Review*. Washington. Publicación de Freedom House. Publica informes sobre violaciones y progresos en derechos humanos. Esta misma organización publica otro informe anual que sí presenta mayor interés, como veremos.
- *Human Rights Watch Quarterly*. Nueva York. Casi una carta para los miembros de la organización Human Rights Watch. Informes puntuales sobre derechos humanos y sobre actividades de la organización.
- *Persecution of Human Rights Monitors*. Publicado también por la organización Human Rights Watch. Se trata de una publica-

ción anual con cobertura mundial que comenzó a publicarse en 1987. Se centra en la persecución que sufren los observadores e informantes sobre violaciones de los derechos humanos, o personas conectadas con los anteriores. Obtiene su información de las ONG afiliadas al proyecto.

- *Freedom in National Press Systems*. Publicado por el International Press Institute de Vitikonwaldegg (Suiza). Se trata de un estudio comparativo de los diversos sistemas de protección legal de la libertad de prensa en el mundo, y establece correlaciones entre ésta y otras variables como el índice de alfabetización o el ingreso per cápita. No tiene cobertura mundial, y sólo se han publicado dos ediciones (en 1959 y 1964)
- *Journalism Morbidity Survey*. Editado personalmente por Leonard R. Sussman, si bien tiene apoyo de la organización Freedom House. Se publica anualmente desde 1982. Parte del supuesto de que el flujo libre de informaciones es condición esencial para la eficacia de los derechos humanos, y por tanto analiza la persecución o los obstáculos puestos a la actividad profesional de los periodistas como variable indicativa de la vigencia de tales derechos.
- *International Newsletter on Treatment and Rehabilitation of Torture Victims*. Copenhague. Intenta promover intercambio internacional de información entre profesionales dedicados a víctimas de la tortura, aportando experiencias sobre programas de rehabilitación y reinserción social, etc.
- *Interights bulletin* (International Centre for the Legal Protection of Human Rights. London). Básicamente contiene decisiones judiciales en aplicación de la legislación internacional sobre derechos humanos, pero está casi totalmente dedicada a cuestiones europeas y del Tribunal Europeo.
- *Reports of the Task Force on Detainees of the Philippines*. Publicado por la organización del mismo nombre. Documenta sistemáticamente los casos de detenciones ilegales y torturas durante los mandatos de F. Marcos y C. Aquino. Mensual desde 1986 (antes irregular).

- *Human Rights Update of South Africa*. Equivalente al anterior aunque referido a Sudáfrica. Incluye también violaciones de los derechos humano, especialmente las que eran consecuencia de las leyes de discriminación racial. Tiene una periodicidad trimestral.
- *African Human Rights Newsletter*. Publicación del African Centre for Democracy and Human Rights. Banjul (Gambia). Trimestral. Información sobre la situación de los derechos humanos en Africa, actividades de ONGs africanas, de la Comisión Africana para los Derechos Humanos y de los Pueblos, etc.
- *Constitutional Rights Journal*. Editado por el Constitutional Rights Project de Lagos (Nigeria). Trimestral. Aunque teóricamente se refiere a toda Africa, se centra principalmente en la situación de los derechos humanos en Nigeria.
- *Aliran Monthly*. Publicado por la organización Aliran Kesedaran Negaran en Malasia. Informa sobre derechos humanos en Malasia atendiendo particularmente a los abusos cometidos contra los derechos políticos, y la protección estatal de los derechos económicos y culturales.
- *China Rights Forum*. Publicado por la organización Human Rights in China de Nueva York. Trimestral. Es la publicación de un grupo de activistas chinos exiliados y documenta abusos específicos en curso, además de artículos de opinión.
- *Turkish Yearbook on Human Rights*. Publicado por el Instituto de Administración Pública de Turquía y Oriente Medio/Centro de Investigación y Documentación sobre Derechos Humanos, de Ankara. Es un volumen anual que se ocupa sobre cuestiones de derechos humanos en Turquía, pero sirve más bien a propósitos propagandísticos. Turquía también ha creado un Ministerio de Derechos Humanos sin dejar por ello de perseguir al pueblo Kurdo.
- *HRCP Newsletter*. (Human Rights Commission of Pakistan). Lahore (Pakistan). Trimestral. Se trata también de la publicación de un grupos de activistas pakistaníes, que documenta abusos y discriminaciones sistemáticas de ciertos grupos políticos y étnicos.

La lista se podría continuar, pero basta como muestra del campo de publicaciones de ámbito particular, tanto en lo geográfico como en el objeto. La información contenida en ellas es muy extensa, pero no suficiente para un análisis riguroso por diversas razones, como ya se apuntaba más arriba. En primer lugar, entre las publicaciones periódicas hay muchas que pertenecen a organizaciones ideológicas o de grupos de activistas, y que cumplen una función de difusión de información sobre una casos concretos, merecedores de atención y solución —algo que evidentemente no tiene nada de censurable—, pero en el ánimo de sus autores no existe interés, lógicamente, por llevar su análisis a un nivel general de comparación. El problema que esto representa es que, por ejemplo, puede contarse con información puntual sobre violaciones de los derechos humanos en Pakistán o Filipinas, pero no en Bangla Desh o en Mongolia, donde no se editan publicaciones equivalentes, y por tanto sobre estos datos no pueden hacerse seguimientos sistemáticos y comparativos en los diversos Estados de la escena internacional. Por otra parte, cuando el objetivo prioritario es llamar la atención sobre una situación grave de abuso de los derechos humanos para que la comunidad internacional adopte medidas prácticas (como es el caso en muchas de estas publicaciones), se emplea un estilo narrativo en el que los conceptos empleados no necesariamente están definidos con la suficiente precisión, y cada publicación emplea su propia terminología. Esto representa un problema para cualquier análisis comparativo o estadístico, donde se precisa que las variables medidas sean expresión de un mismo tipo de fenómenos, pues en otro caso se trataría de elementos no conmensurables y no susceptibles de un análisis cuantitativo.

Otro problema es que en la mayor parte de estas publicaciones tampoco hay propósito de sistematicidad en cuanto a que la información cubra todo el espectro de derechos fundamentales reconocidos en la Declaración Universal de los Derechos Humanos. Así, por ejemplo, algunas publicaciones se centran en problemas relativos a detenciones arbitrarias y falta de garantías procesales, mientras otras se ocupan de violaciones por discriminación racial, o la persecución de periodistas, lo que también impide hacer un seguimiento comparativo entre diversos Estados, pues en estos casos se está aportando información

sobre elementos distintos e inconmensurables. Hay que contar, además, con el inconveniente de que no puede emplearse esta información para extraer conclusiones sobre las variaciones temporales en cuanto a la eficacia de las normas internacionales sobre los derechos humanos, puesto que en pocos casos hay un compromiso de los autores de obtener y suministrar información sobre un mismo asunto con periodicidad regular. Si, por ejemplo, en un determinado lugar deja de haber casos de violaciones de algún derecho fundamental, también cesa el interés de informar sobre ese punto, pues lo decisivo es destacar los casos negativos que requieren solución, y las situaciones positivas no son entonces relevantes. Pero con ello cesa también la información que permitiría hacer un seguimiento temporal y buscar las correlaciones causales con otros factores como las variaciones en las tasas de urbanización, escolarización, etc. Por último, hay algunas publicaciones periódicas que no se dirigen a informar sobre la situación de los derechos humanos de manera directa, sino a suministrar información sobre documentación legal, encuentros académicos internacionales, ratificaciones de tratados, etc. Sin dejar de ser importante, esta información abarca un aspecto parcial de la cuestión y tampoco permite análisis comparativos relevantes.

De lo expuesto se concluye que únicamente podría obtenerse un rendimiento práctico de estas fuentes de datos en orden a análisis particulares que no implicasen estudios comparativos de largo alcance. Así, por ejemplo, estas publicaciones podrían tal vez servir de base para fijar los patrones tipo de violaciones de derechos humanos (dado que muchas de ellas se centran precisamente en documentar los casos más graves de estas violaciones), a base de sintetizar los elementos coincidentes en los casos documentados en correlación con factores económicos, políticos, jurídicos, etc. Esta posibilidad ha sido propuesta por varios autores (así, por ejemplo, V. Ferrari¹⁵), pero la tarea debería emprenderse con suficientes cautelas en atención a los problemas de falta de sistematicidad mencionados.

¹⁵ FERRARI, Vincenzo. *Sociologia dei diritti umani: riflessione conclusive*. En «Sociologia del diritto». XVI/1989/1. p. 168

4.2.3. Primeros estudios monográficos

En un nivel más avanzado de generalidad en el tratamiento de la información hay una serie de estudios en que se comienza ya a intentar describir en términos globales la situación de los derechos humanos. Estos estudios deben entenderse en el contexto de los orígenes del *social indicators movement*, y son por tanto las primeras tentativas, desarrolladas todavía con una metodología relativamente sencilla. Su valor reside en haber sido los precursores de la investigación cuantitativa contemporánea sobre los derechos humanos. Por esta razón, simplemente los mencionaré, junto con una breve descripción de sus características básicas

- Coleman¹⁶. Analiza el nivel de democracia en 75 Estados pertenecientes al grupo de los países en desarrollo, clasificándolos en una escala de tres niveles. Sin embargo, sólo cubre el período 1958-59.
- Cutright¹⁷. Emplea un sistema por intervalos de nivel acumulativos para comparar el desarrollo democrático de los sistemas políticos en 76 Estados, cubriendo el período 1940-60. Sin embargo, los Estados no han sido seleccionados siguiendo un patrón regular.
- Naubauer¹⁸. Sigue también un sistema por intervalos de nivel acumulativos, y también referido al período 1940-60, pero sólo cubre 23 Estados, siendo el criterio de selección el de aquellos países con un más alto nivel de desarrollo democrático.
- Smith¹⁹. Aunque cubre un espectro de 110 Estados, lo que lo aproxima a un nivel casi completo de generalidad en el objeto

¹⁶ COLEMAN, James S. *Conclusions: the political systems of the developing area*. En COLEMAN, Almond y J.S. *Education and political development*. Princeton University Press, 1960. pp. 532-581

¹⁷ CUTRIGHT, Phillips. *National political development: its measures and analysis*. *American Sociological Review*, n° 28 (Abril de 1963). pp. 253-264.

¹⁸ NEUBAUER, Deane E. *Some social conditions of democracy*. *American Political Science Review*. n° 61 (Diciembre de 1967). pp. 1002-1009.

¹⁹ SMITH, Arthur K (Jr). *Socioeconomic development and political democracy: a causal analysis*. *Midwest Journal of Political Science*. n° 13/1 (Febrero de 1969). pp. 95-125.

de la investigación, emplea tanto datos económicos como políticos para relacionarlos entre sí. Se trata de la hipótesis de la determinación económica de los derechos humanos que habremos de ver más adelante. Cubre el período 1945-65 siguiendo un método por intervalos de nivel acumulativos y temporales.

- Dahl²⁰. Emplea un sistema tendente a establecer una escala por orden acumulativo, que cubre 114 Estados, pero limitado al año 1969.
- Coulter²¹. Sigue un método de intervalos de nivel acumulativos, que sólo cubre 85 Estados seleccionados sin patrón uniforme, y abarcando los años 1950-70.
- Bollen²². Emplea una doble tabulación. La primera incluye 123 Estados sin criterio selectivo regular, y compara la situación de todos ellos alrededor del año 1965 en una tabla semiacumulativa. La segunda incluye 113 Estados y corresponde al año 1960, pero la estratificación de los datos es en intervalos de nivel.

El carácter común a estos estudios, y que los diferencia de estudios del tipo de los mencionados en las publicaciones periódicas, es que trataron de describir la situación general de los derechos humanos a nivel internacional, más que fijarse en violaciones localizadas en situaciones particulares en el espacio o en el tiempo, o particularizar el objeto de análisis por algún otro criterio no general.

Aun así, no todos tuvieron un nivel suficiente de amplitud. En general se trata de tentativas no demasiado ambiciosas de hacer una tabulación de Estados en función de su nivel de respeto por los derechos humanos, pero empleando medios insuficientes en la

²⁰ DAHL, Robert A. *Poliarchy, participation and opposition*. Yale University Press, New Haven, 1971.

²¹ COULTER, Philip. *Social mobilization and liberal democracy*. Lexington books. Lexington (Massachussets), 1975.

²² BOLLEN, Keneth. *Issues in the comparative measurement of political democracy*. American Sociological Review, nº 45 (Junio de 1980). pp. 370-390

mayor parte de los casos. Entre otras cosas, casi ninguno opera con su propio sistema de obtención de información, por lo que su fiabilidad depende de terceras fuentes de datos. Por otra parte, se plantearon empleando una metodología relativamente sencilla, por lo que hacen frecuentemente segmentaciones en el tiempo tan amplias que las diferencias en intervalos menores desaparecen, impidiendo hacer seguimientos más pormenorizados, o bien se refieren a períodos limitados que impiden también comparaciones fiables. Además, muchas de ellas seleccionan los Estados objeto de análisis en función de variables como los niveles de desarrollo, que impiden comparaciones con los Estados excluidos de tales categorías. Y, fundamentalmente, emplean definiciones operativas de las cuestiones objeto de análisis demasiado ambiguas. Como habrá de verse, este es un punto crucial de la metodología de un análisis sociológico de los derechos humanos basado información cuantitativa y técnicas estadísticas.

Dentro de los estudios monográficos cabría situar también otras tres fuentes de más difícil clasificación. Se trata de los estudios *Minorities at risk* y *Genocides and politicides since World War II* y la *Liste de descripteurs des droits de l'homme*. El segundo de ellos fue un proyecto de investigación truncado, que sólo dio lugar a un artículo y no continuó luego como estaba previsto. Su clasificación como monografía tiene que ser, por tanto, provisional. Por otra parte, ninguno responde a criterios de generalidad en el objeto en sentido estricto porque se refieren a aspectos bastante concretos de los derechos humanos, pero sí en el sentido de que cubren la totalidad de Estados y hacen seguimientos temporales suficientemente significativos. Al no cubrir el espectro completo de los derechos humanos no ofrecen una base óptima para un análisis comparativo, salvo en aquellos puntos de los que se ocupan. Pero se trata de estudios elaborados a partir de planteamientos metodológicos más rigurosos que los de muchos trabajos de los mencionados hasta ahora, y que tienen interés independientemente del resultado, como criterios eventualmente aprovechables para la definición de futuros sistemas de indicadores sobre derechos humanos. Y con el planteamiento metodológico no se trata

sólo de una cuestión formal, sino que el modelo empleado tiene implicaciones teóricas que habrán de discutirse más adelante. Por esta razón ambos estudios merecen una atención algo más completa. Los dos son fruto del trabajo de un equipo compuesto por Ted Robert Gurr, Barbara Harff y James R. Scarritt. El primero trata sobre minorías en riesgo²³, y el segundo sobre genocidios y politicidios²⁴. Cubren 126 Estados (con población de más de un millón de personas), y abarcan un período comprendido entre 1945 y 1989.

La idea de partida de estos trabajos es que, hasta el momento, todos los estudios cuantitativos sobre derechos humanos se han centrado bien en el status o victimación de individuos, bien en las políticas públicas o los caracteres estructurales de los Estados, con lo que se cubren los niveles micro y macrosociológico. Pero nadie se ha ocupado del nivel intermedio, esto es, el de los grupos políticos o comunitarios. A partir de esta observación se iniciaron dos proyectos, cada uno marcado a su vez por determinadas reflexiones acerca de tendencias habituales en la violación de los derechos humanos. En el caso de «Minorities at risk», la consideración fue que el enfoque habitual en el estudio de los derechos humanos consiste en plantear el desarrollo de la capacidad coercitiva del Estado como un medio para mantener el poder frente a los

²³ GURR, Ted R. y SCARRITT, James R. *Minorities at risk: a global survey*. En «Human Rights Quarterly». n.º 11/3 (Agosto de 1989). pp. 375-405. GURR, Ted y HARFF, Barbara. *Minorities at risk*. United States Institute of Peace. Washington, 1993.

²⁴ HARFF, Barbara y GURR, Ted R. *Victims of the state: genocides, politicides and group repression since 1945*. En «International Review of Victimology». n.º 1/1 /1989. pp. 23-41. Tanto para este como para el anterior, vid. también HARFF, Barbara y GURR, Ted R. *The rights of collectivities: principles, and procedures in measuring the human rights status of communal and political groups*. En JABINE, Thomas B. y CLAUDE, Richard P (eds). *Human Rights and Statistics. Getting the record stright*. University of Pennsylvania Press. Philadelphia, 1992. pp. 159-187. HARFF, Barbara y GURR, Ted. R. *Ethnic conflict in world politics*. Westriew Press. Boulder, 1994. GURR, Ted R. *People versus states: minorities at risk in the new century*. Unived States Institute of Peace. Washington, 2000.

potenciales conflictos intercomunitarios. Por tanto se trata de un enfoque «de arriba a abajo». La premisa de Gurr y Scarritt fue la inversa: ¿por qué se rebelan las minorías? Consecuentemente, el objeto de estudio no fueron los Estados, sino las minorías en situación potencial de amenaza de los derechos humanos. Partiendo de esta consideración se inició, en 1986, un inventario de todos los grupos que en el mundo se encontraban en riesgo de futuras violaciones de derechos fundamentales. En el caso de «Genocides and politicides since World War II» el planteamiento también supuso un cambio de rumbo frente al tipo de estudios hasta entonces dominantes. Estos, o bien se centraban en el holocausto del pueblo judío como único acontecimiento, o bien referían al análisis de textos legales internacionales para prevenir la repetición de genocidios. La cuestión que, por el contrario, llamó la atención de los autores no fue tanto la actitud de los gobiernos responsables de genocidios y politicidios, gobiernos que normalmente eran conscientes de practicar, a los ojos de la comunidad internacional, actos criminales. Más bien el hecho que necesitaba respuesta era por qué los individuos, alentados por sus gobiernos, emprendían el asesinato sistemático de otros ciudadanos sin una especial sensación de culpa, de manera semejante a los soldados, que reciben instrucción militar del Estado y matan en la guerra, convirtiéndose de nuevo en ciudadanos normales cuando regresan a la vida civil en tiempos de paz. El aspecto más problemático de la capacidad de personas «normales» para matar es no sólo la capacidad del Estado para entrenarlos, sino incluso para conseguir que tales acciones se presenten como justificadas ante ellos, hasta el punto de que los escuadrones de la muerte que practican genocidios se diferencian en poco del sentido del deber de los policías en su papel de garantizar el orden público o de los militares en el suyo de garantizar la integridad territorial del Estado.

A pesar de que los resultados obtenidos en esta investigación han dado lugar a publicaciones diferenciadas sobre genocidios o politicidios, evidentemente la metodología empleada tiene una base común en ambos casos, pues el equipo de investigación era el mismo. Se trataba, en primer lugar, de poder identificar grupos

diferenciados en situación de riesgo, y después de poder clasificar como politicidio o genocidio los casos de asesinatos masivos. Ejemplos de minorías en riesgo de potenciales violaciones de los derechos humanos son los kurdos de Irak, Irán o Turquía, los musulmanes de Filipinas o los Albaneses de Kosovo en Yugoslavia. Así, sobre una base común, se particularizan después criterios diferenciales específicos. Porque, ciertamente, hay casos en que el criterio clasificatorio se presta a dudas interpretativas. Por ejemplo, la minoría musulmana de los Chams en Camboya fue perseguida y masacrada por sus caracteres diferenciales como comunidad religiosa, pero en el contexto de la masacre emprendida por los Jemeres Rojos durante el régimen de terror de los años 70. Esta minoría era percibida por el gobierno como peligrosa por su potencial de oposición política, pero sin que tal identificación política pudiese ser netamente separada de la identificación religiosa del grupo. Por tanto, la tarea preliminar es la identificación de grupos minoritarios en riesgo, y sólo como un segundo paso la clasificación de la causa de la persecución explícita.

Por minorías se entenderán grupos cuyos miembros compartan una identidad colectiva, identidad que además estará basada en rasgos distintivos culturales, políticos, históricos o adscriptivos, y de modo que tales rasgos sean reconocidos como socialmente importantes no sólo por el propio grupo sino también por otros grupos o por la mayoría de la sociedad en que el grupo se integra. No todo grupo de estas características, sin embargo, se encuentra en situación de riesgo respecto a los derechos humanos, luego de esta primera definición se pasa a una segunda que delimita con mayor precisión el objeto. Para ello se establecen cuatro categorías de situaciones de riesgo:

1. Grupos sujetos a discriminación política. Tanto si encuentran limitados sus derechos políticos (libertad de voto, de asociación, etc.) como el acceso a posiciones de responsabilidad política. Pero para evaluar estas situaciones se emplea una categorización interna de medida, esto es, cuál es el status del grupo por comparación con otros grupos o con la generalidad de la ciudadanía del propio Estado, y no un patrón universal como los empleados

habitualmente en los sistemas de indicadores pensados para establecer un ranking internacional de derechos humanos.

2. Grupos sujetos a discriminación económica. Se entiende ésta como la exclusión sistemática del acceso a bienes, posiciones o condiciones económicas deseables que están abiertas a otros grupos del mismo Estado. Se emplea también una categorización independiente para cada Estado, y no una escala universal.

3. Grupos o regiones separatistas. Son grupos que reivindican una historia, una cultura y un idioma separados del Estado en que se integran y que han buscado reiteradamente un mayor grado de autonomía respecto a éste.

4. Minorías aventajadas. Paradójicamente, estos grupos están en situación de riesgo más por sus ventajas que por sus desventajas. Sin embargo, históricamente puede comprobarse que las situaciones de privilegio de un grupo terminan por condicionar a menudo movimientos sociales de rechazo que llevan a la persecución de los privilegiados. Ruanda ha sido un buen ejemplo reciente. Pero también las ventajas económicas de las minorías chinas en Malasia e Indonesia, inestablemente compensadas por las ventajas políticas de los grupos dominantes, representan situaciones de riesgo ciertas.

Con estas definiciones se delimitan conceptualmente los objetivos de la investigación. El siguiente paso es definir indicadores para hacer cuantitativamente operativos estos conceptos. Los autores los desglosan en 22 indicadores que evalúan desde la población y dispersión geográfica del grupo o su forma de estructuración política hasta el número de acciones antigubernamentales realizadas en cada década (manifestaciones, acciones armadas o campañas de insurrección). Con criterio realista, la dimensionalidad de cada indicador se preve en función de la información disponible, para que las comparaciones no resulten desfiguradas precisamente por la falta de datos. A su vez, se desiste de emplear escalas homogéneas (como escalas con intervalos de nivel, si bien en variables como la población, donde su empleo no es problemático, si son usadas), lo que redundaría en beneficio de

la exactitud estadística pero a costa de la fidelidad a la información disponible y de la conceptualización realizada. De modo que cada variable tiene su propio criterio de dimensionalidad, y no se prevé una escala unificada que sintetizase los resultados de todas las variables. En consecuencia, tanto en el caso de «Minorías en riesgo» como en el de «Genocidios y politocidios», no es posible hacer rankings de Estados, lo que en absoluto es una objeción, y más bien este tipo de tratamiento de la información cuantitativa resulta más objetivo.

El elemento sustancial en la definición conceptual tanto del politocidio como del genocidio es una política coherente, dirigida por un único grupo que ocupa el poder, y que resulta en la muerte de numerosos miembros de grupos perseguidos. A partir de la definición conceptual de las minorías en riesgo y de la definición de indicadores cuantitativos que hagan operativa esta conceptualización se trata, a continuación, de poder identificar cuándo se da un episodio de genocidio/politocidio. Los criterios para identificar adecuadamente estos casos implican: a) la muerte deliberada de no combatientes (para diferenciarlo de casos de guerra civil). b) un número suficientemente significativo de personas asesinadas (a partir de varios miles. Este criterio no presupone que en casos menos numerosos no se pueda hablar de politocidio/genocidio, pero los autores reconocen no disponer de información tan completa como para hacer un análisis suficientemente fiable. Por ejemplo, el caso «mínimo», respecto al genocidio, es el de los indios Ache de Paraguay, con menos de 1000 víctimas, pero no se considera comparable a casos como el de los Kurdos en Irak), y c) una campaña extendida en el espacio y el tiempo, y en este sentido se sigue un criterio temporal mínimo de seis meses.

Localizados los episodios con la suficiente precisión se puede aplicar el criterio clasificatorio que decide si se trata de casos de genocidio o de politocidio. En los genocidios se trata de eliminar grupos definidos principalmente por caracteres étnicos, religiosos, nacionales o raciales. Mientras en el politocidio las personas perseguidas lo son por su pertenencia a organizaciones políticas o su ideología. Los casos mixtos donde el criterio clasificatorio es

dudoso se mencionan explícitamente como ambivalente, pero en general el criterio definido puede resolver la inmensa mayoría de los casos. Por ejemplo, los kurdos son una minoría étnica y cultural. Pero los kurdos integrados en el sistema social, económico y político turco y que no apoyan la causa separatista no son objeto de discriminación por parte de las autoridades. Luego los episodios pueden ser clasificados claramente como de politicidio, dado que la causa de la persecución no es cultural, religiosa o lingüística.

La información reunida a partir de los indicadores definidos y la localización de los episodios significativos conforme a los criterios expuestos permiten ya el siguiente paso, que es la codificación cuantitativa de las 22 variables en que se desagregan los indicadores principales mencionados arriba. Se asignan 0 puntos cuando no hay diferencias sociales significativas entre la minoría y el resto de la población del Estado en cuestión, 1 punto cuando hay trato diferenciado en una o dos de las variables mencionadas, 2 puntos cuando incumbe a tres variables, y así sucesivamente hasta 4 puntos. Por ejemplo, en relación con las diferencias culturales (que se desglosan en cuestiones étnicas, lingüísticas, de origen histórico, religión, etc.), con un punto se clasifica la situación de la minoría católica holandesa frente a los protestantes, o los escoceses frente a los ingleses en Gran Bretaña. Con dos puntos los negros brasileños frente a los blancos. En relación a criterios diferenciales políticos (que incluyen el acceso a cargos públicos, reclutamiento militar, derecho de voto, etc.), se clasifica con cero puntos la situación de los francocanadienses frente a los anglocanadienses, mientras que merece cuatro puntos el caso de la población rural negra de Sudáfrica frente a los blancos, o el de la minoría Bahai en Irán. Lo mismo se aplica a tratos diferenciales económicos.

Por último, los autores añaden a sus estudios una serie de criterios finales acerca de la validez y fiabilidad de los mismos. En primer lugar, admiten carencias en cuanto a las fuentes de información. Incluso en cuestiones aparentemente más «neutras», como las demográficas, hay disparidad de criterios que a veces respon-

den ya a problemas discriminatorios. Por ejemplo, oficialmente no existen kurdos en Turquía, sino turcos de las montañas. Igualmente, tampoco existen en Bulgaria turcos, sino «búlgaros islamizados». En consecuencia, los respectivos gobiernos no se preocupan de elaborar censos fiables sobre la demografía de ambas minorías, que son incluidas en el conjunto de la población indiferenciadamente²⁵. Esto explica que la información sobre el tamaño de las minorías no esté disponible. Lo mismo ocurre con muchos países africanos, donde la identidad tribal es percibida como un vestigio del pasado y una amenaza a la identidad nacional emergente, por lo que tampoco se elaboran censos diferenciales de tales minorías tribales. En la misma línea, también es un obstáculo respecto a la fiabilidad el hecho de que hay mucha más información disponible sobre aquellos sucesos que han llamado la atención de la opinión pública internacional (como los desaparecidos de Argentina o el genocidio de Camboya), que otros sucesos que han pasado más desapercibidos (como la persecución de los nacionalistas ucranianos por parte de la Unión Soviética en 1947-48, o las campañas del gobierno pakistaní contra la minoría Baluchi en los años 60 y 70).

A pesar de que los autores reconocen ciertas lagunas en cuanto a la información disponible, creen poder garantizar la fiabilidad de los procedimientos de definición, categorización y cuantificación empleados. Y para avalarlo emplearon un procedimiento consistente en encomendar a diferentes investigadores que codificaran por separado un mismo cuerpo de información, y después comparar los resultados, que fueron satisfactorios. De todo lo expuesto se deduce que, si bien estos estudios no son tan generales por su objeto como para constituir indicadores de la situación de los

²⁵ El caso de Bosnia es una triste muestra de este mismo criterio, pues para los serbios nacionalistas la minoría musulmana no era más que la parte «traidora» de la población étnicamente serbia, por haber adoptado y mantenido la religión de los invasores turcos. Se les identificaba, por tanto, primariamente por su identidad étnica con la mayoría sin admitir su derecho a la libertad de conciencia para constituirse como grupo diferenciado.

derechos humanos, sino sólo de un aspecto de éstos, sin embargo la metodología empleada y los criterios de fiabilidad son ejemplares, y desde luego notablemente más exigentes que los de la mayoría de los estudios mencionados o de los que se verán a continuación.

Ambas investigaciones, *Minorities at risk* y *Genocides and politicides since World War II*, representan una forma muy apropiada para evaluar cuantitativamente el monopolio de la violencia legítima, que en la primera parte quedó definido como la primera condición sociológica para la consolidación de los derechos humanos.

El último de los estudios de esta sección, la *Liste de descripteurs des droits de l'homme*, no tiene un carácter directamente sociológico, sino más bien de informática jurídica, aunque al propio tiempo no deja de ser una investigación cuantitativa sobre los derechos humanos. De ahí que su clasificación tampoco resulte fácil. Se trata de un proyecto de investigación llevado a cabo por A.C. Kiss, J.P. Massué y K. Vasak en el Institut International des Droits de L'Homme de Estrasburgo. En realidad, la *Liste de descripteurs* forma parte de un proyecto más amplio, que consta de tres partes, siendo la tercera la Liste, que además es la única que tiene relevancia como estudio cuantitativo de los derechos humanos. La primera consiste en la elaboración de un glosario de los derechos humanos, planteado sobre la premisa de que esta materia exige una terminología precisa, y que los riesgos de la traducción en ocasiones provocan imprecisiones conceptuales. Para ello los autores se proponen redactar un glosario trilingüe (francés, inglés y alemán) con 1361 conceptos directamente relacionados con los derechos humanos. La segunda parte es una Enciclopedia de los Derechos Humanos, con carácter interdisciplinar y formulada sin prejuicios ideológicos o políticos. Los términos del Glosario servirían como material de trabajo básico para la elaboración de la lista de entradas de la Enciclopedia. La tercera parte es una lista de descriptores de derechos humanos susceptible de tratamiento informático.

Esta última parte del proyecto se ideó siguiendo el modelo de la lista de descriptores para la documentación relativa al desarrollo económico y social y que es común al BIT, al Comité Internacional para la Documentación en Ciencias Sociales, la FAO, la OCDE y la Deutsche Stiftung für Entwicklungsländer. De manera semejante, en previsión de que la documentación sobre derechos humanos tenderá a ser procesada informáticamente, puede realizarse una lista de descriptores de uso común para todos los organismos vinculados a ellos de una u otra forma (ONU, Consejo de Europa, OIT, UNESCO, Cruz Roja Internacional, OEA, etc.). Pero el listado no agota la tarea. El procesamiento informático del mismo fue diseñado como un medio para localizar la frecuencia de uso de los términos en los textos internacionales y nacionales sobre derechos humanos. Se hizo una búsqueda informática de los conceptos, predefinidos en el glosario, en textos normativos de jurisprudencia y de doctrina sobre derechos humanos, para comprobar con qué frecuencia eran empleados. Además, el programa localizó nuevos términos no incluidos en el glosario, y que se añadieron también al mismo.

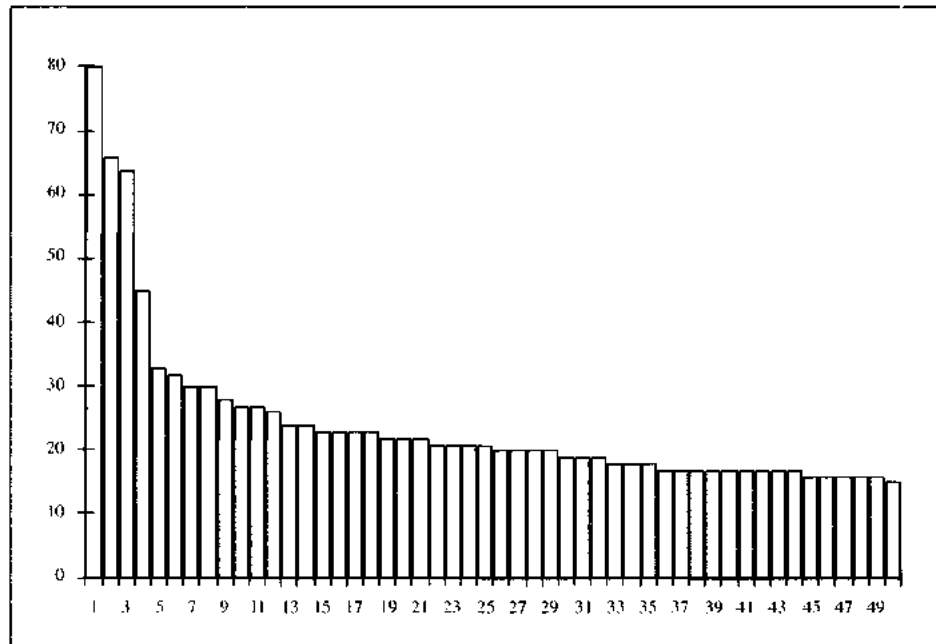
Los textos seleccionados para el análisis se distribuyeron en un 60% de textos internacionales y un 40% de los nacionales. A su vez, en el primer grupo se repartió el 60% en tres subgrupos (cada uno con un 20%), correspondientes a instrumentos jurídicos (tratados y declaraciones), jurisprudencia y doctrina. En los textos nacionales se hizo una división de un 10% de textos legislativos, un 15% de jurisprudencia y un 15% de doctrina. Sobre este material se llevó a cabo una búsqueda de 600 términos para comprobar la frecuencia de aparición. El resultado fue una lista de conceptos en orden decreciente de aparición, que se expone a continuación.

1. Ley
2. Convenio Europeo para la protección de los Derechos Humanos
3. Constitución
4. Estado
5. Disposiciones

6. Declaración Universal de los Derechos del Hombre
7. Comisión Europea de Derechos Humanos
8. Refugiado
9. Decisión
10. Persona
11. Protección
12. Derecho
13. Libertad
14. Estados
15. Religión
16. Demanda
17. Información
18. Comisión Interamericana de Derechos Humanos
19. Tribunal («Cour» en el original francés)
20. Gobierno
21. Vida privada
22. Sentencia
23. Aplicación
24. País
25. Secretario General del Consejo de Europa
26. Procedimiento
27. Tribunal Constitucional
28. Ciudadanos
29. Jueces
30. Competencia
31. Leyes
32. Provincias portuguesas de ultramar
33. Discriminación
34. Poderes públicos
35. Tribunales
36. Detención
37. Guerra
38. Orden público
39. Tribunal («Tribunal» en el original francés)
40. Restricciones
41. Tribunal Europeo de Derechos Humanos
42. Nacionalidad

43. Perjuicio
44. Gobierno demandado
45. Juez
46. Vejez
47. Derechos humanos
48. Ley Fundamental de la República Federal Alemana
49. Estados miembros de la OUA
50. Individuo

La frecuencia de aparición de cada concepto se expone en el siguiente gráfico.



No obstante, estos resultados responden a un primer ensayo tentativo, en que los autores trataron de comprobar la viabilidad de la investigación. Según los cálculos de fiabilidad estadística que hicieron, sería necesario analizar un mínimo de 50.000 términos para que el estudio tuviese garantías. Por esta misma razón, insisten en que estos primeros resultados no deben ser considerados como totalmente válidos para sacar de ellos conclusiones firmes. Sin embargo, el proyecto de investigación no continuó más

allá de esta primera tentativa, con lo que por el momento la obtención de datos más fiables se ha quedado en una posibilidad abierta. Aun así, los autores señalan entre sus conclusiones que es necesario definir y ampliar el glosario de términos para trazar los límites de una ciencia de los derechos humanos, pero que además

«Todavía conviene corregir esta conclusión un tanto desalentadora al constatar que los términos que tienen una frecuencia más alta (por ejemplo, «ley», «Estado», «decisión», etc.) no son propiamente específicos de la materia de los derechos humanos sino que componen la trama de toda rama jurídica de las ciencias humanas»²⁶.

En realidad, esta conclusión no resulta extraña vista desde el modelo weberiano que he expuesto en la primera parte, y según el cual los derechos humanos están intrínsecamente unidos al proceso de formalización y racionalización del derecho contemporáneo. Desde ese punto de vista no tiene nada de particular que tanto los instrumentos jurídicos, textos legislativos, la jurisprudencia y la doctrina vinculen la noción de los derechos humanos a conceptos como «ley», «decisión», «Estado», «tribunal» o «aplicación». Es más, si se tiene en cuenta que la jurisprudencia y la doctrina empleadas se ocupaban de demandas presentadas en la época, hay algunos conceptos que tienen una alta frecuencia de aparición por ser cuestiones planteadas en aquel momento, pero que lógicamente desaparecerían si se hiciese un estudio más prolongado en el tiempo. Es el caso, evidentemente, del concepto «Provincias portuguesas de ultramar», que se repite por haber sido hecho el estudio en un momento próximo a la descolonización. Hoy en día ese concepto no aparecería. Lo mismo cabe suponer en el caso del concepto «vejez». Por otra parte, el ordenador trata como conceptos distintos el singular y el plural de una misma palabra («Juge» y «Juges»; «Loi» y «Lois»), y también los sinónimos o próximos («Cour» y «Tribunal»). Si se eliminan los términos cuya alta frecuencia es puramente circunstancial y se agrupan los diferen-

²⁶ KISS, A.C., MASSUÉ, J.P. y VASAK, K. *L'ordinateur au service des droits de l'homme*. En *Revue des droits de l'homme*. Vol III, nº 3 (1970). p. 542.

ciados en plural y singular o los sinónimos, entonces la conclusión de los autores todavía se acentúa más. La suma de «ley» (que es el concepto más frecuente) con «leyes» ofrece un resultado en que la diferencia con otros conceptos aun se incrementa más. Incluso teniendo en cuenta que los datos son incompletos, los resultados no dejan de ser bastante significativos. Volveremos sobre este punto en el capítulo siguiente.

4.2.4. Fuentes con caracteres de periodicidad y generalidad en la información

En un tercer nivel podemos encontrar fuentes de información más operativas. Se trata de fuentes documentales que presentan características de generalidad (varias de ellas cubren la totalidad de países), periodicidad (lo que permite análisis comparativos en función de la variación temporal) y en algunos casos tratamiento sistemático de la información, si bien no son todavía plenamente aprovechables por algunas otras razones.

El primero de estos informes, y tal vez el más conocido, es el *Amnesty International Reports*, publicado por Amnistía Internacional con periodicidad anual. Cubre, además, prácticamente la totalidad de países. Sin embargo, no cumple exigencia alguna de tratamiento sistemático de la información. De hecho, la organización renuncia expresamente a tal tratamiento, y ante las insistentes solicitudes que le demandaban tal análisis hizo público un comunicado en su informe anual de 1987 en el que alegaba que...

«...el secreto gubernamental y la censura obstruyen el flujo de información desde muchos países e impiden los esfuerzos por verificar las alegaciones. Las comparaciones estadísticas u otras formas de generalización no pueden medir el impacto de los abusos de los derechos humanos sobre las víctimas (...). Las comparaciones de las prácticas gubernamentales sobre los derechos humanos pueden ser manipuladas y tergiversadas para fines políticos»²⁷.

²⁷ *Amnesty International Reports, 1987*. Amnesty International Publications. London. p. 2.

La honradez de este reconocimiento contrasta con otros informes que con bastante probabilidad tienen los mismos problemas de verificación de la información, pero no los hacen explícitos. Además de que Amnistía Internacional se autoimpone determinadas condiciones a la hora de obtener y documentar la información para garantizar su fiabilidad y el que no esté desfigurada por prejuicios subjetivos (como que los informes sobre un Estado sean realizados y verificados por personas distintas a los nacionales de tal Estado), mientras otras fuentes se basan en informaciones periódicas sin ulterior verificación. Estos criterios honran el procedimiento de Amnistía Internacional, que si no aporta una información completa es precisamente porque prefiere emplear únicamente aquella de la que puede ofrecer garantías de fiabilidad. Este tipo de cautelas son destacables, además, si se considera el fin primordial al que sirve la organización, que es el activismo en relación a los derechos humanos, y no un interés científico. Pero lo que es elogiado en este aspecto no obsta para reconocer, sin embargo, que la falta de tratamiento sistemático de la información impide emplear los informes de Amnistía Internacional como un indicador suficiente de la situación de los derechos humanos a nivel internacional, y lleva a clasificar sus informes en el grupo mencionado arriba, caracterizado por una presentación casuística de la información, poco susceptible de análisis generalizable. Harry M. Scoble y Laurie S. Wiseberg²⁸ han señalado, por ejemplo, que Amnistía Internacional está imposibilitada para obtener información cuando los gobiernos cierran completamente las fronteras, como ocurrió en Camboya entre 1975 y 1978 durante el régimen de los Jemeres Rojos. Evidentemente, que no hubiese datos disponibles no significa que no hubiese violaciones de los derechos humanos (y probablemente el caso de Camboya es uno de los episodios más negros de la historia en ese sentido).

²⁸ SCOBLE, Harry M. y WISEBERG, Laurie S. *Problems of comparative research on human rights*. En NANDA, Ved P., SCARRITT, James R. y Shepherd, George W. (Jr). *Global Human Rights: public policies, comparative measures and NGOs strategies*. Westview Press. Boulder (Colorado), 1981. p. 149

No obstante, hay otra objeción que puede hacerse a la información suministrada por Amnistía Internacional. La organización tiene un interés preferente por los abusos de derechos humanos relativos a los presos de conciencia y la tortura, pero no presta la misma atención a otros derechos humanos, como los sociales y económicos. Aparte del hecho de que muchos gobiernos optan por la ejecución extrajudicial o el exilio de los disidentes antes que la prisión, y en tal caso este Estado no aparecerá mencionado en los informes sobre presos de conciencia, lo que no significa que no se den violaciones de los derechos humanos. Esta segunda limitación afecta al criterio de generalidad, no tanto en su sentido geográfico sino en relación con el objeto. Por estas dos razones alegadas, esta fuente de información debe quedar en un plano secundario frente a otras más completas y ambiciosas. No obstante, estas limitaciones no tienen en ningún caso el sentido de una crítica a la organización.

Una segunda fuente de datos es *Human Rights in Developing Countries*. Se trata de una publicación anual editada por los Centros para los derechos humanos de Noruega, Dinamarca, Holanda, Finlandia y Canadá. El propósito de esta obra es orientar la adjudicación de fondos de ayuda al desarrollo en función del respeto a los derechos humanos en los países receptores, siendo los donantes los países arriba mencionados. Sin embargo, esto limita el ámbito de análisis a trece países: Bangladesh, Botswana, India, Kenia, Mozambique, Nicaragua, Pakistán, Filipinas, Surinam, Tanzania, Zambia y Zimbabwe, que son los principales receptores de la ayuda externa de los Estados antes mencionados. A pesar de que la información está expuesta en estilo narrativo, incluye suficiente información cuantitativa, relativa a la posición del gobierno sobre los derechos humanos, el sistema político y los derechos de participación, derechos civiles (vida, libertad e integridad de las personas y libertad de tránsito y residencia) y administración de justicia. El interés de este anuario está en los fines prácticos derivados de estudios basados en indicadores cuantitativos, una cuestión que se plantea en general como una de las posibilidades de un eventual sistema de evaluación de los

derechos humanos cuya metodología haya merecido suficiente consenso. Sin embargo, esta posibilidad será tanto más fiable cuanto mayor sea el número de partícipes entre los que se alcance dicho consenso, y este sentido el presente indicador está limitado a los cinco países que han contribuido a su elaboración. Esta misma razón podría ser un indicio de parcialidad, si la metodología empleada estuviese condicionada por la política exterior de los Estados partícipes²⁹. Sin embargo, este indicador no es particularmente cuestionable por razones metodológicas. En el capítulo 5 volveremos sobre él y tendremos ocasión de comprobarlo. Aparte de esto, hay otra objeción más clara en el ámbito tan restringido de análisis, lo que limita a este anuario a un interés menor.

En la misma línea que el anterior, el *Country Reports on Human Rights Practices* es una publicación del Departamento de Estado de Estados Unidos, que también nació con la intención de vincular la ayuda externa al nivel de respeto de los derechos humanos de los Estados receptores. Se trata de un anuario que se publica desde 1976. Incluye información sobre respeto a los derechos humanos y sobre la situación económica, social y cultural de cada Estado. Sin embargo, no hay un patrón sistemático para la evaluación de los derechos humanos, de forma que aspectos analizados en un país no son ni siquiera mencionados en otros. En todo caso, contrasta más bien la situación de los países en relación con el Bill of Rights norteamericano que con la Declaración Universal de los Derechos del Hombre de la ONU. Ha sido bastante criticado por su cuestionable objetividad. Por otra parte, no analiza la situación de los propios Estados Unidos, pues

²⁹ En general, no sólo en el caso de este estudio, sino también en relación con los programas de ayuda al desarrollo de la ONU, la Unión Europea y los Estados Unidos, se reconoce que el objetivo de tales programas debe ser el desarrollo de los derechos humanos. Sin embargo, la ayuda se envía también a países que no los respetan, y aun cuando otros los respeten en parte, la ayuda no se orienta específicamente a programas previstos para la mejora de la situación de tales derechos. Sobre este punto, Cfr. TOMASEVSKI, Katarina. *Development aid and human rights*. St. Martin's Press. New York, 1989.

evidentemente no puede ser receptor de su propia ayuda externa. Cumple con el requisito de generalidad geográfica de un modo bastante amplio, pero no de generalidad en cuanto al objeto, y tampoco ofrece un análisis sistemático de la información. El hecho de supeditar la objetividad de la información a la política exterior norteamericana es más evidente aún que en el caso anterior. Cada año, tras la publicación de este informe, el Lawyers Committee for Human Rights de Estados Unidos publica un «contrainforme», llamado *Critique*, en el que revisa las informaciones allí contenidas para poner de manifiesto en qué medida responden a criterios políticos más que a una pretensión de objetividad. En este sentido, los autores de la edición de *Critique* de 1992 afirman que:

«Cada año, la introducción al *Country Reports* repite el señuelo de que el informe es «remitido al Congreso por el Departamento de Estado en cumplimiento de las secciones 116 (d) y 502 (b) de la Foreign Assistance Act de 1971 y sus enmiendas». La sección 502 b prohíbe la ayuda externa a países implicados en un «patrón consistente de graves violaciones de los derechos humanos internacionalmente reconocidos». Sin embargo, y a pesar de que la sección 502 b fue aprobada hace 14 años, nunca ha sido formalmente invocada, pese a la evidencia de graves violaciones de los derechos humanos documentadas en muchos de los informes del Departamento de Estado»³⁰.

Un cuarto informe de interés es el *Human Suffering Index*, publicado por la organización Population Crisis Committee (Washington). Cubre un espectro de 130 países, y se publica cada 5 años. Intenta presentar un escala de las diferencias en las condiciones de vida en el mundo. Para ello considera 10 variables: 1) PIB, 2) Tasa de inflación, 3) Tasa de productividad, 4) Crecimiento de la población urbana, 5) Mortalidad infantil, 6) Consumo de calorías en la dieta alimenticia, 7) Porcentaje de población con acceso a agua corriente, 8) Consumo de energía per capita, 9) Tasa de alfabetización y 10) Libertad personal. Sin embargo, esta última

³⁰ Lawyers Committee for Human Rights. *Critique. Review of the U.S. Department of State's Country Reports on Human Rights Practices for 1991*. New York, 1992. p. 6.

variable está tomada de otro informe (de Freedom House, que veremos más adelante). Por esta razón, no es un indicador directamente referido a los derechos humanos, pero es citado por otras fuentes frecuentemente.

Un caso semejante es el del *World Handbook of Political and Social Indicators*, editado por la Virginia Polytechnic and State University. Se han publicado tres volúmenes (1948-67, 1967-77 y 1977-83). Su cobertura es mundial. Intenta un estudio comparado de los países en relación con determinados índices políticamente relevantes. Entre los indicadores más relacionados con los derechos humanos incluye libertad de prensa, regularidad electoral, manifestaciones de protesta, muertes por causas políticas y sanciones gubernamentales contra la oposición. Pero hay otras 57 variables más en el sistema de evaluación. Debe señalarse que buena parte de los datos que aporta provienen de informaciones periódicas, con la limitación de fiabilidad que ello supone. Como en el caso anterior, su interés directo en relación con los derechos humanos es secundario, pero sus datos son empleados por otros informes que sí están dirigidos específicamente a esta cuestión.

Las siguientes fuentes son las más representativas de intentos de información cuantitativa sobre derechos humanos, cumpliendo las características de generalidad geográfica y del objeto de análisis, periodicidad y tratamiento sistemático de la información. Por ello merecerán una atención más detallada.

4.3. LOS INFORMES DE FREEDOM HOUSE Y CHARLES HUMANA

The Annual Survey of Political Rights and Civil Liberties, editado por la organización Freedom House de Nueva York/Washington es un informe anual publicado ininterrumpidamente desde 1973, que cubre todos los Estados independientes del mundo y los territorios dependientes. Incluye informes particulares sobre la situación de los derechos civiles y políticos de cada Estado, a lo

que añade una breve historia política del país, y aporta además información estadística sobre demografía, PIB, grupos étnicos, expectativa de vida, etc. No obstante, la mayor parte de la información se expone en estilo narrativo, lo que representa una cierta dificultad para establecer conclusiones precisas. En otro orden de cosas, los estudios de Freedom House están diseñados para evaluar la situación de los derechos civiles y políticos en todos los Estados a lo largo del mundo, es decir, que no declara entre sus propósitos hacer una evaluación de los derechos humanos como tales, si bien luego emplea frecuentemente a éstos como sinónimos de derechos civiles y políticos. Esta ambivalencia no es casual. Al igual que en otros casos ya mencionados (como el Country Report on Human Rights Practices), el patrón de medida empleado por Freedom House se corresponde mucho más con el Bill of Rights de Estados Unidos que con la Declaración Universal de los Derechos Humanos o el Pacto Internacionales de Derechos Civiles y Políticos, y la diferencia tal vez se acentúa más respecto del Pacto Internacional de Derechos Económicos, Sociales y Culturales. Ciertamente esta definición de objetivos del estudio es síntoma de una clara orientación ideológica. Así, por ejemplo, en la edición del año 1994-95, en un apartado titulado «Oportunidades para la expansión de la libertad», se dice respecto de las democracias incipientes que...

«...estos países representan la mejor esperanza para el ensanchamiento de la libertad humana. Las democracias frágiles y emergentes merecen ser el punto de atención central de los esfuerzos de la ayuda externa estadounidense. Tal asistencia debería estar más rigurosamente dirigida a la promoción de organizaciones no gubernamentales, al fortalecimiento del sector privado y los sindicatos libres, a promover el diálogo y la cooperación interétnicos, a fortalecer la promoción de la sociedad civil, a asistir a la diversificación de los medios de comunicación y a consolidar las bases de un sistema judicial independiente e imparcial»³¹.

Del texto se deduce claramente que Freedom House es una organización políticamente activa e interesada en influir en la

³¹ VV.AA. *Freedom in the world. The annual survey of political rights and civil liberties. 1994-95*. Freedom House. New York, 1995. p. 8

política exterior de Estados Unidos, que además se orienta por ideas liberales en cuanto a la prevalencia de la sociedad civil frente al Estado, a la promoción de un modelo económico de libre mercado, etc³². Más adelante habrá ocasión de considerar con mayor detenimiento en qué medida esta orientación ideológica resta fiabilidad al informe, punto este en el que coinciden la mayor parte de los autores que se han ocupado de estudiar y comparar la metodología de los diversos sistemas de evaluación cuantitativa de la eficacia de las normas internacionales sobre los derechos humanos.

La metodología empleada por Freedom House consiste en una doble tabla de variables, la primera para evaluar la situación de los derechos civiles y la segunda referida a los derechos políticos. Por cierto que la propia definición de qué se entiende por ambos conceptos es también muy representativa de la orientación de esta institución. Derechos civiles significan para Freedom House las «libertades para desarrollar puntos de vista, instituciones y autonomía personal al margen del Estado»³³. Los derechos políticos son entendidos como aquellos que «posibilitan a la población a participar libremente en los procesos políticos»³⁴. Cada una de las tablas correspondientes se compone de un cuestionario diseñado para medir el nivel de respeto por los derechos civiles y políticos. Los cuestionarios son los siguientes:

CUESTIONARIO DE DERECHOS POLÍTICOS

1. ¿Es elegido el jefe del Estado y/o jefe del gobierno por medio de elecciones libres y no manipuladas?
2. ¿Son elegidos los representantes legislativos por medio de elecciones libres y no manipuladas?

³² Freedom House ha sido calificada por algunos críticos como un baluarte del pensamiento conservador estadounidense. Cfr., por ejemplo, NAGLE, John D. *Introduction to comparative politics*. Nelson Hall, Chicago, 1985. p. 95.

³³ Idem. p. 672

³⁴ Idem. p. 672

3. ¿Hay leyes electorales equitativas, igualdad de oportunidades en la campaña electoral y votaciones y publicación de resultados libres de sospecha?
4. ¿Están capacitados los electores para conferir a sus representantes libremente elegidos un poder real?
5. ¿Tienen los ciudadanos el derecho a organizarse en partidos políticos u otros grupos políticos competitivos de su elección? ¿Está abierto el sistema electoral al ascenso y caída de tales grupos o partidos?
6. ¿Se da un porcentaje significativo de voto a las opciones de la oposición? ¿Tiene la oposición un poder real? ¿Existe la posibilidad realista para la oposición de incrementar su apoyo o de conquistar el poder a través de las elecciones?
7. ¿Tiene la nación derecho de autodeterminación? ¿Son libres los ciudadanos respecto del dominio de los militares, poderes extranjeros, partidos totalitarios, jerarquías religiosas, oligarquías económicas o cualquier otro grupo poderoso?
8. ¿Tienen las minorías culturales, étnicas, religiosas u otros grupos minoritarios un razonable derecho de autodeterminación, de autogobierno, autonomía o participación en la toma de decisiones a través de un consenso informal?
9. ¿Está descentralizado el poder político, permitiendo que las administraciones locales, regionales y/o provinciales sean dirigidas por representantes libremente elegidos? (Este punto no es considerado en caso de Estados de extensión limitada, como ocurre por ejemplo en algunas islas-Estado muy pequeñas)

CUESTIONARIO DE DERECHOS CIVILES

1. ¿Hay medios de comunicación, literatura u otras expresiones culturales libres e independientes? (En Estados con medios de comunicación exclusivamente públicos, se evalúa el nivel de pluralismo permitido en dichos medios)
2. ¿Hay discusión pública abierta y discusión privada libre?
3. ¿Hay libertad de asociación y manifestación?

4. ¿Hay libertad de asociación política o quasi-política? (incluye partidos políticos, asociaciones civiles y grupos ad hoc)
5. ¿Son los ciudadanos iguales ante la ley, con acceso a un poder judicial independiente y no discriminatorio, y son respetados por las fuerzas de seguridad?
6. ¿Existe protección frente al terrorismo político, la prisión injustificada, el exilio o la tortura, tanto por grupos que apoyen al sistema como que se opongan a él, y libertad frente a la insurgencia o la guerra?
7. ¿Hay sindicatos libres y organizaciones de agricultores o equivalentes, así como una eficaz negociación colectiva?
8. ¿Hay organizaciones profesionales libres y otras organizaciones privadas?
9. ¿Hay negocios libres y cooperativas?
10. ¿Hay instituciones religiosas libres y libertad de expresión religiosa pública o privada?
11. ¿Hay libertades sociales personales, como igualdad entre los sexos, derechos de propiedad, libertad de movimientos, elección del lugar de residencia y libertad para contraer matrimonio y elegir el tamaño de la familia?
12. ¿Hay igualdad de oportunidades, incluyendo la libertad para no ser explotado o depender de terratenientes, empleadores, sindicalistas, burócratas o cualquier otro tipo de obstáculo para compartir los beneficios económicos legítimos?
13. ¿Existe libertad frente a la indiferencia gubernamental extrema o la corrupción?³⁵.

Sobre cada uno de los elementos de este cuestionario, y en función de la situación del Estado de que se trate, se asignan puntos en una escala de 1 a 7, representando 7 el mínimo de derechos civiles y políticos y 1 el máximo. A continuación se agregan los resultados obtenidos en cada apartado de cada una de las tablas, y se calcula la media aritmética. El informe incluye tanto un listado

³⁵ Cfr. VV.AA. *Freedom in the world. The annual survey of political rights and civil liberties, 1994-95*. pp. 672-74

de Estados con sus correspondientes puntuaciones en la escala de derechos civiles como en la de derechos políticos, además de una escala unificada de ambos que resulta de la media entre los dos. Aparte de esto, la lista total de países se divide, en función de su posición, en tres grupos: Estados libres, parcialmente libres y no libres, comprendiendo la primera aquellos que han obtenido una puntuación entre 1 y 2,5, la segunda quienes cuentan entre 3 y 5, y la tercera los que obtienen entre 5,5 y 7. A continuación se expone la lista de Estados en la escala unificada, es decir, la que sintetiza la situación de los derechos civiles y los políticos. No obstante, Freedom House incluye en su listado países de un tamaño mínimo (como San Vicente y las Granadinas, Tuvalu, San Kitts-Nevis, etc.), que además de ser muy poco representativos incrementan notablemente la relación, por lo que la mayor parte de estos casos no serán mencionados.

Estados libres		Parcialmente libres		No libres	
Andorra	1.0	Bahamas	1.5	Polonia	2.0
Australia	1.0	Cabo Verde	1.5	Uruguay	2.0
Austria	1.0	Costa Rica	1.5	Samoa Occid	2.0
Barbados	1.0	Dominica	1.5		
Bélgica	1.0	Eslovenia	1.5	Argentina	2.5
Belice	1.0	España	1.5	Benin	2.5
Canadá	1.0	Francia	1.5	Bolivia	2.5
Chipre	1.0	Granada	1.5	Botswana	2.5
Dinamarca	1.0	Hungría	1.5	Ecuador	2.5
EE.UU	1.0	Irlanda	1.5	Estonia	2.5
Finlandia	1.0	Is. Mauricio	1.5	Eslovaquia	2.5
Holanda	1.0	Italia	1.5	Jamaica	2.5
Islandia	1.0	Mónaco	1.5	Letonia	2.5
Liechtenstein	1.0	Reino Unido	1.5	Malawi	2.5
Luxemburgo	1.0	Rep. Checa	1.5	Mongolia	2.5
Malta	1.0	Trinidad T.	1.5	Namibia	2.5
Micronesia	1.0			Panamá	2.5
Noruega	1.0	Bulgaria	2.0	Sudáfrica	2.5
N. Zelanda	1.0	Chile	2.0		
Portugal	1.0	Corea del Sur	2.0	Bangladesh	3.0
San Marino	1.0	Grecia	2.0	Brasil	3.0
Suecia	1.0	Guayana	2.0	El Salvador	3.0
Suiza	1.0	Israel	2.0	Honduras	3.0
		Japón	2.0	Madagascar	3.0
Alemania	1.5	Lituania	2.0	Mali	3.0

Papua N. Guinea	3.0	Ghana	4.5	Oman	6.0
Surinam	3.0	Guatemala	4.5	Tanzania	6.0
Taiwan	3.0	Malasia	4.5	Yugoslavia	6.0
Venezuela	3.0	Nicaragua	4.5		
Túnez	5.5	Perú	4.5	Brunei	6.5
Albania	3.5	Senegal	4.5	Burundi	6.5
Armenia	3.5	Sri Lanka	4.5	Gambia	6.5
Colombia	3.5			Indonesia	6.5
Filipinas	3.5	Georgia	5.0	Irán	6.5
Guinea-Bissau	3.5	Haití	5.0	Laos	6.5
Islas Fidji	3.5	Kuwait	5.0	Liberia	6.5
Kirguicia	3.5	Marruecos	5.0	Nigeria	6.5
Macedonia	3.5	Singapur	5.0	Qatar	6.5
Nepal	3.5	Turquía	5.0	Sierra Leona	6.5
Paraguay	3.5	Uganda	5.0	Zaire	6.5
Rep. Centroatricana	3.5	Zimbabwe	5.0		
Rep. Dominicana	3.5			Afganistán	7.0
Rumanía	3.5	Líbano	5.5	Angola	7.0
Rusia	3.5	Camerún	5.5	Arabia Saudí	7.0
Ucrania	3.5	Chad	5.5	Argelia	7.0
Zambia	3.5	Costa Marfil	5.5	Bhutan	7.0
		Emiratos Ar.	5.5	Birmania	7.0
Bielorrusia	4.0	Eritrea	5.5	China	7.0
Congo	4.0	Etiopía	5.5	Corea del N.	7.0
Croacia	4.0	Guinea	5.5	Cuba	7.0
India	4.0	Kazajstán	5.5	Guinea Ec.	7.0
Jordania	4.0	Swazilandia	5.5	Iraq	7.0
Lesotho	4.0	Togo	5.5	Libia	7.0
México	4.0	Yemen	5.5	Mauritania	7.0
Moldavia	4.0			Ruanda	7.0
Mozambique	4.0	Azerbaiyan	6.0	Siria	7.0
Niger	4.0	Bahrain	6.0	Somalia	7.0
Pakistán	4.0	Bosnia Herz.	6.0	Sudán	7.0
Tailandia	4.0	Djibuti	6.0	Tayikistán	7.0
		Egipto	6.0	Turkmenistán	7.0
Burkina Faso	4.5	Kenia	6.0	Uzbekistán	7.0
Camboya	4.5	Maldivas	6.0	Vietnam	7.0
Gabon	4.5				

El *World Human Rights Guide* es un informe periódico elaborado por Charles Humana, un investigador académico cuyo trabajo tiene a su favor que no es el encargo de alguna institución o asociación de activistas, lo que contaría como indicio de sesgo ideológico y que, como se ha visto, resta fiabilidad a otros trabajos.

No obstante, este simple dato no garantiza que el resultado del informe este libre de toda crítica, y de hecho se le han hecho objeciones que habrán de examinarse más adelante. Este informe se publicó por primera vez en 1983, con el propósito de editar siguiendo una periodicidad de 3/4 años nuevos informes actualizados. Una segunda edición apareció en 1986 y la tercera en 1991. Sin embargo, lo más destacable es que en 1991, el Programa de Naciones Unidas sobre Desarrollo incorporó la Guía de Humana correspondiente a 1986 en el Human Development Report, un documento anual de Naciones Unidas que evalúa un amplio conjunto de indicadores estadísticos que se sintetizan en el denominado Índice de Desarrollo Humano.

Humana denomina a cada uno de los elementos del cuestionario con una abreviatura (que aquí se reproduce en su versión original inglesa por la imposibilidad de traducir una abreviatura). Inmediatamente después de esta abreviatura se menciona el contenido de la cuestión concreta, y a continuación el artículo de las Declaraciones o Pactos internacionales a que corresponde esta cuestión. Esto no quiere decir que cada elemento del cuestionario sea siempre idéntico al contenido del artículo con el que se corresponde. Hay casos en que un mismo artículo sirve de base para dos elementos diferenciados del cuestionario (así, por ejemplo, el artículo 23.2 de la Declaración Universal de los Derechos Humanos se desglosa en dos elementos del cuestionario: SOCFEM y ETHEQ, el primero de los cuales evalúa el derecho de las mujeres a la igualdad social y económica, y el segundo el derecho de las minorías étnicas a la misma igualdad social y económica). En otros casos, el contenido de un elemento del cuestionario no agota el contenido del artículo correspondiente. Por ejemplo, el ítem BRTHCONT evalúa el derecho a emplear píldoras y otros métodos anticonceptivos, y se basa en el artículo 15.1b del Pacto Internacional de Derechos Económicos, Sociales y Culturales que reconoce el derecho a gozar de los beneficios del progreso científico. Obviamente, este artículo incluye mucho más que el simple derecho al uso de métodos anticonceptivos, pero es esto último lo que Humana se propone examinar. En relación con las Declaracio-

nes o Pactos se emplearán las siguientes abreviaturas: DUDH (Declaración Universal de los Derechos del Hombre de 1948), PIDCP (Pacto Internacional de Derechos Civiles y Políticos de 1966) y PIDESC (Pacto Internacional de Derechos Económicos, Sociales y Culturales de 1966). El orden de los elementos en la tabla asignado por Humana tal como se reproduce a continuación no presupone una jerarquización.

1. **MOVEIN.** Derecho a circular libremente y a elegir residencia en el territorio de un Estado (art. 13.1 de la DUDH)
2. **MOVEOUT.** Derecho a salir de cualquier país, incluido el propio, y a regresar a su país (art. 13.2 de la DUDH)
3. **ASSEMBLY.** Derecho de reunión pacífica (sin otras restricciones que las impuestas por la ley para garantizar la seguridad nacional, la seguridad y el orden público, la salud o la moral pública y los derechos y libertades de los demás) y libertad de asociación. (art. 21 del PIDCP)
4. **FREEINFO.** Derecho a investigar y a recibir información y opiniones y a difundirlas. (art. 19 DUDH)
5. **MONITOR.** Libertad para controlar e informar sobre violaciones de los derechos humanos. (art. 19 DUDH)
6. **ETHLANG.** Derecho de las minorías étnicas, religiosas o lingüísticas a emplear su propio idioma, y a recibir educación o educar a sus hijos y publicar también en su propio idioma. (art. 27 PIDCP)
7. **SLAVLABR.** Libertad de la servidumbre, la esclavitud y del trabajo forzado infantil. Obligación de los Estados de fijar límites de edad por debajo de los cuales quede sancionado y prohibido por ley el empleo a sueldo de mano de obra infantil (art. 10.3 del PIDESC)
8. **MURDER.** Evalúa el número de ejecuciones extrajudiciales y «desapariciones». (art 6.1 del PIDCP)
9. **TORTURE.** Derecho a no ser sometido a torturas ni tratos crueles, inhumanos o degradantes. (art. 5 de la DUDH)
10. **FREEWORK.** Derecho al trabajo, a la libre elección de profesión, a condiciones equitativas y satisfactorias de trabajo

y a la protección contra el desempleo. Prohibición de trabajos forzados. (art. 23.1 de la DUDH)

11. CAPPUN. Prevalencia de la pena capital por parte del Estado. Si bien la pena de muerte no está excluida en la DUDH, en esta cuestión se evalúa la proporción en que se emplea y las medidas que se adoptan para su abolición. (art. 6.6 PIDCP y Segundo Protocolo Opcional al PIDCP)
12. PUNISH. Derecho a no recibir penas corporales aplicadas por el Estado. (art. 7 del PIDCP)
13. DETENTN. Derecho a no ser detenido por las fuerzas de orden público sin cargos. (art. 9.1 del PIDCP)
14. MEMBER. Libertad de toda afiliación obligatoria a partidos estatales. (art. 20.2 de la DUDH)
15. NOIDEOL. Nadie será obligado a recibir una educación basada en la religión oficial del Estado o en la ideología estatal oficial. (art. 18.1 del PIDCP)
16. FREEART. Libertad de la producción artística del control o la censura estatal. (art. 19 DUDH)
17. FREEPRESS. Libertad de toda censura política de la prensa. (art. 19 DUDH)
18. FREEMAIL. Derecho a comunicaciones postales libres de censura. (art. 12 de la DUDH)
19. POLTCOPP. Derecho a la oposición política pacífica. Ausencia de controles y persecuciones de los grupos opositores. (art. 25 del PIDCP)
20. BALLOT. Derecho de todos los ciudadanos a votar y ser elegido en elecciones periódicas, multipartidistas, realizadas por sufragio universal y voto secreto. (art. 25 PIDCP)
21. LEGFEM. Derecho de las mujeres a la igualdad legal y política. (art. 2 de la DUDH)
22. SOCFEM. Derecho de las mujeres a la igualdad social y económica. (art. 23.2 de la DUDH)
23. ETHEQ. Derecho de las minorías étnicas a la igualdad social y legal. (art. 2 de la DUDH)
24. PAPER. Libertad para la publicación de periódicos independientes. (art. 19 de la DUDH)

25. **BOOK.** Libertad para la publicación de libros. (art. 19 de la DUDH)
26. **TVRADIO.** Libertad de las emisoras de radio y televisión del control o la censura estatal. (art. 19 de la DUDH)
27. **INDCOURT.** Derecho a la independencia de los tribunales que hayan de examinar los derechos y obligaciones de las personas, o que examinen las acusaciones penales que se formulen contra ellas. (art. 10 de la DUDH)
28. **UNION.** Derecho a fundar y afiliarse libremente a sindicatos, que puedan operar sin interferencias del poder político. (art. 8 del PIDESC)
29. **KEEPCIT.** Derecho a tener una nacionalidad, y a no ser privado arbitrariamente de ella o del derecho a cambiar de nacionalidad. (art. 15 de la DUDH)
30. **PRVGUILT.** Derecho de los acusados de algún delito a la presunción de inocencia mientras no sea probada la culpabilidad conforme a la ley. (art. 11.1 de la DUDH)
31. **LEGALAID.** Derecho de los acusados de algún delito a recibir asistencia legal gratuita si carece de medios para pagarla. (art. 14.3 del PIDCP)
32. **PUBTRIAL.** Derecho a un proceso público cuando un tribunal haya de examinar los derechos y obligaciones de las personas, o cuando examine las acusaciones penales que se formulen contra ellas. (art. 10 de la DUDH)
33. **QKTRIAL.** Derecho de los detenidos a causa de una infracción penal a ser llevados sin demora ante un juez o tribunal, y a ser juzgados en un plazo razonable. (art. 9.3 del PIDCP)
34. **NOSEARCH.** Derecho a no ser objeto de injerencias arbitrarias en el domicilio. Prohibición de registros policiales del domicilio sin garantías judiciales. (art. 17.1 del PIDCP)
35. **PROPERTY.** Derecho a la propiedad, individual y colectivamente, y a no ser privado arbitrariamente de ella. (art. 17 de la DUDH)
36. **MIXMARRY.** Derecho a contraer matrimonio interracial, interconfesional y al matrimonio civil. (art. 16.1 de la DUDH)

37. **EQMARRY.** Igualdad de derechos entre varones y mujeres para contraer matrimonio, durante el matrimonio y en caso de disolución del matrimonio. (art. 16.1 de la DUDH)
38. **ANYRELGN.** Derecho a practicar cualquier religión, incluyendo la libertad para cambiar de religión o de creencia, la libertad para manifestar su religión o creencia, individual o colectivamente, en público o en privado, por la enseñanza, la práctica, el culto y la observancia (art. 18 de la DUDH)
39. **BIRTHCONT.** Derecho a emplear píldoras y otros métodos anticonceptivos. (art 15.1b del PIDESC)
40. **HOMOSEX.** Derecho a practicar la homosexualidad entre adultos que consienten. (art. 12 de la DUDH)

Cada uno de estos elementos es examinado en los Estados objeto de estudio, y en función de la situación se asigna un valor en una tabla de 1 a 4, de modo que, por ejemplo, respecto de LEGFEM el Estado donde no exista discriminación alguna hacia las mujeres y se de una absoluta igualdad legal y política recibiría cuatro puntos en este apartado. Las puntuaciones obtenidas se van agregando para obtener la puntuación total correspondiente. No obstante, Humana considera que ciñiendo el cuestionario a los textos legales internacionales no quedan márgenes para una interpretación interesadamente desviada, pero no pueden valorarse igual cada uno de los cuarenta elementos resultantes, algunos de los cuales son claramente más importantes que otros para valorar la situación de un Estado en relación a su respeto de los derechos humanos. De este modo, Humana selecciona siete de los cuarenta elementos y les atribuye un valor especial. Estos siete elementos son: TORTURE, PUNISH, CAPPUN, MURDER, DETENTN, FREEWORK y SLAVLABR. Humana les asignó un valor de ponderación triple al del resto de elementos del cuestionario, lo que en su opinión no tergiversaba los resultados finales más allá de un 3-4%, ni alteraba gravemente el orden final de los países.

A continuación se expone la tabla de Estados en función de los puntos obtenidos siguiendo los criterios arriba expuestos para que pueda comprobarse el grado de correspondencia general con los resultados del informe de Freedom House. Los datos son los

correspondientes al informe de 1991, que es el que Naciones Unidas asumió como propio en su Informe sobre el Desarrollo Humano:

Finlandia	99	Hong Kong	79	El Salvador	53
Alemania	98	R. Dominicana	78	Ghana	53
Dinamarca	98	Israel	76	Mozambique	53
Holanda	98	Costa de Marfil	75	Egipto	50
N. Zelanda	98	Nicaragua	75	Sudáfrica	50
Suecia	98	Venezuela	75	Nigeria	49
Checoslov.	97	Filipinas	72	Oman	49
Hungría	97	Jamaica	72	Yemen	49
Noruega	97	Bolivia	71	Ruanda	48
Bélgica	96	Senegal	71	Togo	48
Suiza	96	Papua N. Guinea	70	Sri Lanka	47
Austria	95	Paraguay	70	Kenia	46
Canadá	94	Brasil	69	Uganda	46
Francia	94	Nepal	69	Turquía	44
Irlanda	94	Sierra Leona	67	Paquistán	42
Reino Unido	93	Argelia	66	Tanzania	41
Portugal	92	Honduras	65	Zaire	40
Australia	91	Jordania	65	Indonesia	34
Benín	90	Zimbabwe	65	Camboya	33
Costa Rica	90	Méjico	64	Kuwait	33
EEUU	90	Guatemala	62	Malawi	33
Italia	90	Tailandia	62	Cuba	30
Uruguay	90	Malasia	61	Siria	30
España	87	Colombia	60	Arabia Saudí	29
Grecia	87	Singapur	60	Afganistán	28
Argentina	84	Túnez	60	Angola	27
Trinidad	84	Bangladesh	59	Vietnam	27
Bulgaria	83	Corea del Sur	59	Libia	24
Ecuador	83	Zambia	57	Irán	22
Polonia	83	Camerún	56	China	21
Japón	82	Marruecos	56	Corea Norte	20
Rumanía	82	Yugoslavia	55	Sudán	18
Panamá	81	India	54	Birmania	17
Chile	80	Perú	54	Iraq	17
Botswana	79	Unión Soviética	54		

Si se contrastan los resultados de Humana con los de Freedom House se puede observar que no existen excesivas disonancias. De hecho, David L. Banks ha estudiado la fiabilidad de los estudios de

Humana y Freedom House comparando los resultados obtenidos por ambos, y esto pese a reconocer que tales estudios no son «directamente commensurables»³⁶, y que la metodología de Humana es posiblemente superior a la de Freedom House, dado que el primero analiza libertades específicas y concretamente definidas, mientras el segundo hace una tabla «impresionista» de derechos civiles y políticos. A lo que habría que añadir que Humana tiene además el mérito de haber corregido su metodología a lo largo de sus sucesivos estudios, en el sentido de que los ítems contenidos en su cuestionario se han ido adaptando en función de las objeciones que se le han hecho (algo que no puede decirse del informe de Freedom House, al menos al mismo nivel) y se corresponden todos (en el estudio de 1991) con derechos reconocidos en la Declaración Universal de los Derechos Humanos, el Convenio Internacional sobre Derechos Económicos, Sociales y Culturales y el Convenio Internacional sobre Derechos Civiles y Políticos³⁷. Por el contrario, los estudios de Freedom House están orientados, con un cierto sesgo ideológico, hacia un sentido de las libertades públicas

³⁶ BANKS, David L. *New patterns of oppression: an updated analysis of human rights data*. En JABINE, Thomas B. y CLAUDE, Richard P (eds). *Human Rights and Statistics. Getting the record stright*. University of Pennsylvania Press, Philadelphia, 1992. p. 366.

³⁷ Algunos ítems empleados en su cuestionario de 1983 fueron eliminados precisamente por no corresponder con derechos explícitamente reconocido por las declaraciones y convenios internacionales, como por ejemplo el derecho a comprar y consumir bebidas alcohólicas, el castigo por posesión de literatura prohibida, la ratio entre policías y militares y la población civil, etc. Algunas de estas cuestiones servían de indicadores indirectos de la situación general de disfrute de derechos fundamentales en los diversos Estados, pero la inferencia de conclusiones sobre indicadores indirectos se presta a juicios personales que restan fiabilidad a un estudio que se presenta como resultante de la aplicación de metodologías cuantitativas. Otros elementos del cuestionario fueron mejor precisados en las sucesivas ediciones del estudio; así, por ejemplo, el derecho a la igualdad de las mujeres contenido en el estudio de 1983 fue sustituido en 1986 por dos nuevos apartados: derecho a la igualdad civil y política de las mujeres, y derecho a la igualdad social y económica de la mujeres. De esta manera Humana trataba de adecuar su metodología a la línea del Convenio sobre la eliminación de todas las formas de discriminación contra la mujer.

derivado del Bill of Rights Estadounidense más que de la Declaración Universal de los Derechos Humanos. La consecuencia es que la tabla de variables que cada uno emplea difiere de la del otro, y por tanto no pueden contrastarse directamente los resultados obtenidos.

Por otra parte, tampoco son directamente comparable los resultados de Humana y Freedom House porque cada uno emplea una escala de valores distinta. Freedom House puntúa a cada Estado según dos tablas que cubren una escala de siete puntos: la primera para medir los derechos civiles y la segunda para los derechos políticos. Humana, por su parte, asigna valores entre uno y cuatro a cada uno de los cuarenta ítems que componen su cuestionario. Para eludir estos obstáculos, que impiden una comparación lineal de resultados, Banks emplea dos recursos. Primero elabora una tabla de conversión de los valores numéricos de Freedom House y Humana, calculando las proporciones de variación, de forma que los resultados finales se vuelven conmensurables. Y segundo, selecciona de las tablas de variables de ambos estudios aquellas que tienen una clara equivalencia (por ejemplo, selecciona de la tabla de Humana ANYRELGN, DETENTN, ETHLANG, FREEINFO, entre otros, como variables directamente comparables con la tabla de derechos civiles de Freedom House, y INDCOURT, KEEPCIT, PUBTRIAL, NOIDEOL, etc. como correlatos de la tabla de derechos políticos de Freedom House), de modo que también en este aspecto se posibilite la conmensurabilidad. Aquellos ítems que difieren claramente son eliminados.

Con este procedimiento previo, se analizan los resultados obtenidos en ambos estudios, y Banks encuentra correlaciones muy altas entre ambos. Por ejemplo, en los estudios correspondientes a 1983, la correlación entre los derechos civiles evaluados por Humana y Freedom House es de 0.895 y la de derechos políticos 0.900. En los estudios de 1986 la correlación es aún más alta: 0.914 para derechos civiles y 0.912 para derechos políticos. La conclusión de Banks es que la fiabilidad de ambos estudios es, por tanto, muy alta, pues trabajando cada uno con su propia

metodología llegan a resultados equivalentes, lo que cumple el requisito de rigor científico de que cualquier experimento será fiable si un segundo investigador, repitiendo el protocolo realizado por el primero, llega a idénticas conclusiones. Como habrá de verse, esta fiabilidad es empleada por Banks como piedra de toque para emprender análisis estadísticos más ambiciosos, como reducciones de variables o árboles de clasificación y regresión, a partir de lo cual extrae algunas conclusiones teóricas acerca de los factores más determinantes de la eficacia de las normas internacionales sobre derechos humanos. Pero también habrá que considerar, en el apartado sobre objeciones a los estudios cuantitativos sobre derechos humanos, las críticas que otros autores han hecho a la fiabilidad de los análisis de Humana y Freedom House.

4.4. OBJECIONES A LAS FUENTES CUANTITATIVAS Y A LAS CONCLUSIONES TEÓRICAS

Antes de adentrarnos en una revisión de las objeciones que se han hechos a los sistemas cuantitativos de tratamiento de la información sobre los derechos humanos y a las hipótesis teóricas desarrolladas a partir de ellos, debe adelantarse ya como una primera conclusión que el simple examen realizado sobre las fuentes disponibles muestra que estos sistemas se encuentran en una fase de desarrollo bastante incipiente. Por tanto, el enfoque que ha de emplearse no es el de una revisión crítica de un sistema consolidado, sino más bien el de un conjunto de consideraciones acerca de qué base aprovechable hay hoy en día de cara a las posibilidades futuras de un procedimiento riguroso de obtención y procesamiento de datos, así como una metodología fiable para la inferencia de conclusiones teóricas suficientemente fundadas. En otras palabras, hay que considerar los modelos cuantitativos de información sobre los derechos humanos como una tarea por realizar, y lo emprendido hasta ahora en este campo como una

simple piedra de toque para evaluar las deficiencias a enmendar. Consecuentemente, las observaciones críticas que se hagan tendrán el sentido de señalar las correcciones necesarias para que esta metodología pueda rendir en el futuro mejores frutos. También en esa misma línea se considerarán en el siguiente capítulo las propuestas de diseño de modelos alternativos.

Antes de entrar en observaciones pormenorizadas sobre los sistemas disponibles de tratamiento cuantitativo y estadístico sobre los derechos humanos hay una primera consideración de carácter general que merece ser mencionada. La mayoría de los estudios realizados hasta el momento provienen de Estados Unidos y han sido realizados por investigadores estadounidenses. Este dato no es irrelevante, tal como afirman Scoble y Wiseberg:

«En los típicos análisis estadounidenses, los derechos civiles y políticos individuales están concretamente definidos, son numerosos y tienen la más alta prioridad. La mayoría de los relacionados con el grupo de los derechos económicos, sociales y culturales son añadidos en una consideración más bien vaga, están definidos genéricamente y aparecen en un contexto negativo: no deben condicionar o comprometer los derechos civiles y políticos. Esta aproximación al problema de la definición de los derechos humanos sitúa a Estados Unidos en una posición intelectual y políticamente diferenciada respecto a otros países, incluidos los de Europa occidental al definir el contenido de los derechos humanos. Los derechos socioeconómicos reciben un mucho mayor énfasis en las definiciones de otros países. Si el investigador acepta acríticamente una definición estadounidense etnocéntrica como universal, el producto de la investigación inevitablemente estará vinculado a prejuicios culturales, y por tanto hasta el punto de invalidar su propósito declarado»³⁸.

También William Safran insiste en que «los investigadores americanos [léase estadounidenses] están inclinados a ver el sistema de libertades civiles de Estados Unidos como un mode-

³⁸ SCOBLE, Harry M. y WISEBERG, Laurie S. *Problems of Comparative Research on Human Rights*. En NANDA, Ved P., SCARRITT, James R. y Shepherd, George W. (Jr). *Global Human Rights: public policies, comparative measures and NGOs strategies*. Westview Press. Boulder (Colorado), 1981. p. 148. (La traducción es mía)

lo»³⁹. Y Kenneth A. Bollen coincide igualmente en el potencial de desfiguración por prejuicios que supone la desproporción de investigadores estadounidenses en esta materia⁴⁰. Ciertamente el hecho de que la sociología de los Estados Unidos tenga una marcada tendencia al uso de metodologías estadísticas, mayor que la habitual en otros contextos, es un hecho que no necesariamente tendría que afectar a la fiabilidad de los estudios cuantitativos sobre derechos humanos. Sin embargo, el substrato cultural de los investigadores de aquel país termina por influir, a través de ciertos prejuicios, en el propio planteamiento de la investigación. Este factor no debe ser desconsiderado como un condicionante general de la exactitud de los resultados que se obtengan, pero evidentemente afecta al modo de conceptualización de los derechos.

4.4.1. Objeciones a los sistemas de obtención de datos

Al margen de prejuicios a la hora de establecer qué datos son relevantes y cómo interpretarlos, lo que sería un condicionante externo al uso mismo de una metodología estadística para analizar la situación internacional de los derechos humanos, puede señalarse un primer problema ya directamente relacionado con el procesamiento de los datos. Se trata de la fiabilidad que puede tener la información cuando se consideran las dificultades de acceso a la misma, lo que permite sospechar que buena parte de los datos que actualmente se manejan pueden estar distorsionados por la existencia de importantes lagunas. George A. López y Michael Stohl⁴¹

³⁹ SAFRAN, William. *Civil liberties in democracies: constitutional norms, practices and problems of comparison*. En NANDA, Ved P., SCARRITT, James R. y Shepherd, George W. (Jr). Op. Cit. p. 195

⁴⁰ Cfr. BOLLEN, Keneth A. *Political rights and political liberties in nations: an evaluation of human rights measures. 1950 to 1984*. En JABINE, Thomas B. y CLAUDE, Richard. P. Op. Cit. p. 192.

⁴¹ LÓPEZ, George A. y STOHL, Michael. *Problems of concept and measurement in the study of human rights*. En JABINE, Thomas B. y CLAUDE, Richard. P. Op. Cit. pp. 217-221.

—ya mencionados en el capítulo 1— han llamado la atención sobre este punto en relación con los datos relativos a las violaciones de los derechos humanos. En primer lugar señalan que la ausencia de conductas detectables contrarias a los derechos humanos en un determinado momento puede no ser un indicador fiable de verdadera ausencia de violaciones. Así, en muchos Estados autoritarios dotados de mecanismos eficientes para violentar los derechos humanos, se consolida un efecto persuasivo en la población que sigue operando por «inercia» incluso bastante tiempo después de que se produzcan violaciones explícitas y públicas de los derechos fundamentales. Esta interiorización del terror basta a los gobiernos autoritarios para ejercer mecanismo coercitivos «silenciosos» que pueden pasar desapercibidos en un superficial recuento de abusos. También Ted R. Gurr y Barbara Harff han coincidido en este mismo punto de crítica⁴².

Otro problema puede presentarse cuando se emplea un sistema comparativo de la situación de los derechos humanos que resulte en una clasificación de los diversos Estados en una tabla por puntos (como es el caso de Humana o Freedom House). Esta metodología trata a los Estados como unidades indivisibles, pero a veces las violaciones no son emprendidas por el Estado en su conjunto, sino por determinados órganos. Si las violaciones se producen en un período prolongado, los agentes que las ejecutan desarrollan objetivos por sí mismos, con lo que un cambio en el gobierno en ocasiones no termina con las prácticas de estos cuerpos represivos. Es lo que ha ocurrido en algunos países africanos en la década de los 70, y en Guatemala o El Salvador entre 1975 y 1984. No hay normas, ni decretos del gobierno, ni estados de excepción abusivamente prolongados que atenten contra los derechos humanos, pero determinados cuerpos policiales, militares o paramilitares siguen desarrollando una actividad represiva por su cuenta y

⁴² HARFF, Barbara y GURR, Ted R. *The rights of collectivities: principles, and procedures in measuring the human rights status of communal and political groups*. En JABINE, Thomas B. y CLAUDE, Richard P (eds). Op. Cit. p. 160.

frecuentemente contra las órdenes del gobierno. En tales casos, si se evalúa la acción del Estado por el tipo de decisiones que adoptan sus más altas instancias pasan desapercibidos hechos decisivos.

Añaden López y Stohl que otro problema resulta de las situaciones de guerra civil, donde la falta de protección de los derechos humanos puede deberse a la inestabilidad política y a excesos de algunos de los bandos beligerantes, sin que pueda asignarse claramente a las autoridades públicas la responsabilidad por estas violaciones. En países con un largo historial de falta de respeto a los derechos fundamentales, y donde se persigue sistemáticamente a los disidentes, esta represión puede cesar como tal cuando se desata una guerra civil, ocasión que el gobierno emplea como coartada, encubriendo entonces la misma represión como una simple campaña militar contra los «terroristas» que se han alzado en armas contra el poder «constitucional». Por último, queda el problema de la definición de «arbitrariedad» que se emplee. El concepto es decisivo para tipificar y calificar supuestos de falta de garantías procesales, detención arbitraria y, en general, toda decisión pública que no se atenga al principio de legalidad. Sin embargo, hay muchas formas de acomodar las violaciones de los derechos fundamentales a la legalidad vigente y, a la inversa, existen tantas variaciones en los sistemas legales de los diversos Estados que cualquier calificación de los hechos como violación de los derechos humanos en función de la ilegalidad de las decisiones resulta muy difícil, sobre todo si se trata de emplear un único patrón de medida supuestamente unívoco e igualmente válido para todo Estado. Como habrá de verse más adelante, el caso del derecho de asociación y de manifestación pacífica en Estados Unidos y Japón es sintomático, y en el caso de la competencia de determinadas autoridades administrativas para autorizar la celebración de manifestaciones pacíficas, se han podido tomar en Japón decisiones restrictivas dentro de la legalidad que, en igualdad de condiciones, han motivado en Estados Unidos recursos ante los tribunales que han anulado tal tipo de decisiones.

Keneth Bollen⁴³ añade a las observaciones anteriores otro problema no menos importante relativo a los canales de obtención de la información. La situación ideal respecto de los datos sería aquella en que estuviesen contenidas todas las características relevantes para expresar la situación de los derechos humanos en un determinado Estado. Pero no todas las características son accesibles. Pongamos por caso que en un Estado hay deficiencias en el censo electoral, deficiencias que el gobierno conoce pero que no tiene interés en corregir porque incluiría a una parte de la población previsiblemente opuesta a la política gubernamental, por ejemplo, campesinos de una zona ocupada por la guerrilla y que apoya a ésta. En casos como este resulta imposible saber quién no está en el censo y, por tanto, tiene anulado su derecho de participación política. Luego del total de información de la situación ideal hay que descartar la parte que no queda registrada en ningún documento.

Pero además, no toda la información registrada es accesible. Informes policiales o militares que son declarados secretos, estadísticas de uso exclusivo del gobierno que no se hacen públicas, etc. Por tanto, hay una segunda restricción que limita más la disponibilidad de información. Además, de la información registrada y accesible, no toda es empleada de hecho por las organizaciones no gubernamentales (desde la prensa hasta grupos de activistas) que, en caso de haber un gobierno autoritario, serían las que podrían denunciar violaciones, algo que el propio gobierno no va a hacer. A esta nueva limitación hay que añadir que sólo una parte de la información registrada, accesible y empleada por la prensa u organizaciones del país atraviesa las fronteras y se difunde a nivel internacional. Y por último, con cierta ironía, Bollen constata que sólo una parte de la información disponible a nivel internacional es empleada por los investigadores en Estados

⁴³ BOLLEN, Keneth A. *Political rights and political liberties in nations: an evaluation of human rights measures, 1950 to 1984*. En JABINE, Thomas B. y CLAUDE, Richard. P. Op. Cit. pp. 188-216

Unidos (pues es allí donde se hace la inmensa mayoría de estudios estadísticos sobre los derechos humanos). Teniendo en cuenta todas estas limitaciones, Bollen expone sus dudas acerca de la fiabilidad de estudios como los de Freedom House, Humana, el World Handbook of Political and Social Indicators, etc., y añade que «irónicamente, es posible que una nación que sea relativamente abierta pueda aparecer más baja en derechos y libertades [en los rankings arriba mencionados] simplemente porque es más fácil que llegue información al exterior sobre las violaciones de los derechos humanos»⁴⁴. Teniendo en cuenta este problema, el Programa de Acción de la Conferencia Mundial sobre Derechos Humanos, organizada por Naciones Unidas en Viena, del 14 al 25 de Junio de 1993, estableció en el punto 39:

«La Conferencia Mundial de Derechos Humanos, haciendo hincapié en la importancia de disponer de información objetiva, responsable e imparcial sobre cuestiones humanitarias y de derechos humanos, pide una mayor participación de los medios de información, a los que la legislación nacional debe garantizar libertad y protección»

4.4.2. Objeciones desde los criterios estadísticos de fiabilidad

Uno de los problemas más reiteradamente detectados en los sistemas desarrollados hasta el momento para evaluar comparativa y cuantitativamente la situación internacional de los derechos humanos es el de su escasa fiabilidad atendiendo a los criterios habituales requeridos por las técnicas estadísticas. Esta circunstancia se hace tanto más grave cuantas más conclusiones teóricas se extraigan de informes lastrados por deficiencias de este tipo en el procesamiento de la información, como ocurre, por ejemplo, en el estudio de Arat que veremos en el siguiente capítulo. Russel Lawrence Barsh⁴⁵ ha hecho

⁴⁴ Idem. p. 200.

⁴⁵ BARSH, Russel Lawrence. *Measuring Human Rights: problems of methodology and purpose*. En «Human Rights Quarterly», n.º 15/1 (Febrero de 1993). pp. 87-121.

un seguimiento de los sistemas de indicadores sobre derechos humanos existentes analizándolos desde estos criterios de rigor técnico, que se sintetizan en seis: la definición y medida, la precisión, la consistencia, la validez, la equivalencia y la agregación.

La definición y medida es un conjunto de exigencias que, a diferencia de otros criterios como la fiabilidad, a menudo ni siquiera son planteadas como problemáticas, lo que puede dar imagen de seriedad metodológica a un estudio que, sin embargo, tal vez está viciado desde el origen. Los sistemas de indicadores responden al propósito de cuantificar conceptos que, de suyo, o bien no son directamente cuantificables, o bien son incluso imposibles de observar. Evidentemente no se puede «medir» la democracia, como tampoco puede medirse la satisfacción en abstracto. Por eso se emplean indicadores, que miden indirectamente estos conceptos. Muchos sondeos regulares de opinión, bien conocidos por aparecer publicados en la prensa, incluyen cuestiones como la satisfacción de los ciudadanos ante la eficacia de un determinado servicio público, la percepción personal de las expectativas de la propia economía doméstica, la sensación personal de seguridad en las calles, etc. Se supone que este conjunto de cuestiones (junto con otras) puede ser un indicio del concepto abstracto «satisfacción». Las respuestas obtenidas en diferentes encuestas periódicamente repetidas pueden mostrar entonces variaciones que pueden ser correlacionadas con variaciones en otros indicadores, por ejemplo la tasa de desempleo, el precio de la vivienda, la valoración del gobierno o las expectativas de voto de determinados partidos. De estas correlaciones pueden extraerse conclusiones significativas y relativamente fiables, pero en cualquier caso no se está midiendo la satisfacción en sí, pues esto es imposible. La cuestión es, entonces, localizar cuál es el mejor indicador para el concepto que se quiere evaluar, y mostrar que, efectivamente, el concepto se encuentra apropiadamente representado en el indicador. A esto se refieren las tareas de definición.

De la misma manera, la democracia o los derechos humanos son conceptos no mensurables en cuanto tales. La cuestión es definir un sistema de indicadores que puedan servir como indicios

de las variaciones en el concepto subyacente. Pero qué relación haya entre el indicador y el concepto es algo que no puede obviarse, y para probar que un indicador es un buen indicio de un concepto hace falta una teoría que explique el por qué de esta relación. Este es el aspecto que muchos estudios descuidan, dando ingenuamente por supuesto que la relación es directa y autoevidente. Y este es el problema del modelo del normativismo jurídico que vimos en el primer capítulo; la suposición de que por estar los derechos humanos expresados en un conjunto cerrado de normas, la misión de los indicadores es comprobar su eficacia, como si no fuese necesario establecer en primer lugar qué relación existe entre el indicador y el concepto. Sin embargo, incluso la demografía, que aparentemente podría tener menos problemas de este tipo, necesita previamente establecer acuerdos acerca de criterios precisos a la hora de incluir personas en los censos. Volviendo sobre cuestiones concretas sobre la medición de indicadores relativos a los derechos humanos, Spierer⁴⁶, por ejemplo, menciona el problema de cuantificar los prisioneros políticos. ¿Cuál es el criterio empleado para definir cuándo un prisionero es político? ¿A través de las razones oficiales dadas para su encarcelamiento? Pero muchos Estados autoritarios sortearán esta práctica censurable encubriéndola como delincuencia común. ¿Qué ocurre cuando se detiene a una persona por razones políticas, pero concurre además una infracción contra la ley vigente? Puede ocurrir que las verdaderas razones sean políticas pero que el detenido haya participado en una manifestación violenta no autorizada, y la razón oficial alegada será esta última. Por otra parte, los gobiernos autoritarios suelen caracterizarse por una muy amplia y estricta legislación reguladora de la seguridad del Estado, por lo que es fácil imputar delitos de este tipo a opositores políticos, respetando en apariencia las formalidades propias de un Estado de derecho. ¿Se computará entonces como delincuencia común? Como habrá de verse en las propuestas de

⁴⁶ SPIRER, Herbert F. *Violations of human rights-How many?* En «American Journal of Economics & Sociology», n.º 49/2 (1990). pp. 199-206. Citado por Barsh, R. L. Op. Cit. p. 93.

modelos alternativos, McCamant entiende que hay tantas dificultades en la definición operativa de los derechos humanos para un sistema de indicadores que éste resulta inviable, por lo que propone un cambio radical hacia otros conceptos.

La definición y validación de conceptos debería ser el primer paso para cualquier investigación cuantitativa, y más en el caso de que se quieran emplear técnicas estadísticas para extraer conclusiones generalizables que tengan un mínimo de fiabilidad. Sin embargo, el examen de la definición conceptual de los sistemas de indicadores desarrollados hasta la fecha muestra que este primer criterio ha sido ampliamente descuidado. Humana, por ejemplo, ha ido variando a lo largo de los años sus sistema para corregir los defectos señalados por sus críticos. Su criterio final ha sido el de ceñirse al máximo a los textos legales internacionales sobre los derechos humanos, de modo que lo que se evalúa parece incuestionable, por ser una aplicación directa de los mismos derechos humanos jurídicamente reconocidos. Sin embargo, y por debajo de esta apariencia, todavía hay decisiones sobre el sistema cuya razón no es discutida. Por ejemplo, de los 40 apartados que componen su cuestionario, 7 están referidos a la libertad de prensa o de las comunicaciones (lo que les da un peso relativo del 17,5% sobre el total), mientras sólo 3 se refieren al derecho al trabajo y los derechos sindicales, es decir, menos de la mitad. ¿Por qué casi el 20% de la información relevante sobre los derechos humanos tiene que referirse a la libertad de información, de prensa y de las comunicaciones? Si se consideran tan relevantes estos apartados debe haber alguna razón que lo explique, pero Humana no hace objeto de discusión este punto. De acuerdo con el modelo propuesto en la primera parte, un primer factor a considerar debería ser la eficacia del monopolio de la violencia legítima (lo que se verá confirmado, como tendremos ocasión de ver, por algunas investigaciones empíricas), y hay buenas razones para considerar prioritario este factor, pues sin él es imposible que arraigue eficientemente un sistema jurídico garante de los derechos humanos, como he mostrado siguiendo a Weber y Elias. Si el modelo es otro, como el que implícitamente propone Humana dando tanta importancia a la

libertad de prensa, hay que exponer qué conceptualización de los derechos humanos se hace para que este derecho sea tan prioritario, y avalarlo con razonamientos convincentes, pero no puede darse por supuesto.

La precisión de los estudios estadísticos se refiere a los criterios de medición. Todo sistema que opere con indicadores cuantitativos presupone que la asignación numérica resultante es una expresión de características del objeto estudiado. En qué medida exista correspondencia entre valores numéricos y caracteres observables es una cuestión de precisión. Hay tres formas clásicas de establecer sistemas de medición: el nominal, el ordinal y el sistema de intervalos o ratios. El primero sigue una tipología clasificatoria de las características, siendo éstas sustancialmente distintas, y se emplea cuando se dan propiedades no conmensurables. Por ejemplo, el estudio de Scarritt citado anteriormente se preocupa de establecer criterios para poder calificar determinados sucesos como politocidios o genocidios. La inclusión en un tipo excluye, naturalmente, la inclusión en el otro, y se considera que sucesos de politocidios no son directamente comparables, empleando una misma escala, que los genocidios. El sistema ordinal emplea una tabla de mayor a menor con niveles fijos y estandarizados de variación, por tanto, presupone una propiedad que puede ser unidimensionalmente cuantificada como función discontinua, por «pasos». El sistema de medida por ratios es una escala con variación continua (como la de los grados que miden la temperatura, que permiten subdividir indefinidamente las unidades). El problema relativo a la precisión es que frecuentemente estos niveles se confunden, y tras establecerse un sistema nominal, luego es combinado con otro ordinal o por ratios sin establecer las correspondencias. Por ejemplo, Freedom House define sus criterios clasificatorios respecto de los derechos civiles y políticos, y al margen de la vaguedad con que lo hace, esto supone emplear un sistema nominal. Pero además aspira a traducir numéricamente la situación de cada Estado en relación con cada uno de los tipos definidos. Para ello emplea un sistema ordinal con una escala de 1 a 7. Pero ¿cuál es el criterio para que un Estado reciba una

puntuación específica? John F. McCamant menciona el estudio correspondiente a 1973, y reproduce el criterio aducido por Freedom House. Los puntos correspondientes a los derechos civiles son asignados de la manera siguiente:

1 punto. El Estado de derecho es incontestable. La variedad de los medios de comunicación y la libertad de expresión es tanto posible como evidente.

2 puntos. Tiende a lo anterior, pero a causa de la falta de medios de comunicación libres o suficientes, los derechos civiles son menos efectivos (...).

3 puntos. Tiene todos los atributos de la libertad civil, y el gobierno es contestado con éxito ante los tribunales. Pero todavía hay serias imperfecciones. Estas pueden incluir el apoyo reiterado en la ley marcial, la prisión por sedición o la supresión de publicaciones.

La tabla sigue de manera semejante hasta...

7 puntos. El mundo exterior apenas tiene noticias de críticas hacia el gobierno, salvo porque éste las condena. Los ciudadanos no tienen derechos vis-à-vis el Estado.

Ciertamente, el criterio para traducir numéricamente propiedades de los derechos civiles es tan subjetivo, vago y confuso que no resiste la más mínima revisión de fiabilidad. De hecho, McCamant hizo la prueba de eliminar el ranking de Freedom House, desordenar los resultados y entregar esta información a cinco estudiantes de estadística para que intentasen reproducir la clasificación en función de la gravedad de las violaciones de los derechos humanos. Ninguno coincidió en la prueba con la clasificación hecha por Freedom House⁴⁷. Es cierto que esta prueba se hizo respecto al estudio de 1973, pero desde entonces los criterios

⁴⁷ McCAMANT, John F. *A critique of present measures of human rights development and an alternative*. En NANDA, Ved P., SCARRITT, James R. y Shepherd, George W. (Jr). Op. Cit. p. 132

de Freedom House no han variado sustancialmente, y siguen siendo objeto de crítica por ello.

La consistencia, o fiabilidad, es un criterio de repetición. Cuando un observador distinto al que ha hecho la primera cuantificación aplica los criterios de medición establecidos por éste y obtiene los mismos resultados que el primero, el sistema de medición es fiable y no susceptible de tergiversaciones subjetivas. Como se ha visto en el caso citado de McCamant, la prueba realizada sobre el informe de Freedom House se aproximaba mucho a una de consistencia, con el resultado mencionado. No obstante, parte de la información empleada por los sistemas de evaluación de los derechos humanos depende de los propios medios de los investigadores, o de la selección que ellos hagan. Hacer pruebas de consistencia requiere habitualmente un esfuerzo no menor que el de los propios informes. Por esta razón, y salvo los casos ya mencionados anteriormente, no se han hecho este tipo de pruebas de manera habitual en los estudios cuantitativos sobre derechos humanos.

Como señala Kenneth Bollen, «la fiabilidad se refiere a la consistencia de las medidas. Esto es, en qué grado dos medidas de la misma cuestión aportan los mismos resultados. La validez concierne a si uno está realmente midiendo un concepto»⁴⁸. Por tanto, ambas difieren. El estudio hecho por Banks⁴⁹ acerca del rigor estadístico de los estudios de Humana y Freedom House encuentra correlaciones muy significativas entre los resultados de ambos, pero con esto, en todo caso, está comprobando la consistencia de tales estudios, no su validez. Una cosa no presupone a la otra, y Banks no se preocupa de preguntarse por la segunda. Probablemente porque la discusión sobre la validez no resulta tan simple. Como señala Barsh;

⁴⁸ BOLLEN, Keneth A. *Political rights and political liberties in nations: an evaluation of human rights measures, 1950 to 1984*. En JABINE, Thomas B. y CLAUDE, Richard. P. Op. Cit. p. 207.

⁴⁹ Vid. supra notas 28 y 29.

«La relación entre la medición o indicador y la característica o variable a medir debe estar ella misma basada en una teoría, que debería ser puesta a prueba independientemente de la hipótesis que haya sido diseñada para ayudar a su comprobación. Por el contrario, las publicaciones de las ciencias sociales abunda en arbitrariedades, opiniones comunes o reglas a priori de correspondencia, que deberían ser sospechosas como sistemáticamente cargadas de prejuicios o de etnocentrismo»⁵⁰.

He aquí un punto crucial. Salvo en el caso de los estudios «Minorities at risk» y «Genocides and politicides since World War II», y algún otro caso menor, este es el aspecto más descuidado de las investigaciones empíricas sobre derechos humanos. Sin una mención explícita a los supuestos teóricos que subyacen a la formalización de indicadores, la validez de estos sistemas resulta en una petición de principio. Toda apariencia de rigor basada en el empleo de sistemas cuantitativos (el fetichismo de los números) se desvanece en este punto. Ya ha habido ocasión de comentar la carga ideológica que inspira el trabajo de Freedom House. La validez de su sistema de cuantificación está lastrada por la decisión puramente personal de adherirse a los valores que los autores, más o menos explícitamente, sostienen. El caso de Humana no es, probablemente, tan grave. Pero como habrá ocasión de ver en el siguiente capítulo, las cargas de etnocentrismo que indica Barsh tienen una buena prueba en la respuesta que Humana dio cuando el Human Development Programme de las Naciones Unidas decidió no volver a asumir su informe. En el caso de otros informes, los presupuestos teóricos, que no se hacen suficientemente explícitos, van en la línea de buscar únicamente correlaciones monocausales entre desarrollo económico y derechos humanos.

La equivalencia es una cuestión relativa a la medida en que los resultados cuantitativos pueden ser asignados con un mismo

⁵⁰ BARSH, Russel Lawrence. *Measuring Human Rights: problems of methodology and purpose*. En «Human Rights Quarterly», n.º 15/1 (Febrero de 1993), p. 96.

significado a situaciones muy diversas entre sí. Las objeciones a la equivalencia de los estudios cuantitativos sobre derechos humanos abundan en el argumento de que no hay un significado unívoco de los derechos humanos a través de la variedad de culturas. Algunos autores consideran esto un obstáculo insalvable, que impide todo tratamiento cuantitativo de esta cuestión, y por tanto sólo cabrían estudios cualitativos de tipo etnográfico. De todos modos, este problema se tratará separadamente en el siguiente apartado, dedicado a los problemas del multiculturalismo.

La agregación, por último, se ocupa del modo de organización y síntesis de los resultados. Cualquier sistema de indicadores reúne una cantidad muy importante de información, con el objeto de poder inducir a partir de ella un conjunto más limitado de datos relevantes. Pero, ¿cómo se sintetiza el conjunto total de la información? El Índice de Desarrollo Humano, por ejemplo, mezcla datos de ingresos brutos per cápita con datos relativos a la educación y la salud. De suyo estas son categorías heterogéneas entre sí, por lo que no podría hacerse una suma lineal de ambas. ¿Se pueden mezclar en un indicador cuantitativo factores heterogéneos? También esta es una cuestión que requiere una fundamentación teórica que justifique que la agregación está midiendo parámetros que pueden ser indicios de una realidad única. Y, por tanto, también aquí hay un peligro potencial de tergiversaciones por prejuicios personales del investigador.

Otro problema relacionado con la fiabilidad de los criterios estadísticos es el de las unidades de imputación de datos. Este problema fue considerado, por ejemplo, por Naciones Unidas en su Informe sobre Desarrollo Humano de 1993. Dicho informe nació como una alternativa al sistema clásico que evaluaba el nivel de desarrollo de los Estados en función del PIB, y en todo caso considerando las desigualdades de distribución de la renta. Por el contrario, el Informe sobre Desarrollo Humano emplea otros indicadores relativos al nivel educativo y la protección de la salud (vinculados a derechos sociales). Relacionando estos factores se obtiene un índice internacional de desarrollo en el que éste no es ponderado en términos exclusivamente económicos, sino que

incluya también en el concepto de desarrollo el derecho a la cobertura de necesidades básicas. Pero en el año 1993 los autores del informe incluyeron un apartado dedicado a considerar el problema de las unidades de imputación de datos. Las medias estadísticas ocultan a menudo situaciones heterogéneas, que precisamente desaparecen porque se trata de obtener una media. En el caso del informe mencionado se probó a desagregar los datos de desarrollo por grupos y regiones en tres Estados: México, India y Estados Unidos⁵¹. Este último ocupa el segundo lugar mundial en el índice de desarrollo humano, después de Japón. Pero desagregando los datos resulta que la población blanca de Estados Unidos tiene un índice de desarrollo humano superior al de los japoneses, mientras la minoría negra estadounidense tiene una posición equivalente a la de Trinidad Tobago (y que aún es superior a la de los hispanos de Estados Unidos). Diferencias semejantes se producen entre distintas regiones de la India o de México. E igualmente se producen disparidades en función de los sexos. El ejemplo es directamente traducible a los sistemas de indicadores sobre derechos humanos. Esta misma cuestión es también señalada por Russel L. Barsh:

«Si la cuestión objeto de investigación está relacionada con el comportamiento de los gobiernos, entonces las unidades de análisis deben ser los gobiernos. Si está relacionada con el interjuego entre el entorno socio-político y las actividades económicas, las unidades de análisis apropiadas son los actores económicos, tales como los individuos, corporaciones, o comunidades relativamente autosuficientes. El análisis debería ser llevado a cabo al nivel en que se toman las decisiones relevantes. Ciertos individuos y grupos dentro de un mismo Estado están sujetos a niveles de riesgo mucho mayores, como por ejemplo los indígenas de Guatemala. Un estudio comparativo que compara Guatemaltecos con Nigerianos parece poder producir menos resultados válidos que uno que compare Guatemaltecos de alto y bajo riesgo»⁵².

⁵¹ United Nations Development Programme. *Informe sobre desarrollo humano 1993*. Editado por el Centro de comunicación, investigación y documentación entre Europa, España y América Latina. Madrid, 1993. pp. 19-23.

⁵² BARSH, Russel Lawrence. Op. Cit. p. 120.

La indicación de Barsh podría concretarse cruzando los datos del estudio *Minorities at risk*, ya mencionado, con los sistemas de indicadores generales de Humana o Freedom House. Evidentemente, dentro de las medias resultantes en estos informes se producirían diferencias internas enormes en función de la existencia de minorías como los kurdos en Turquía, Irak o Irán, los tibetanos en China, los cristianos y animistas del sur de Sudán, etc. E incluso otras minorías no necesariamente étnicas o culturales, como los inmigrantes indocumentados de muchos países desarrollados. En cualquier caso, el problema de las unidades de imputación siembra una duda más que razonable sobre un aspecto de la fiabilidad de los estudios de tipo cuantitativo sobre los derechos humanos. Ciertamente, la solución «técnica» a este problema supone un incremento notable de la información a procesar, información que en el caso de muchos Estados será de difícil acceso o incluso inexistente. Pero esta razón no impide relativizar en gran medida el tipo de trabajos cuyo resultado es un ranking en que se expresa numéricamente la situación de un país, y esta media es poco más que una tapadera para un buen número de problemas que quedan ocultos tras ella.

4.4.3. Objeciones desde el multiculturalismo

Una de las objeciones más decisivas respecto a los sistemas de evaluación de la eficacia de las normas internacionales sobre los derechos humanos es el problema de la diversidad cultural frente a la univocidad de tales normas. Cuando, en 1991, el Programa de Naciones Unidas para el Desarrollo hizo propio el informe de Charles Humana, incluyendo su índice sobre los derechos humanos en el *Human Development Report*, el embajador de Ghana, Kafi Awoonor —hablando en representación del grupo de los 77— objetó que:

«La libertad es un concepto con una carga valorativa que encuentra expresión en diferentes formas de sociedad en sociedad. Por esta razón podría haber pensado que el Programa de las Naciones Unidas para el Desarrollo sería más cauteloso y selectivo al identificar las nociones de libertad que iban a ser incluidas en el índice. Usar el trabajo de un

investigador académico particular, representante de una cultura particular vista por muchos en la historia reciente como vinculada a la opresión y la explotación de una vasta parte de nuestro mundo, y desarrollar un índice que debería ser aplicable a todas las sociedades y culturas, es mostrar una falta de sensibilidad difícilmente aceptable en un cuerpo universal como el Programa de las Naciones Unidas para el Desarrollo»⁵³.

La protesta del Grupo de los 77 obligó al administrador del Programa de las Naciones Unidas para el Desarrollo (PNUD), William Draper, a salir en defensa de la fiabilidad del trabajo de Humana, pero para entonces ya corrían rumores entre los representantes de los Estados acerca de que las decisiones sobre la asignación de fondos de ayuda al desarrollo se iban a vincular a índices sobre el nivel de respeto a los derechos humanos como el de Humana, que la ONU había respaldado. De hecho, pocas semanas después el Banco Mundial publicó su *World Development Report* correspondiente a 1991⁵⁴, que contenía una matriz de correlaciones estadísticas basadas en el informe de Freedom House, y en el que intentaba demostrar la vinculación entre libertades públicas y crecimiento económico. La publicación de este informe acrecentó los rumores y, con ella, la movilización en contra del respaldo de la ONU hacia el trabajo de Humana. Al hacerse mayor la presión de muchos Estados, el Consejo de Gobierno del PNUD decidió desarrollar su propio sistema independiente de indicadores sobre los derechos humanos, abandonando por tanto el de Humana por las susceptibilidades que había levantado, pero en la reunión celebrada en Febrero de 1992 se acordó no volver a publicar ningún *ranking* de Estados, aunque la investigación sobre un sistema propio de indicadores podría continuar. También en 1992 Danilo Türk, relator especial de la Comisión de Derechos Huma-

⁵³ Discurso del Dr. Kofi Awoonor, embajador y representante permanente de Ghana en el debate general del Consejo de Gobierno del Programa de las Naciones Unidas para el Desarrollo, New York, 11 de Junio de 1991. Citado por Russel Lawrence Barsh. *Measuring Human Rights: problems of methodology and purpose*. En *Human Rights Quarterly*, n° 15/1 (Febrero de 1993), p. 87-88.

⁵⁴ Cfr. WORLD BANK. *World Development Report, 1991; the challenge of development*. Oxford University Press. New York, 1991.

nos de Naciones Unidas, en su informe final presentado a la 44ª sesión de dicha comisión, volvió a insistir en la necesidad de desarrollar un sistema de indicadores fiable para evaluar la realización de los derechos económicos, sociales y culturales. De hecho esta era una de las tres recomendaciones generales con que concluía su informe⁵⁵, y no sin haber recordado además que esta recomendación se venía haciendo desde el primer informe presentado a la Comisión en 1990⁵⁶. El asunto volvió a la luz en la Conferencia de Viena sobre Derechos Humanos, de Junio de 1993, cuyo Programa de Acción recoge, en el punto 98, aquel proyecto:

«Para fortalecer el disfrute de los derechos económicos, sociales y culturales, deberían examinarse otros métodos, como un sistema de indicadores para medir los avances hacia la realización de los derechos enunciados en el Pacto Internacional de Derechos Económicos, Sociales y Culturales».

Pero, entre tanto, no se ha hecho ningún avance en tal sistema de indicadores. Volvemos así al problema planteado al comienzo de este capítulo; que cuando median intereses políticos de los Estados resulta difícil lograr garantías de objetividad en un sistema de evaluación de los derechos. Y dado que los informes del Programa de las Naciones Unidas para el Desarrollo deben ser aprobados por los Estados miembros, aquellos que resulten mal catalogados en tal sistema de indicadores tenderán a descalificar la metodología⁵⁷. Pero lo que merece destacarse a este respecto es que el argumento empleado para invalidar el sistema de indicadores no atacaba la existencia de un patrón universal de los derechos humanos (la Declaración de 1948), sino que cada sociedad y cultura tiene su propio concepto de la libertad. Ciertamente, una

⁵⁵ TÜRK, Danilo (Special rapporteur for the Commission of Human Rights of the United Nations) *The realization of economic, social and cultural rights*. Final report submitted to the 44th session of the Commission of Human Rights. 3 de Julio de 1992. United Nations document GE.92-12667/4656B. p. 53

⁵⁶ Idem. pp. 1-2

⁵⁷ Incidentalmente, Ghana ocupaba la posición 72, de un total de 104, en el ranking de Humana.

noción tan genérica como la de la libertad —tal como la empleaba en su argumento el embajador de Ghana— es difícil de concretar en un sistema de evaluación cuantitativa universalmente aceptado. Pero aparentemente, como argumentaba Cassese, ya mencionado al inicio de este capítulo, debería ser más fácil llegar a un acuerdo universal acerca de cuestiones más concretas, como el hecho de informar al detenido de los cargos que se le imputan, que nadie pueda ser detenido sin cargos, o que la legislación de un Estado no establezca discriminación en los derechos de un determinado grupo étnico. Para este tipo de cuestiones, más concretas que la noción genérica de libertad, sí debería ser posible llegar a un consenso por encima de la diversidad cultural. La Declaración Universal de los Derechos del Hombre es el punto de partida, y por esa razón Humana ha ido reformulando su metodología para que se corresponda máximamente con derechos formalmente reconocidos en dicha Declaración o en los Pactos Internacionales. Ahora bien, tal vez el optimismo de Cassese no sea tan fácil de llevar a la práctica. Dejando de lado argumentos sobre el relativismo cultural como el empleado por el embajador de Ghana, hay un segundo nivel de análisis en que el problema del multiculturalismo se presenta como una objeción más seria a la posibilidad de un sistema de indicadores sobre los derechos humanos.

A raíz de la polémica levantada por la inclusión de su informe en documentos oficiales de Naciones Unidas y la posterior decisión de no volver a recurrir a él hasta que no se elaborase un sistema de indicadores propio de Naciones Unidas y que resultase menos controvertido ante los Estados miembros, Charles Humana reaccionó contestando en su siguiente informe a quienes habían protestado contra su sistema de evaluación. Esta respuesta es interesante por lo que revela acerca de lo que el autor se proponía con estos trabajos y de los presupuestos que subyacían a ellos.

«La oposición a la medición de derechos y a la clasificación de los Estados se centra en la percepción de que el objetivo está marcado por un prejuicio occidental. Esto es innegablemente cierto. Los tratados de Naciones Unidas sobre los derechos humanos reflejan tradiciones y valores liberales occidentales porque fueron fundamentales para la creación de la organización mundial. Sus tratados siguieron los precedentes de

tratados occidentales anteriores, y lo que puede denominarse moralidad internacional social y política de la segunda mitad del siglo XX tiene por tanto un origen similar (...). Pero en el seno de Naciones Unidas hay una seria contradicción. Muchos países que disfrutaban de plena condición de miembros y que participan en todas sus actividades se adhieren a tradiciones, religiones y creencias que realmente no pueden coexistir con las ideas liberales occidentales. Este conflicto ha sido resuelto mediante un experimentado ejercicio de tolerancia, olvidos voluntarios, diplomacia e insinceridad, que permite seguir a su aire a los Estados que ni hacen honor a sus obligaciones para con los instrumentos de Naciones Unidas, ni pueden tampoco hacerlo (...). El cuestionario de esta guía no ha sido construido para adaptarse a los patrones y las prácticas que podrían ser más aceptables para los países que faltan a las obligaciones que han contraído en los tratados. Si hay que soportar que los indicadores sean etiquetados de «liberales occidentales», o adaptados a los «países prósperos», o incluso de ser «intrusiones ilícitas», entonces la guía tendrá que sobrevivir con tal criticismo»⁵⁸.

Esta reacción de Humana no parece concordar bien con el propósito de alcanzar consenso sobre una metodología de trabajo, y deja la impresión de que su labor ha estado marcada por la intención de hacer la apología de los valores occidentales a cualquier precio. Sin embargo, ni el trabajo de Humana queda retrospectivamente empañado por esta observación ni necesariamente hay que vincular de esta manera una cuestión metodológica con posicionamientos doctrinales. Si se trata de averiguar cuál es la situación fáctica de los derechos humanos en el mundo, habrá que preguntarse, entre otras cosas, por qué la eficacia de tales normas varía en relación con determinadas tradiciones culturales. Pero esta cuestión nos lleva de nuevo al problema del modelo del normativismo jurídico, tratado en el capítulo 1. Humana se ha esforzado para poder garantizar la validez metodológica de su indicador a base de vincularlo lo más estrechamente posible a las declaraciones de derechos de Naciones Unidas. Supone así que el objeto que se trata de evaluar está definido en términos máximamente objetivados, y por ello merecedores de un consenso

⁵⁸ Humana, Charles. *World human rights guide*. Third edition. Oxford University Press. Oxford/New York, 1992. p. 8

incuestionable, que se deriva directamente del carácter universal de las declaraciones de derechos. No parece darse cuenta de que lo cuantificable no es la norma en cuanto tal, sino su eficacia. Pero si de la norma puede predicarse validez (eventualmente universal), no puede hacerse lo mismo con la eficacia.

La constatación de que la Declaración de 1948 se formuló a partir de valores y tradiciones occidentales no puede confundirse con la suposición a priori de que por ser tales valores los únicos legítimos, la diversidad restante puede ser ignorada por completo. Entre el argumento de Humana y el del embajador de Ghana hay una vía intermedia que apunta a la verdadera objeción a los sistemas hasta ahora ensayados de cuantificar los derechos humanos; el multiculturalismo. Si es un hecho que la Declaración de 1948 está marcada por la tradición cultural occidental, no puede dejar de tenerse en cuenta que la sociedad internacional es multicultural, y esto sí que es un hecho cierto. Me refiero a que si en el seno de cada Estado particular el hecho de la diversidad cultural es cada día más patente, cuando nos referimos a normas que tienen por destinatario al género humano, el multiculturalismo debería ser un punto de partida todavía más evidente. Cualquier sistema de evaluación cuantitativa de los derechos humanos debe plantearse cómo abordar metodológicamente que un mismo enunciado normativo experimente sustanciales variaciones en su operatividad determinadas por el contexto cultural en que la norma deba aplicarse. Y esto sin necesidad alguna de cuestionar la validez de la norma. Vamos a analizarlo, en primer lugar, a través de una serie de casos.

Lawrence W. Beer ha hecho un análisis comparativo de las diferencias en la aplicación del derecho fundamental a la libertad de expresión en Estados Unidos y Japón⁵⁹. En su estudio aborda en primer término el derecho vigente, para examinar más adelante cómo el sustrato cultural modula la interpretación jurisprudencial

⁵⁹ BEER, Lawrence W. *Freedom of expression in Japan with comparative reference to the United States*. En *Comparative Human Rights*. Richard P. Claude (ed). The John Hopkins University Press. Baltimore, 1976. pp. 99-127.

e incluso social de tales normas. Merece la pena examinar detenidamente esta relación entre patrones culturales de comportamiento y normas formales porque afecta a la esencia del concepto sociológico del derecho, es decir, a la relación entre el aspecto convencional, el orden legítimo, y el aspecto específicamente jurídico de los modos de organización política del respaldo coactivo a tal orden legítimo. De hecho, la cuestión es central para la sociología de los derechos humanos. La elección de Japón y Estados Unidos por parte de Beer es, en este sentido, muy acertada, porque es un ejemplo muy claro de esta relación entre patrones culturales y normas positivas. La razón es que la Constitución Japonesa de 1947, como es sabido, fue redactada con la «inspiración» de las fuerzas de ocupación estadounidenses, y refleja con claridad el constitucionalismo norteamericano. De modo que en cuanto al derecho positivo se da una sustancial coincidencia, que permite evaluar exclusivamente en términos del factor cultural (o *convencional* en sentido weberiano) las diferencias detectadas en la eficacia de ambos sistemas legales. En otras palabras, si un análisis comparativo tuviese que examinar dos situaciones en que se diesen divergencias tanto en el plano cultural como en el de las normas vigentes, no habría base para asignar fiablemente al factor cultural o al del derecho positivo las diferencias observables. Por el contrario, en el caso estudiado la variable normativa queda despejada y significa, en consecuencia, un ejemplo muy representativo de la importancia de las diferencias culturales a la hora de evaluar la eficacia de las normas de derechos humanos.

La constitución japonesa garantiza un amplio abanico de derechos fundamentales, entre los cuales se encuentra la libertad de reunión, asociación, expresión y de prensa (art. 21). Si bien se establece, por una parte, que los derechos humanos serán el criterio supremo que oriente tanto la legislación como la administración pública, se añade en el texto constitucional que los derechos individuales deben ejercerse para y dentro de los límites del bienestar público (artículos 12, 13, 22 y 29). Como observa Beer...

«...a diferencia del *bienestar general* en la Constitución estadounidense, las referencias al *bienestar público* en la Constitución japonesa son

por lo común algo más que didácticas, y han aportado un criterio judicial para decidir sobre la constitucionalidad de las leyes y reglamentos reguladores de libertades fundamentales desde finales de los años cuarenta»⁶⁰

Esta diferencia no nace de la letra de la Constitución que, como ya se ha comentado, garantiza derechos individuales. La diferencia proviene de la falta de un sentido individualista de las relaciones sociales en la tradición cultural japonesa, lo que por el contrario sí está presente en la sociedad norteamericana. Un concepto clave que impregna dicha tradición cultural es el de *amae*, sustantivo que significa dependencia y cuya forma verbal es *amaeru*. La libertad de expresión está fuertemente limitada por el *amaeru*, según el cual lo decisivo es el sentido de pertenencia al grupo, de modo que, en el seno de los grupos, invocar razones personales para oponerse a lo que consensualmente se ha fijado como objetivo común es visto como un ejercicio injustificable de egoísmo. De hecho, hasta mediados del siglo XIX no existía en japonés una palabra para designar «derecho» (en su acepción de derecho subjetivo, evidentemente). Por el contrario, lo decisivo en la identidad individual es el sentido del deber, y éste se deriva de la situación que uno ocupa en la jerarquía social. *Amae* se refiere a la necesidad individual de sentirse amado, protegido, a recibir indulgencia del grupo. Esta expectativa nace de la reciprocidad con que los otros miembros del grupo ejercen sus deberes para con uno. Por todo ello, el deseo se orienta más a la cálida integración en un sistema de lealtades, deberes y dependencias que hacia la exigencia de autonomía individual (identificada con la pérdida de los vínculos seguros con el grupo). Sociológicamente, este sentido está reforzado además con medidas convencionales de boicot frente a quien pese a todo trata de hacer valer pretensiones individuales. Por supuesto, se trata de medidas no escritas, que coexisten con normas de derecho positivo que garantizan lo contrario. Y a su vez, como afirma Beer, este sustrato cultural influye en las ideas y el comportamiento de

⁶⁰ BEER, Lawrence W. Op. Cit. p. 100 (La traducción es mía)

los responsables de la administración pública y de los jueces, al tiempo que resulta persuasivo para que cada individuo se someta al consenso establecido en sus grupos de referencia.

Por otra parte, y también como consecuencia de la tradición cultural nipona, el papel de la conciliación y arbitraje tiene mucha mayor importancia en Japón que en Estados Unidos. Y no sólo es que haya una mayoría social que opte preferentemente por la conciliación antes que por la resolución judicial de conflictos; es que además existe una presión social en el seno de los grupos para que los individuos resuelvan los conflictos por esta vía. Reuniendo el factor social del *amaeru* y la institución tradicional de la conciliación, preferente frente a la resolución judicial, se explica la relativamente baja litigiosidad en causas relativas a la libertad de expresión, como señala Beer, que emplea como ejemplo al escaso número de demandas que llegan a los tribunales en relación con el derecho de reunión en el caso de manifestaciones de protesta⁶¹. Tal derecho está limitado constitucionalmente en aquellos casos en que las manifestaciones alteren el bienestar público, pero las autoridades locales han tendido en sus ordenanzas hacia una interpretación restrictiva de tal libertad de expresión precisamente para garantizar el orden público, que como se vio anteriormente es algo más que una referencia didáctica de la Constitución y que está reforzado por la tradición cultural japonesa influyente en la actividad de los funcionarios públicos y en los criterios de los tribunales. Es más, incluso antes de cualquier intervención administrativa, y dada la falta de arraigo en la cultura japonesa de la libertad de reunión, los ciudadanos se han sentido obligados en ocasiones a formular peticiones de reunión que eran completamente innecesarias (y Beer cita el ejemplo⁶² de una conferencia sobre nutrición en un domicilio privado, o de una película proyectada a los miembros de un club social en el jardín de uno de ellos). Por más que esté en vigor una constitución democrática, el tradicional sentido

⁶¹ Cfr. BEER, Lawrence W. Op Cit. p. 104

⁶² Cfr. BEER, Lawrence W. Op Cit. pp. 109-10

paternalista del Estado seguía presente en la conciencia de muchos ciudadanos.

No obstante, hay una diferencia de grado entre la conducta espontánea de las personas y la actividad reglada de las autoridades públicas. En este sentido, las ordenanzas que, a nivel local (de *prefecturas*), regulan en Japón reuniones y manifestaciones presentan un espectro variado de limitaciones, optando unas por un sistema de notificación y otras por el de autorización. Este último puede resultar cuestionable, pues restringe la libertad de reunión en función de lo que la autoridad del distrito estime como posibles alteraciones del bienestar público, denegando así discrecionalmente la autorización y restringiendo con ello libertades públicas. Cuando tales ordenanzas fueron sucesivamente recurridas hasta llegar al Tribunal Supremo (que entiende en causas de inconstitucionalidad), éste reiteradamente desestimó las demandas y estableció que una ordenanza es constitucional en tanto regule el lugar y procedimiento bajo criterios claros y razonables encaminados a mantener el orden social y a proteger el bienestar público. Como observa Beer, con este argumento, ni tan siquiera consideró la alternativa notificación/autorización como la raíz del problema, a la vez que no indicó con mayor precisión qué criterios claros y razonables resultaban legítimos y a partir de qué punto vulneraban la constitucionalidad, de modo que dejaba abierto un amplio campo de discrecionalidad para las autoridades administrativas a la hora de denegar manifestaciones y reuniones. Beer muestra además que este argumento reaparece en varias sentencias entre 1954 y 1970, a pesar del voto particular de algunos magistrados que intentaron seguir el ejemplo de los sistemas de notificación de Alemania, Francia, Italia y Estados Unidos, donde el procedimiento de autorización había sido desestimado.

Termina el autor concluyendo que «es más la sociedad que la ley la que determina el status de la libertad de expresión»⁶³, y extrae una consecuencia básica; el análisis de derecho comparado no

⁶³ BEER, Lawrence W. Op. Cit. p. 119

puede ser desarrollado ni comunicado apropiadamente a través de las barreras culturales a menos que integre factores extralegales y trascienda una perspectiva monocultural. En tal sentido, un análisis transcultural debería incluir según el autor cuatro elementos de comparación:...

«...el ordenamiento vigente y la doctrina jurisprudencial, las condiciones sociales generales que tienden a afectar a la libertad de expresión, los patrones de comportamiento y la conciencia social de los derechos en el país bajo estudio y los presupuestos teóricos de la democracia constitucional en cualquier tiempo y lugar»⁶⁴.

William Safran⁶⁵ cita algunos otros ejemplos del diferente significado cultural de los derechos que, sin ánimo de hacer un análisis tan detallado, avalan la misma tesis expuesta por Beer. Su propósito es también poner de manifiesto la diferencia entre derechos constitucionales y tradiciones culturales, de modo que ante normas positivas muy semejantes o idénticas en diversas constituciones, el ejercicio de los derechos de ellas derivados puede resultar completamente distinto en función de factores culturales, revelándose por tanto que el mero análisis comparado de la letra de las constituciones es insuficiente. Así, tanto la Constitución de Estados Unidos como la de Suiza garantizan el derecho a llevar armas, aunque tal derecho no está reconocido como tal en otros países. Pero en Suiza este derecho proviene de la tradición de un ejército popular («el pueblo en armas», en la línea de la Revolución francesa, como una consecuencia de la igualdad y frente a los ejércitos de mercenarios del absolutismo o el privilegio exclusivo de la nobleza en el ejercicio de las armas). Por el contrario, el mismo derecho en Estados Unidos, a pesar de que establece como titular al pueblo (Enmienda 2ª a la Constitución: «...no se atentará contra el derecho del pueblo a llevar armas») es interpretado en sentido individualista, como un vestigio de las necesidades de autodefensa personal de la época colonial, carac-

⁶⁴ BEER, Lawrence W. Op. Cit. p. 119

⁶⁵ SAFRAN, William. Op. Cit. pp. 195-206

terizada por la inexistencia del imperio de la ley en buena parte del territorio. De modo que en Suiza el ejercicio de este derecho es colectivo, está institucionalizado en el ejército regular y tiene el significado de una igual participación en los asuntos públicos, mientras en Estados Unidos es ejercido como parte de la esfera inviolable de la individualidad en la que el Estado no puede interferir. El efecto social de este derecho es, en el caso suizo, un refuerzo de la socialización político-estatal de los ciudadanos, mientras en Estados Unidos el efecto es inverso, y provoca conflictos e inseguridad ciudadana a pesar de que la segunda enmienda justifica el derecho del pueblo a llevar armas en la necesidad de «seguridad de los Estados libres».

Israel también cuenta con un ejército «popular», de reemplazo, aún cuando en su texto constitucional no se reconoce derecho alguno a llevar armas. Los ciudadanos drusos de Israel consideran que su deber de realizar el servicio militar es un motivo de orgullo en la medida en que simboliza su estatuto de ciudadanía en igualdad con los ciudadanos israelíes de origen hebreo. Por el contrario, para muchas mujeres judías ortodoxas este mismo deber es visto como una carga inaceptable, de acuerdo con sus convicciones religiosas. Mientras en Estados Unidos el sistema de las escuelas «segregadas pero iguales» para estudiantes blancos y de color terminó siendo declarado inconstitucional por denegar la igualdad en el sistema educativo, en Israel un sistema educativo integrado para judíos y musulmanes sería considerado inaceptable, y en este caso no se invoca el derecho a la no discriminación en función de la raza o el igual derecho a la educación, sino el derecho de las minorías étnicas, religiosas o lingüísticas mantener su propia identidad cultural, así como el derecho de los padres a elegir la educación que desean para sus hijos (artículo 27 del Pacto Internacional de Derechos Civiles y Políticos y artículo 26.3 de la Declaración Universal de los Derechos Humanos, respectivamente).

La tesis de la diversidad de sentidos culturales en Estados que cuentan con normas constitucionales semejantes o equivalentes (con lo que ello comporta en orden a la diversa operatividad de

tales normas) está avalada por Edmund Cahn⁶⁶ e Ivo Duchacek⁶⁷, desde la perspectiva del derecho comparado el primero y desde la politología el segundo. Ambos autores han llegado a conclusiones muy semejantes en el sentido de que los repertorios de derechos reunidos en las constituciones de los diversos Estados son invariablemente «acusaciones» hechas retrospectivamente contra excesos o deficiencias de regímenes anteriores, y que con objeto de enmendar las experiencias negativas del pasado son incorporados en el texto constitucional para evitar que se repitan en el futuro. De esta manera, la constitución de cada Estado está vinculada a particulares circunstancias de su propia historia que quedan arraigadas a través de la socialización política de los ciudadanos. Así, por ejemplo, la cultura política francesa tiende a identificar el imperio de la ley con la expresión de la voluntad popular (en una tradición que claramente sigue a Rousseau) frente a privilegios particulares. Por eso la centralización administrativa no es vista como un impedimento a la libertad, sino como una garantía de la misma. Y por eso se explica la declaración de inconstitucionalidad del proyecto de estatuto de autonomía para Córcega, pactado entre el gobierno central y los nacionalistas corsos, bajo el argumento empleado por el Consejo Constitucional de que «sólo existe la nación francesa», es decir, que la voluntad popular es una e indivisible y no puede quebrarse por privilegios locales. Un reconocimiento de derechos históricos por parte de una pretendida nacionalidad se presenta como un intento de volver a situaciones quasifeudales. Por el contrario, en muchos Estados federales se verá cualquier intento de fortalecimiento del poder del gobierno federal como un atentado intolerable a las libertades públicas.

Hay otro tipo de tradiciones culturales que también determinan el modelo constitucional elegido. En esta línea, en un análisis cuantitativo de derechos constitucionales que será analizado más

⁶⁶ CAHN, Edmund. *The sense of injustice*. New York University Press, 1949

⁶⁷ DUCHACEK, Ivo A. *Rights and liberties in the world today*. Clio Press, Santa Bárbara (California), 1973.

detalladamente en el próximo capítulo, Boli-Bennett aprecia una variación sustancial en la definición de los derechos de los ciudadanos entre países latinoamericanos, musulmanes de oriente medio, asiáticos y europeos. Y, particularmente, observa que...

«...los países que cuentan con una herencia católica definen los derechos constitucionales de los ciudadanos mucho más extensamente de lo que lo hacen los países con una herencia protestante. El grupo más alto es el de los países latinoamericanos, todos dentro de la tradición católica del derecho natural. Entre los países africanos, aquellos que cuentan con una herencia católica (los colonizados por Francia o Bélgica) definen más extensamente los derechos [fundamentales] que los de herencia protestante (colonizados por Alemania o Gran Bretaña)»⁶⁸.

Sin necesidad de plantear ahora las causas de esta diferente forma de positivación constitucional, algo sobre lo que volveremos más adelante, puede tomarse nota de cómo hay otro elemento cultural que determina las diversas maneras de percibir el significado de los derechos fundamentales. Una cuestión equivalente es planteada por Olusola Ojo⁶⁹ en relación con la Carta Africana de los Derechos Humanos y de los Pueblos de 1981. Teniendo en cuenta que la experiencia colonial es esencialmente determinante del modo de reconocimiento de los derechos en esta Carta, observa que por esta razón el derecho de autodeterminación de los pueblos recibe mayor atención que en otras normas o pactos internacionales, mientras que la protección de derechos individuales está contemplada en términos más laxos al ser vista como una consecuencia de típicos valores de las sociedades occidentales que ejercieron el papel de metrópolis en la colonización.

⁶⁸ BOLI-BENNETT, J. *Human Rights or State expansion? Cross-National definitions of constitutional rights, 1870-1970*. En NANDA, V.P., Scarritt, J.R. y Sheperd, G.W. (eds). *Global Human Rights: Public Policies, Comparative Measures and NGO strategies*. Boulder (Colorado). Westview Press, 1981. p. 182

⁶⁹ OJO, Olusola. *Understanding Human Rights in Africa*. En BERTING, Jan y otros (eds.), *Human Rights in a pluralist world*. Roosevelt Study Center Publications. n° 10. Middelburg (Holanda), 1990. pp. 115-123.

En una línea semejante, Fletcher N. Baldwin⁷⁰ compara el reconocimiento constitucional de los derechos fundamentales en Gran Bretaña y algunas de sus ex-colonias. En Gran Bretaña no existe, en el sentido habitual para los demás Estados, constitución ni un catálogo de derechos fundamentales explícitamente establecidos como tales en un texto jurídico positivo. Los hitos en la historia constitucional de Gran Bretaña, como la *Petition of Rights* de 1628 o el *Bill of Rights* de 1688 no tienen un status legal superior al de otras leyes vigentes. Sin embargo, estas objeciones de «técnica» jurídica no impiden que los derechos fundamentales sean claramente identificables y además respetados. En ello pesa el factor de una tradición poderosamente internalizada y transmitida en la educación y el comportamiento general. En la práctica, esta socialización de los derechos fundamentales impide que el Parlamento intente enmendarlos como si fuesen otra ley cualquiera o que la práctica jurisprudencial ignore su carácter orientativo del ordenamiento. Sin embargo, en Estados como la India, Kenia, Nigeria, Uganda o Malasia, que fueron en su momento colonias británicas y que, por tanto, se inscriben en la tradición jurídica anglosajona, la tendencia ha sido elaborar una enumeración de derechos fundamentales y libertades públicas en las constituciones, y a hacerlo de manera bastante pormenorizada.

Esta tendencia la explica Baldwin porque en aquellos Estados donde el poder judicial ha sido visto por la población como un elemento represivo durante la lucha por la independencia, debe ser contrapesado con limitaciones cuando se alcanza la soberanía. En realidad, se trata del mismo proceso que ya se produjo durante la Revolución francesa. Durante la época de la codificación, la supresión final del artículo I del título I, y del artículo II del título V del proyecto de Código Civil de Portalis, donde se establecía un amplio sistema de fuentes que incluía al mismo derecho natural, se

⁷⁰ BALDWIN, Fletcher N (Jr). *Constitutional limitations of government in México, the United States and Uganda*. En CLAUDE, Richard P. (ed). *Comparative Human Rights*. The John Hopkins University Press. Baltimore (Maryland), 1976. pp. 76-99.

hizo por temor al poder arbitrario de los jueces, ya demostrado durante las monarquías absolutas, poder que no quedaría suficientemente limitado si se les dejaban amplios márgenes para la interpretación de las normas. Por tanto, es lógico que ocurra algo semejante en la descolonización. Pero la conclusión que de aquí se deduce es que la historia política del Estado es determinante para comprender el por qué del modo de proteger los derechos fundamentales. Este tipo de hechos, así como las observaciones de Cahn y Duchacek, vienen a confirmar la idea weberiana de la legitimación afectivo-tradicional que va implicada en el nacionalismo, idea que siguiendo el modelo propuesto en la primera parte, constituye un factor a tener en cuenta en cualquier análisis sociológico de los derechos humanos.

El modo de correlación que se dé en cada Estado entre los patrones culturales y las normas positivas representa, sin embargo, un problema frecuentemente ignorado por los estudios cuantitativos sobre derechos humanos. Pero la solución metodológica que se establezca en relación con este punto resulta crucial. Consideremos, por ejemplo, el caso del informe de Freedom House. Algunos de los elementos de su cuestionario están planteados para averiguar cuáles son las normas positivas vigentes (¿hay leyes electorales equitativas?, ¿tienen los ciudadanos derecho a organizarse en partidos políticos?), mientras otras se preguntan directamente por hechos sociales (¿existe un porcentaje de voto significativo para las opciones de la oposición?, ¿hay asociaciones profesionales libres?). Aquí se produce una mezcla indiscriminada de niveles de análisis que perturba los posibles resultados. Bien podría ocurrir que en un Estado no hubiese impedimento legal alguno para constituir asociaciones profesionales, pero que por la tradición de los modos de organización económica no hubiese habido necesidad de constituir tales organizaciones en la misma proporción que en otros lugares, y que la representación de intereses de los grupos profesionales se canalizase por otra vía. ¿Significaría esto que este Estado tiene un menor nivel de libertades públicas? ¿Deben evaluarse las libertades públicas por lo que la ley permite o por el pluralismo que de hecho exista? En el primer

caso mediríamos la eficacia de normas positivas, en el segundo hechos sociales, y mientras no se discriminen se produce una mezcla acrítica entre el deber ser y el ser. Cuanto más pesen criterios de consideración del segundo tipo, más determinantes resultan los factores culturales, pero en tal caso, al optar por una metodología que evalúa efectos sociales con preferencia sobre las normas formalmente válidas, es preciso definir previamente un sistema de criterios unívoco que sirva para evaluar la situación internacional por encima de la diversidad cultural. Tal criterio ya no será entonces deducible de las normas internacionales sobre derechos humanos, pues no se están evaluando normas. No obstante, ni los informes de Freedom House, ni el de Humana, ni el de Arat, ni el World Handbook of Political and Social Indicators, ni ningún otro parecen tener esta objeción en cuenta.

Y, a la inversa, el problema también se plantea si se adopta como criterio preferente la evaluación de normas de derecho positivo. La literalidad del texto legal no necesariamente es una base fiable para inferir consecuencias. Puede considerarse, por ejemplo, la cuestión SOCFEM del cuestionario de Humana, que evalúa la igualdad social y económica de las mujeres. Si se trata de examinar qué impedimentos legales existen para tal igualdad podríamos comprobar que en China, por causa de la ideología revolucionaria, no hay barreras legales para la promoción social y económica de la mujer. Sin embargo, China tiene una altísima tasa de población rural, que vive en buena parte orientándose por tradiciones culturales muy arraigadas. El cambio legal introducido por la Revolución ha añadido a las mujeres las cargas de la participación igualitaria en la producción industrial, en el ejército o en la educación, pero sin descargarlas de sus responsabilidades tradicionales en el ámbito doméstico rural (porque el Estado no puede cambiar a corto o medio plazo y voluntaristamente tradiciones milenarias transmitidas en el medio familiar). Entre otras cosas, esto explica las altas tasas de infanticidio de niñas, porque sigue vigente la tradición patrilineal y patrilocal en la estructura familiar china. Por otra parte, basta con echar un vistazo a la cúpula dirigente del Partido Comunista y contar el número de mujeres

para comprobar que esa igualdad que la ley proclama es a menudo simplemente formal. Si se examinan los principios legales obtendremos un resultado muy distinto que si examinamos la situación de hecho, muy determinada por factores culturales. Pero, ¿qué tipo de relación puede establecerse entre lo normativo y lo fáctico? Esta cuestión no está ni siquiera planteada en los informes estadísticos examinados.

Por tanto, en relación con los problemas del multiculturalismo en la interpretación de los derechos humanos, al igual que en otros casos, se observa una falta de definición conceptual que oriente con más claridad la investigación cuantitativa. La falta de discusión sobre qué indicadores pueden reflejar apropiadamente aquello que se trata de evaluar indica algo más profundo; la falta de una verdadera reflexión sobre el modelo sociológico de los derechos humanos que debería subyacer a este tipo de trabajos. El análisis de Beer es una buena prueba de ello. Pero un modelo sociológico de este tipo implica también fijar con claridad cuestiones más elementales, como qué tipo de concepto sociológico de derecho se emplea. Por más que haya normas unívocas, en un sentido jurídico, sobre los derechos humanos, resulta decisivo saber qué tipo de estructura política soporta la vigencia mayor o menor del Estado de derecho en cada país. Donde, bajo la capa aparente de un Estado de derecho, subsisten formas de dominación basadas en los clanes (como ocurre en parte, por ejemplo, en Irak) es imposible que haya posibilidades reales de igualdad en la participación política. Y donde subsiste la tradición del *amaeru* en la constitución de todo tipo de agrupaciones sociales es difícil que sea materialmente operativa la libertad de conciencia o de expresión a un nivel equivalente al de países de tradición protestante. Pero estas diferencias culturales deben ser categorizadas en una conceptualización sociológica del derecho, y concretamente de los derechos humanos, que permita el análisis intercultural. Mientras tanto, el multiculturalismo seguirá siendo una objeción insoslayable de los sistemas de verificación empírica de la situación de los derechos humanos. En definitiva, la situación actual de la investigación empírica sobre los derechos humanos no puede evaluarse positiva-

mente más que como una línea de trabajo incipiente. Las deficiencias señaladas en los diversos estudios restan fiabilidad al tipo de sistemas de evaluación hasta ahora ensayados. Como señalan López y Stohl, a pesar de una cierta experiencia de años...

«...es embarazoso pero necesario admitir que se ha hecho un progreso concreto relativamente escaso tanto en la mejora de la discusión transnacional y entre las diversas ideologías sobre los derechos humanos como en el establecimiento de las relaciones entre los derechos humanos y los particulares procesos y estructuras sociales, económicos y políticos»⁷¹.

La escasez de avances, sin embargo, puede imputarse a la falta de fiabilidad de los procedimientos estadísticos, lo que orientaría los intentos dirigidos hacia una futura mejora de estos instrumentos en la línea de un refinamiento de las técnicas empleadas. Pero también puede ser imputable a otra causa, como es la falta de soporte teórico de todas estas investigaciones, lo que llevaría a otro tipo de discusión. A continuación se examinarán propuestas de ambos tipos. Por una parte, los modelos alternativos que algunos autores han avanzado como procedimientos a ensayar en adelante para perfeccionar el sistema de evaluación de la situación fáctica de los derechos humanos en la escena internacional. Pero también se discutirá la posibilidad de que el fallo principal esté en la ausencia de base teórica en todos estos estudios. En cuyo caso habrá que analizar las posibilidades futuras de estudios de este tipo en función de un modelo sociológico de los derechos humanos como el que he propuesto en la primera parte. Este será el objeto del siguiente capítulo.

⁷¹ LÓPEZ, George A. y STOHL, Michael. Op. Cit. p. 217

Investigaciones cuantitativas: alternativas, hipótesis y resultados

5.1 PROPUESTAS METODOLÓGICAS ALTERNATIVAS EN LA INVESTIGACIÓN CUANTITATIVA SOBRE DERECHOS HUMANOS

Tras examinar los principales sistemas de indicadores sobre la situación de los derechos humanos a lo largo de los diversos Estados, y de revisar los criterios de fiabilidad de todo estudio de base cuantitativa sobre la materia, en la práctica puede descartarse la utilidad de los análisis que metodológicamente están orientados sobre todo a la reunión y clasificación de datos para después formular sus resultados en una tabulación en orden decreciente. Es muy poco probable que un estudio de este tipo sea capaz de respaldar con criterios sólidos de fiabilidad aquella apariencia de objetividad científica que el señuelo de los números parece despertar en muchas ocasiones. Por tanto, una adecuada expresión de la situación de los derechos humanos en los diversos Estados deberá intentarse en otras direcciones.

Las que vamos a considerar ahora son fundamentalmente de dos tipos. En primer lugar examinaremos las propuestas de reformulación de los modelos de análisis desarrollados hasta el momento. Su interés no reside únicamente en la mayor o menor capacidad técnica para definir un sistema que pueda superar sin problema los criterios básicos de fiabilidad, sino en la conceptualización de los derechos humanos que definen para la investigación, pues no puede haber estudio alguno, por más que se plantee en términos exclusivamente cuantitativos, que no implique

modelización. En segundo lugar veremos aquellos análisis que, por así decirlo, parten de una inversión de los términos respecto a investigaciones vistas anteriormente: no hacen una labor de búsqueda y ordenación de datos que permitan después deducir de ellos conclusiones generales, siguiendo un [supuesto] esquema inductivo, sino que más bien parten de una hipótesis para cuya comprobación construyen una metodología que permita seleccionar los datos relevantes y exponerlos con sentido. También dentro de este apartado consideraremos modelos teóricos genéricos formulados por especialistas en la investigación cuantitativa como patrón para la redefinición metodológica.

5.1.1. Gupta, Jongman y Schmid

La propuesta de estos autores es un nuevo ejemplo de que depositar toda la confianza en la capacidad de los instrumentos estadísticos puede llevar a un excesivo optimismo, a costa de problemas de conceptualización que están lejos de haber sido resueltos de manera definitiva. Como muestra de ello puede destacarse la convicción de los autores de que el estado actual de la investigación cuantitativa en el campo de los derechos humanos, al haber conseguido unir sólidamente los tres pilares básicos sobre los que se sustenta (es decir, la recogida de datos, la conceptualización y el procesamiento metodológico) está ya en condiciones de alumbrar un sistema de predicción de abusos de los derechos humanos, al que denominan «an early warning system» (algo así como una alerta anticipada). Tal hipotético sistema estaría en condiciones de anticipar situaciones de deterioro de los derechos humanos por medio de inferencias hechas a partir de datos estadísticamente procesados. Supuestamente no hay ya más obstáculos para la puesta en marcha de este sistema que la falta de voluntad política.

«Creemos que ha llegado el momento en que la información, la metodología y los problemas teóricos pueden ser adecuadamente dirigidos a la creación de un sistema de supervisión y de prealerta en el campo de los derechos humanos. El problema ya no es si resulta científicamente posible reconocer con antelación situaciones de deterioro de los derechos

humanos y de estimación de los riesgos crecientes de los grupos vulnerables. Por el contrario, la cuestión es si habrá voluntad política, respaldo financiero y capacidad organizativa para supervisar de manera continua y comprensiva los abusos, y hacer de la prevención en el campo de los derechos humanos una realidad»¹.

En realidad existe bastante disonancia entre este optimismo y la aportación efectiva que los autores hacen al campo de la investigación cuantitativa de los derechos humanos, que se limita a una leve corrección de estudios anteriores, particularmente del de Charles Humana. No obstante, si bien la aportación formal — metodológica— en sí misma tal vez no sea decisiva, sin embargo sí puede tener cierto interés material, por cuanto sirve para destacar una parte de los factores que determinan la eficacia de los derechos humanos. No obstante, veamos en primer lugar el planteamiento general.

Merece la pena resaltar, antes de comenzar propiamente el análisis de la metodología de Gupta, Jongman y Schmid, el tipo de justificación puramente utilitarista que alegan en favor de la investigación cuantitativa. Desde su punto de vista, los derechos humanos han dejado de ser merecedores de atención internacional por simples razones de «compasión», y han comenzado a ser relevantes en función de lo que denominan un «autointerés ilustrado». Los argumentos que invocan son, por ejemplo, que los 20 millones de refugiados crean problemas en los países receptores, que no pueden asumir económicamente este problema, y que además favorece opciones políticas de extrema derecha que explotan sentimientos xenófobos. Por otra parte, el problema afecta no sólo a la política, sino también a la economía internacional. Las teorías neoclásicas sobre el desarrollo tendían a omitir toda referencia a conflictos sociopolíticos en sus modelos, pero poco a poco

¹ GUPTA, Dipak, JONGMAN, Albert J. y SCHMID, Alex P. «Creating a composite index for assessing country performance in the field of human rights: proposal for a new methodology». En *Human Rights Quarterly*, Vol. 16, n° 1 (Febrero de 1994). p. 158.

han llegado a ser conscientes de la importancia de la estabilidad política, y particularmente de la importancia que los abusos sobre los derechos humanos tienen como factor determinante del desarrollo económico. Consecuentemente, cada vez en mayor medida los inversores internacionales tienden a evaluar sus decisiones en función de la situación de respeto de los derechos humanos. Pero tanto los responsables políticos como los económicos necesitan un sistema objetivo de evaluación de la situación de los diferentes Estados para poder tomar decisiones informadas. Esta supuesta correlación entre desarrollo económico y respeto por los derechos humanos es tratada por los autores como un hecho, y el único problema es que sea reconocido como tal por más o menos grupos afectados para obrar en consecuencia, como según su particular opinión ya han hecho los inversores internacionales. Lo que no discuten es si los derechos humanos dejarán de ser merecedores de atención en el supuesto de que llegara a demostrarse que no afectan al desarrollo económico. En cualquier caso, aun cuando no está de más citar esta argumentación justificatoria de la investigación empírica, lo que interesa de forma directa es su propuesta metodológica.

Su aportación puede resumirse en una aplicación de determinadas técnicas estadísticas para mejorar la tabulación llevada a cabo en estudios anteriores, y concretamente en el de Charles Humana, que consideran el más completo. De hecho, los autores no tienen intención de elaborar un nuevo índice del estado de los derechos humanos en la escena internacional, y su estudio sólo sirve por el momento a propósitos ilustrativos, dejando abierta la puerta a una aplicación más elaborada de la metodología que ahora proponen. La primera cuestión que despejan es la de vincular todo estudio cuantitativo a la letra y el espíritu de los tratados internacionales en materia de derechos humanos, por lo que en este punto mantienen idéntico criterio que Humana². Sin embargo, no justifican esta decisión metodológica, que resulta crucial, con ningún argumento.

² Cfr. GUPTA, Dipak, JONGMAN, Albert J. y SCHMID, Alex P. Op. Cit. p. 136.

Sobre esta premisa, los autores citan algunos de los intentos de cuantificación de los derechos humanos, de entre los que seleccionan el de Humana por considerarlo el más fiable técnicamente. Sin embargo, observan que Humana formula un cuestionario de 40 variables, tomadas directamente de la Declaración de los Derechos Humanos de 1948 y de los Pactos Internacionales de 1966, pero que no tiene criterios claros de ponderación entre cada uno de estas 40 variables, a las que no puede atribuirse la misma importancia. El criterio de Humana es asignar un valor triple a algunos elementos que son considerados de mayor importancia al evaluar el nivel de respeto por los derechos humanos (así, por ejemplo, el derecho a no ser sometido a torturas, la existencia de ejecuciones y desapariciones extrajudiciales o de la esclavitud, ya mencionados en el capítulo anterior al examinar la metodología de Humana). Si bien, a juicio de Gupta, Jongman y Schmid este es un paso en la dirección correcta, la selección de tal criterio de ponderación es arbitraria, y no ha sido justificada suficientemente por Humana. Y aducen como prueba el hecho de que en la tabla resultante Japón y Rumanía habían sido calificados con 82 puntos sobre 100, cuando en 1991 ningún conocedor de la situación de los derechos humanos en ambos países los habría situado al mismo nivel.

La raíz del problema, en opinión de los autores, reside en que toda atribución a priori de criterios cuantitativos de ponderación es vulnerable. Para remediarlo proponen el empleo del análisis discriminativo, una técnica estadística que permitiría encontrar un criterio de ponderación objetivo. El análisis discriminativo es empleado cuando los extremos de una escala están bien definidos, pero no así cada uno de los tramos que la componen. En otros términos, la ubicación en una determinada posición de la escala depende de un conjunto n de variables, y es fácil comprobar que la maximización de todas las variables da como resultado la posición más alta en la escala, o la más baja asignando valores mínimos a las variables. El problema se presenta cuando las magnitudes de las variables son intermedias, puesto que cabe suponer fiablemente que no todas tendrán idéntica importancia. Lo que no se sabe entonces es qué peso relativo tiene cada una de estas variables en

el resultado, y por tanto no puede definirse con claridad quién debe estar situado en la posición media de la escala. El análisis discriminativo multivariable identifica aquellos factores que tienen un mayor peso en la determinación de la posición dentro de una escala dada, y por tanto permite evitar resultados paradójicos como el que Humana atribuye a Japón y Rumanía, haciendo el resultado final más objetivo.

Gupta, Jongman y Schmid utilizaron como referencia los sistemas de indicadores de Freedom House y de Stohl, López, Mitchell y Carleton³ referentes al año 1991, y compararon sus resultados para dividir a los Estados en tres grupos: los de nivel alto, medio y bajo de respeto de los derechos humanos. Aquellos Estados en los que ambos sistemas discrepaban en cuanto a su clasificación fueron eliminados por razones de fiabilidad estadística. Reducida la escala a tres posiciones, se aplicó un análisis discriminativo en el que las variables a considerar eran los 40 elementos que componen el cuestionario del sistema de indicadores de Humana, obteniéndose el peso relativo que cada uno de ellos tenía en la determinación de la posición alta, media o baja de la escala establecida con anterioridad.

A continuación se expone la tabla resultante, en la que cada elemento es mencionado por la abreviatura con que lo denomina Humana⁴ y junto a él se expresa el coeficiente que resulta del análisis discriminativo:

³ El primero de ellos ya ha sido mencionado en el anterior capítulo. El segundo, según citan los autores, se trata de STOHL, Michael, LÓPEZ, George, MITCHELL, Christopher y CARLETON, David. *State terrorism: issues of concept and measurement*. En STOHL, Michael, LÓPEZ, George. *Government violence and repression: an agenda for research*. Greenwood. Westport (Connecticut), 1986

⁴ Puesto que ya se detalló en el capítulo anterior el significado de cada una de las abreviaturas con que se denominan los elementos, así como su correspondencia con los artículos de la Declaración universal de los derechos humanos, o los Pactos internacionales de derechos civiles y políticos y de derechos económicos sociales y culturales, y por razón de la brevedad, aquí se añade solamente una referencia sencilla del significado de las abreviaturas.

MURDER (No ejecuciones y desapariciones extrajudiciales)	0,43727
DETENTN (No detención sin cargos)	0,40548
TORTURE (No torturas, tratos crueles, inhumanos o degradantes)	0,39029
NOSEARCH (Garantías judiciales de la investigación policial)	0,38661
POLTCOPP (Derecho a la oposición política pacífica)	0,33307
PRVGUILT (Presunción de inocencia)	0,33042
FREEMAIL (Secreto de las comunicaciones telefónicas y postales)	0,32183
PROPERTY (Propiedad privada sin privaciones arbitrarias)	0,30106
QKTRIAL (Presentación de los detenidos sin demora ante un juez)	0,30101
PUBTRIAL (Proceso público si afecta a derechos o acusaciones penales)	0,29636
TVRADIO (No injerencia estatal en emisoras de radio y televisión)	0,29458
INDCOURT (Independencia de tribunales)	0,29303
FREEPRESS (No censura política de la prensa)	0,27454
ASSEMBLY (Derecho de reunión pacífica y libertad de asociación)	0,27083
BALLOT (Sufragio universal y elecciones periódicas y multipartidistas)	0,26503
MONITOR (Libertad de acción para relatores de derechos humanos)	0,26186
UNION (Libre fundación y afiliación a sindicatos)	0,25564
CAPPUN (Empleo de la pena capital por parte del Estado)	0,25321
FREEINFO (Libre investigación, difusión y recepción de información)	0,24769
BOOK (Libertad para la publicación de libros)	0,24720
PAPER (Libertad para la publicación de periódicos independientes)	0,24363
HOMOSEX (Homosexualidad libre entre adultos que consienten)	0,20322
SLAVLABR (No servidumbre, esclavitud ni trabajo forzado o infantil)	0,19729
MOVEIN (Libre movimiento y residencia en el territorio del Estado)	0,18272
MOVEOUT (Derecho a salir del país y a regresar a él)	0,18080
FREEART (Producción artística sin control o censura estatal)	0,17741
LEGFEM (Derecho de las mujeres a la igualdad legal y política)	0,16580
ANYRELGN (Libertad religiosa)	0,16272
LEGALAID (Derecho a recibir asistencia legal gratuita)	0,16096
NOIDEOL (No educación basada en la religión o ideología estatal oficial)	0,15898
FREEWORK (Libre elección de profesión y protección del desempleo)	0,14674
SOCFEM (Derecho de las mujeres a la igualdad social y económica)	0,14485
EQMARRY (Igualdad de derechos entre cónyuges)	0,12415
ETHEQ (Derecho de las minorías étnicas a la igualdad social y legal)	0,12383
PUNISH (Derecho a no ser condenado a penas corporales)	0,11561
KEEPCIT (Derecho a no ser privado de la nacionalidad y a cambiarla)	0,11408
ETHILANG (Libre uso/educación de la lengua de las minorías étnicas)	0,11385
MIXMARRY (Libre matrimonio interracial o interconfesional)	0,10164
MEMBER (Libertad de toda afiliación obligatoria a partidos estatales)	0,08037
BRTHCONT (Libre uso de píldoras y otros métodos anticonceptivos)	0,03122

Como señalan los autores, el primero de los elementos tiene un peso relativo catorce veces superior al último, y por tanto la asignación por Humana de un valor triple para algunos elementos resulta insuficiente. De igual forma, Humana había seleccionado siete elementos prioritarios a los que asignó dicho valor triple, pero de ellos sólo tres se han mostrado como significativamente más determinantes (MURDER, DETENTN y TORTURE), en tanto que los otros cuatro caen hasta la posición número 18 (CAPPUN), 23 (SLAVLABR), 31 (FREEWORK) y 35 (PUNISH). Se pone así de relieve que las meras estimaciones personales pueden ser una base poco fiable a la hora de establecer la definición, la precisión, la validez y la equivalencia en la metodología estadística.

La aportación metodológica de Gupta, Jongman y Schmid puede representar un avance en ciertos aspectos, pero no deja de abrir nuevas dificultades. Una de ellas es consecuencia de un problema que ya se mencionó al exponer la metodología del sistema de Humana. Vincular cada uno de los elementos de su cuestionario con artículos de la Declaración o de los Pactos internacionales no implica una correspondencia lineal entre ambos. Como ya vimos, Humana elige 7 elementos muy directamente relacionados (PAPER, BOOK, TVRADIO, FREEINFO, FREEPRESS, MONITOR y FREEART) todos derivados del artículo 19 de la Declaración universal de los derechos humanos. Conceder 7 elementos de un total de 40 a un sólo artículo es un criterio de definición que merecería una buena teoría que lo respaldase, lo que Humana no ha hecho. Pero aplicando el criterio estadístico de Gupta, Jongman y Schmid se comprueba que todos estos elementos tienen un peso relativo muy próximo entre sí (salvo en el caso de FREEART, que ya es un indicador discutible en un cuestionario de 40 ítems, están todos dentro de un margen de 0,033 puntos arriba o abajo sobre una media de 0,26158 en el coeficiente). Por lógica, se deduce que no es necesario multiplicar por siete los elementos destinados al artículo 19 de la Declaración cuando los resultados obtenidos en cada uno no difieren sustancialmente respecto de los demás. Si, por el contrario, se mantuviesen en esa proporción, sabemos ahora que el indicador

estaría modificando al alza la puntuación de aquellos Estados en los que se respetase el artículo 19 de la Declaración universal de los derechos humanos. En sentido inverso, aquellos elementos que se han demostrado más determinantes deberían ser acentuados por la vía de desglosarlos en nuevos elementos más específicos.

De esta forma, el análisis discriminativo puede servir para demostrar que la ponderación relativa de cada elemento debe ajustarse, pero también para comprobar que la composición del cuestionario estaba desproporcionada. Siendo así, la solución más evidente apunta a la revisión de la conceptualización inicial, estableciendo nuevos elementos, que sucesivamente deberían ser puestos a prueba. Ahora bien, cada vez que se localicen elementos poco relevantes a los que se les había supuesto alguna virtualidad explicativa será necesario poner en cuestión los argumentos que justificaban el modelo inicial. Y de la misma forma, cada nuevo intento de conceptualización necesitará argumentar por qué se eligen unos elementos y no otros. De otro modo, habría que considerar todos los elementos imaginables para componer el cuestionario, y la prueba de análisis discriminativo se convertiría en un círculo interminable de pruebas y errores. Y esto nos lleva de nuevo a la misma conclusión: no puede haber técnicas cuantitativas sólidas sin hipótesis de trabajo que justifiquen por qué se adopta una metodología u otra. Este es un aspecto de la investigación que, como otros autores, también eluden Gupta, Jongman y Schmid. Ni ellos ni Humana justifican, por ejemplo, por qué no se incluye un elemento relativo al derecho a la seguridad social (artículo 22 de la Declaración universal y artículo 9 del Pacto internacional de derechos económicos, sociales y culturales), ausencia que en la práctica concede prioridad a la posición de los Estados con modelos de desarrollo económico no respaldado por la protección social de los trabajadores, ya que el incumplimiento de este artículo no es computado para evaluar el status del Estado en cuestión. En cambio se incluye un elemento que atiende al derecho a contraer matrimonios interconfesionales o interraciales, que tiene como resultado una rebaja en la posición de los países de tradición islámica, donde está prohibido el matrimonio de la mujer

musulmana con un varón de otra confesión. Este tipo de opciones no dependen de técnicas de análisis discriminativo, sino de otro tipo de razones que convendría no dejar sin responder.

Además de la corrección a la metodología de Humana (y con ella también a sus resultados), hay una segunda conclusión que extraen Gupta, Jongman y Schmid, y es que la aplicación de la técnica del análisis discriminativo despeja toda posible crítica hecha desde el relativismo cultural hacia los sistemas de indicadores. Suponen, en primer lugar, que los derechos humanos deben ser evaluados en función del consenso básico que los sostiene. Según los autores, «a pesar de que las diferencias culturales son importantes, no debería olvidarse que la mayor parte de las culturas tienen patrones comunes»⁵. A continuación sintetizan estos patrones en dos fundamentales: el principio de reciprocidad (debe tratarse a las otras personas como uno mismo desearía ser tratado) y el derecho a la vida y a la integridad física. Estas dos normas son universales puesto que, según los autores, todas las culturas imponen sanciones contra quienes los contravienen. Pero no aportan ninguna prueba ni citan ningún estudio que certifique esta afirmación, cuando más bien las conclusiones de la antropología dirían lo contrario⁶.

Esta tesis, aparentemente, es traída a colación para poder dar soporte teórico a lo que quieren inferir de la aplicación de las mencionadas técnicas estadísticas. Porque si se parte del referente de los tratados y acuerdos internacionales en materia de derechos humanos como premisa, y se añade un consenso «fuerte», teóricamente a salvo de las críticas del relativismo cultural, alrededor de

⁵ GUPTA, Dipak, JONGMAN, Albert J. y SCHMID, Alex P. Op. Cit. p. 135

⁶ De acuerdo con la teoría clásica de Levy-Strauss, la única norma universal, que se da en todas las culturas a pesar de las diferencias de grado, es la prohibición del incesto. Cfr. Levy-Strauss, Claude. *Las estructuras elementales del parentesco*. Paidós. Barcelona, 1988. Por otra parte hay infinidad de culturas que han practicado sacrificios humanos, mutilaciones, raptos, autoinmolaciones y todo tipo de prácticas contra la vida y contra la integridad física que eran considerados positivos tal vez no en sí mismos, pero sí en tanto que medios para obtener fines superiores,

un «núcleo duro» de los derechos humanos, entonces partiríamos de una conceptualización de los mismos que parecería directamente tomada del referente normativo, y que por tanto no requiere ulterior discusión, y tendríamos además un criterio de jerarquización dentro de la misma, justificado implícitamente desde una teoría consensualista cuyo empleo tampoco queda justificado. Ya hemos visto en el capítulo anterior que la diversidad cultural afecta a los problemas de definición de toda metodología cuantitativa con más profundidad de la que queda resuelta con esta propuesta.

A partir de estas premisas, los autores elaboraron tres agregados diferentes a base de reagrupar los elementos del cuestionario de Humana: el índice de integridad vital, el de derechos políticos y el de derechos civiles. El primero se corresponde con los siete elementos priorizados por Humana arriba mencionados. El segundo incluye otros doce, que son ASSEMBLY, MONITOR, POLTCOPP, FREEMAIL, FREEPRESS, TVRADIO, BALLOT, MEMBER, UNION, BOOK, PAPER y KEEPCIT. El tercero contiene los veintiuno restantes. Desglosando los resultados, ya filtrados por el análisis de discriminación, en las tres tablas correspondientes puede comprobarse que no existen grandes diferencias entre ellas. Y de ahí deducen los autores que queda anulada la argumentación del relativismo cultural.

«Los intentos anteriores de construir sistemas agregados de indicadores de los derechos humanos fueron criticados sobre la base de que estos conceptos reflejan prejuicios etnocéntricos, culturales y morales. Nosotros contestamos que mientras tal crítica puede tener alguna relevancia para la medida de los derechos civiles y políticos, pierde su fuerza cuando llega a la valoración de la tortura y otras violaciones graves de los derechos humanos. Sobre esta base, encontramos poca evidencia de relativismo cultural, porque los países situados en el tercio más bajo de la escala general de abusos de los derechos humanos son también aquellos donde se producen las más graves violaciones de tales derechos. En otras palabras, no hay fundamento estadístico para respaldar el argumento principal de los apologetas del relativismo cultural»⁷.

⁷ GUPTA, Dipak, JONGMAN, Albert J. y SCHMID, Alex P. Op. Cit. pp. 155-56

En mi opinión, las críticas del relativismo cultural emplean más argumentos que el que aquí mencionan Gupta, Jongman y Schmid, y consecuentemente no queda tan claro que una prueba estadística como la propuesta sirva para considerarlos superados. En otro orden de cosas, se adhieren también a la hipótesis de que hay una relación directa entre respeto de los derechos humanos y desarrollo económico, y ello a pesar de que, según reconocen, los datos relativos al período 1985-1991 indican más bien que el deterioro de los derechos humanos es peor en *ambos* extremos de la escala del desarrollo económico. Pero resuelven esta contradicción argumentando que por ser éste el período en que se realizaron las transiciones a la democracia del antiguo bloque del Este y de varios países latinoamericanos, «los niveles relativos de dispersión en estos grupos son altos. Por tanto, su diferencia temporal no consigue pasar los tests de fiabilidad estadística»⁸. En realidad, la afirmación anterior sobre la relación entre desarrollo económico y derechos humanos tampoco se fundamenta en investigaciones estadísticas especialmente rigurosas, sino en un simple cruce de datos entre su propia tabulación y los índices de producto interior bruto per cápita del Banco Mundial y los informes sobre desarrollo humano de Naciones Unidas. Volveremos sobre este punto en el apartado dedicado a imagen de los derechos humanos que se desprende de la investigación cuantitativa.

A pesar de que, como los propios Gupta, Jongman y Schmid anuncian en su trabajo, no tratan todavía de construir un indicador de derechos humanos nuevo sino sólo aportar algunas correcciones metodológicas, de su investigación se deducen algunas conclusiones que merecen una reflexión más allá de lo puramente metodológico. No puede dejar de sorprender, por ejemplo, que la posición relativa de los Estados en la escala de respeto de los derechos humanos esté poco relacionada con elementos del cuestionario de Humana como la aplicación de penas corporales por parte de los tribunales, la pertenencia obligatoria a partidos estata-

⁸ Idem. p. 152.

les oficiales o los trabajos forzados. Parece necesario considerar con mayor atención este punto, si bien también esta es una cuestión que pertenece menos a las aportaciones metodológicas que a la definición de los derechos en la investigación cuantitativa. Por tanto, también sobre este asunto volveremos en el tercer apartado.

5.1.2. El informe *Human Rights in Developing Countries*

El anuario *Human Rights in Developing Countries* merece atención en orden a comparar su metodología con las de otros estudios y con las alternativas propuestas. Sin embargo, y como ya se mencionó en el capítulo anterior, no parece tan relevante en cuanto a arrojar datos de interés, puesto que su cobertura es limitada; cada año es examinado un pequeño grupo de países, cuyo número ha ido variando entre 7 y 10, y alterándose también cada año la composición del grupo. No cabe, por tanto, extraer conclusiones generalizables ni hacer seguimientos temporales a partir de la información obtenida. Por otra parte no se hace un tratamiento cuantitativo de la información, que está expuesta en estilo narrativo. No obstante, el debate metodológico sí debe tomar en consideración su aportación en este punto concreto.

Los autores del informe parten de la premisa de que el objeto de estudio debe ser el análisis de las tendencias y los cambios en las políticas gubernamentales referentes a los derechos humanos, más que de los casos individuales de violaciones. Por esta razón consideran que los términos de referencia para evaluar las políticas públicas deben ser la Declaración Universal de 1948 y los Pactos Internacionales de 1966, además de algunos otros instrumentos de Naciones Unidas o sistemas regionales de protección que en la definición metodológica mencionan sin mayor concreción. A partir de aquí, el informe referente a cada Estado se hace siguiendo un esquema dividido en 7 apartados:

1. Sumario. Se trata de una síntesis de los hechos más relevantes acaecidos en el país en los últimos tres años en relación con los derechos civiles, políticos, económicos, sociales y culturales.

2. Posición del gobierno sobre los derechos humanos. Dividido a su vez en tres secciones:

a) Obligaciones internacionales: ratificación de los principales tratados en materia de derechos humanos y reservas interpuestas a los mismos, cumplimiento de la obligación de enviar informes a los comités correspondientes de Naciones Unidas, integración de las obligaciones internacionales en el ordenamiento jurídico interno y posibilidad para los ciudadanos de presentar demandas individuales.

b) Declaraciones básicas del gobierno (manifestaciones programáticas sobre las políticas públicas y respuesta gubernamental a las demandas de investigación de violaciones de los derechos humanos tanto internacionales como nacionales)

c) Existencia de organizaciones (gubernamentales o no) dedicadas a los derechos humanos

3. Sistema de gobierno y derecho a la participación política. A partir de este punto, los autores del informe establecen una relación directa entre las normas internacionales de derechos humanos y los elementos a incluir en el informe. Concretamente exponen que por lo que se refiere a este tercer apartado cubre los artículos 18 a 21 de la Declaración Universal de Derechos Humanos (DUDH), los artículos 18 a 22 y 25 del Pacto Internacional de Derechos Civiles y Políticos (PIDCP), y los artículos 8 y 15 del Pacto Internacional de Derechos Económicos, Sociales y Culturales (PIDESC). Está dividido en 6 partes:

a) Estructura del gobierno y del sistema político. Incluye: el derecho de todos a tomar parte en el gobierno de su país, rasgos generales del sistema político, ideología dominante, tolerancia hacia alternativas ideológicas o de sistemas políticos, papel de los militares y de las fuerzas de seguridad, status legal y fáctico de los partidos políticos y, en los Estados de partido único, garantías de los procedimientos internos democráticos en el partido.

b) Principales conflictos políticos, económicos y sociales

c) Legislación constitucional y ordinaria sobre estados de emergencia (en caso de estar en vigor, cumplimiento de los

requisitos previstos en el artículo 4 del Pacto Internacional de Derechos Civiles y Políticos)

d) Entidades no gubernamentales, grupos de interés o movimientos sociales (distintas a los partidos) con un papel político relevante, así como su grado de aceptación social.

e) Participación política: derecho al voto, elecciones periódicas, genuinas y libres, sistema político pluralista e igual acceso a los cargos públicos.

f) Libertades políticas (respeto a la libertad de pensamiento, de expresión, de información y religiosa y derechos de reunión y manifestación).

4. Derechos civiles. Comprende los artículos 5, 9 y 13 de la DUDH y los artículos 6, 7, 9 a 14 y 26 del PIDCP. Está dividido en tres subapartados.

a) Vida, libertad e integridad de la persona. Dividido a su vez en 5 partes. En la primera se analiza el derecho a la vida básicamente por referencia al uso de la pena capital (si es en aplicación de leyes o no, si cumple o no los requisitos del art. 6 de PIDCP, etc.), las ejecuciones extrajudiciales (investigadas o no, respuesta del gobierno a requerimientos de la ONU, etc.) y otro tipo de amenazas (desastres ecológicos, guerras, etc.). En la segunda, tercera, cuarta y quinta parte se investigan de forma semejante las desapariciones forzadas, la tortura y el trato inhumano, las condiciones de detención o prisión y la libertad personal, respectivamente.

b) Administración de justicia, que incluye cuestiones como la independencia judicial, la existencia o no de tribunales especiales y las diversas garantías procesales.

c) Libertad de movimientos, referida a la libertad de residencia, la existencia de destierros o exilio, y la libertad para entrar y salir del país o circular libremente por él.

5. Derechos económicos y sociales. Las cuestiones de esta parte están tomadas de los artículos 22 a 28 de la DUDH y de los artículos 6 a 13 del PIDESC. Se divide en cuatro partes, cada una de ellas consagrada a un derecho específico:

a) Derecho a un nivel de vida adecuado y a la protección contra el hambre. Se estudia el nivel alimentario general, la composición de la población en niveles de pobreza absoluta (urbana/rural, distribución regional, etc.) y estructura del sistema agrícola (propiedad de la tierra, condiciones del mercado, etc.).

b) Derecho al trabajo bajo condiciones equitativas y satisfactorias. Se analizan niveles de empleo, desempleo y subempleo, tanto agregados como por sectores (mujeres, jóvenes, etc.), políticas públicas de empleo, niveles de discriminación profesional por raza, sexo, opinión política, etc.

c) Derecho al disfrute del más alto nivel posible de salud física y mental. Contempla niveles de salud pública (agregados y por grupos sociales), gasto público en sanidad y medicamentos, índices de mortalidad infantil, acceso a agua corriente, expectativas de vida, prevención de epidemias, minorías con situación sanitaria por debajo de la media, etc.

d) Derecho a la educación. Se analizan las tasas de alfabetización, escolarización en diferentes niveles (elemental, medio y superior) y por grupos (diferentes tasas entre población rural y urbana, o entre varones y mujeres), porcentaje presupuestario en gasto educativo, minorías discriminadas en este aspecto, etc.

6. Igualdad y no discriminación. Si bien esta cuestión aparece tratada a lo largo de los demás apartados, aquí se analiza más específicamente la existencia de patrones estables de desigualdad o discriminación en relación con las mujeres, los niños, los refugiados, los discapacitados, los ancianos, los emigrantes, la población rural o los económicamente desfavorecidos. Está tomado de los artículos 12 de la DUDH, 2, 3 y 26 del PIDCP y 2 del PIDESC.

7. Derechos de los pueblos y las minorías. Referido a los artículos 1 y 27 del PIDCP y al artículo 1 del PIDESC. Se divide en tres apartados: minorías (analizándose el respeto de los derechos lingüísticos, culturales y religiosos de los grupos minoritarios y la política gubernamental hacia ellos), poblaciones indígenas (derechos sobre la tierra, autonomía local y respuesta gubernamental).

mental a sus demandas) y derecho de autodeterminación (grupos demandantes, objeto de la misma y respuesta gubernamental).

En síntesis, las características más destacables de la metodología del informe *Human rights in developing countries* son: vinculación directa a los referentes normativos de los derechos humanos (en la línea de Humana o Gupta), equilibrio entre derechos civiles y políticos y los derechos sociales, económicos y culturales (a diferencia de Humana o Freedom House), y ausencia de criterios de cuantificación y de tratamiento estadístico de la información. Pero la opción por un modelo basado en una descripción cualitativa no implica que los problemas de definición, consistencia y precisión de toda investigación cuantitativa no aparezcan. De hecho, en la metodología utilizada se mezcla indiscriminadamente la existencia de normas con datos puramente fácticos, como el recuento de los casos de tortura con la pregunta por la prohibición, en las leyes del Estado correspondiente, del empleo de la tortura. ¿Deben sumarse unos a otros? ¿Debe valorarse como más grave que se torture de hecho o que exista la prohibición legal de hacerlo? El problema más grave se presenta precisamente en aquello que el informe tiene de positivo por comparación con otros: en el hecho de hacer un cuestionario integrado de derechos civiles, políticos y económicos, sin dejar los últimos al margen como suele ser habitual. Pero puede comprobarse que casi todos los elementos del cuestionario referidos a derechos económicos son datos estadísticos (tasas de alfabetización, de cobertura sanitaria o alimentaria, etc.). Y no se ofrece una clave para referir un tipo de datos al otro. Nos encontramos entonces con los típicos problemas de circularidad: no puede saberse si un buen nivel de respeto a los derechos humanos está en función de la capacidad de gasto público o, a la inversa, si la disponibilidad de fondos públicos depende de la prosperidad que se alcanza al respetar los derechos humanos. Esto no es una crítica si se tiene en cuenta que el estudio está orientado a determinar qué países, de entre un pequeño grupo de Estados en desarrollo, merece prioridad en la ayuda externa. Porque ya se presupone un nivel de desarrollo más o menos equivalente, que opera como variable fija.

Pero la metodología no podría generalizarse para extenderla a un estudio de cobertura universal.

5.1.3. James R. Scarritt

James R. Scarritt⁹, tras analizar la situación actual de la investigación cuantitativa sobre los derechos humanos, concluye que los modelos disponibles hasta el momento resultan incompletos y adolecen de debilidades conceptuales o metodológicas. En consecuencia propone un modelo distinto, si bien la suya es una propuesta que apenas dibuja unas líneas generales. Fundamentalmente se basa en una síntesis de las aportaciones de Flathman y Claude. El primero de ellos no se ocupa directamente de la sociología de los derechos humanos; su trabajo consiste en una definición sociológica general del ejercicio de los derechos, según la cual tal ejercicio es una gama de acciones orientadas normativamente y recurrente en el tiempo.

«Dichas acciones son entendidas por sus partícipes como coincidentes en un grado significativo, e implican interacciones entre los detentadores de los derechos, el contenido de los mismos, personas con obligaciones correlativas a los derechos de los detentadores, personas en posiciones de autoridad y sus obligaciones respecto de estos derechos, y otras personas que participan de las prácticas pero que no ocupan ninguno de los roles antes mencionados»¹⁰.

⁹ SCARRITT, James R. *Definitions, dimensions, data and designs*. En NANDA, Ved P., SCARRITT, James R. y Shepherd, George W. (Jr). *Global Human Rights: public policies, comparative measures and NGOs strategies*. Westview Press. Boulder (Colorado), 1981. pp. 115-123.

¹⁰ FLATHMAN, Richard. *The practice of Rights*. Cambridge University Press, 1976. pp. 1-10 y 161-181. Como puede verse, el concepto sociológico de derecho de Flathman está centrado en el sentido subjetivo de la acción, pero desde un modelo individualista —y centrado en el sujeto— que actualmente resulta difícil sostener. Lo que aquí falta —y esto sería lo objetable desde una perspectiva weberiana— es la determinante estructural de el tipo de dominación política que se dé en cada caso. Basando el derecho en el sentido intersubjetivo de la acción de individuos atomísticamente considerados se ignora la notable diferencia existente, por ejemplo, entre el ejercicio de los derechos en un sistema de

Scarritt considera que esta definición sociológica del ejercicio de los derechos aporta un concepto preciso y empíricamente verificable que puede servir de base para estudios cuantitativos en materia de derechos humanos. Al mismo tiempo reconoce que la metodología de Flathman ignora los derechos socioeconómicos, pero cree que esta deficiencia puede ser salvada con pequeñas modificaciones en el modelo sin disminuir su potencial de utilidad.

Respecto de Claude, Scarritt entiende que puede ser útil su clasificación de los derechos humanos en un esquema clasificatorio que combina la dimensión analítica y la temporal. En el eje horizontal se sitúan las generaciones de derechos: en primer lugar los formulados en el siglo XVII o antes (libertades políticas), después los formulados en el XVIII (libertades civiles), después los del XIX (derechos de participación), y por último los del presente siglo (derechos socioeconómicos). En la columna vertical se sitúan dos categorías básicas: derechos positivos y negativos, que a su vez se subdividen en categorías más precisas. Los primeros incluyen el derecho a la asistencia sanitaria, a la vivienda, al trabajo, etc. Los negativos se dividen en derechos frente al Estado y frente a grupos o personas privadas (aquí, por ejemplo, se encuentran los derechos del menor o la protección del medio ambiente en el siglo XX, la protección contra la servidumbre en el XIX o contra la calumnia y el libelo en el XVIII). Los derechos frente al Estado se dividen a su vez en derechos económicos (no discriminación en el acceso a servicios públicos en el siglo XX, la libertad sindical en el XIX, etc.), de igualdad (trato no discriminatorio en función del sexo, la edad, el nivel educativo o la salud en el siglo XX, trato no discriminatorio en función de la raza, el idioma o el origen étnico o nacional en el XIX, igualdad ante la ley en el XVIII, etc.), derechos políticos y derechos «privados» (según los denomina Claude). En conjunto resulta un

patrimonialismo jurídico o en un sistema de dominación burocratizada. Los derechos de Flathman parecen ejercerse en un espacio atemporal y abstracto en relación con la diversidad de formas de organización política.

cuadro muy detallado que podría entenderse como un desarrollo de la teoría de los derechos públicos subjetivos de Jellinek con sus categorías históricas de derechos¹¹.

La ventaja de esta clasificación está en que evita prejuicios ideológicos en el tratamiento de la información cuantitativa, prejuicios que, según Scarritt, se originan en alguno de los tres grandes modelos de interpretación de los derechos humanos: el individualista, el socialista y el comunitarista¹². La clasificación precisa de los derechos a través de estas categorías representa un esfuerzo por delimitar de la manera más completa el campo de informaciones relevantes y obliga a un análisis global que no se desvirtúe por una orientación interpretativa que resulte incompleta al basarse en un único tipo de derechos.

Sobre la base de las aportaciones de Claude y Flathman, Scarritt argumenta que la perspectiva sociológica desde la que deben ser analizados los derechos humanos es la del cambio social, y particularmente el cambio político. Se deduce entonces, lógicamente, que antes que un sistema de indicadores construido por agregación, debería emplearse un sistema que manejase información sobre sucesos individualizables, seriándolos para detectar patrones de cambio. Se trataría, por tanto, de la medida y explicación del cambio político general. Tal cambio político sería definido con la suficiente amplitud como para abarcar cambios en los papeles políticos, de las estructuras de grupos y de las normas,

¹¹ El esquema completo puede consultarse en CLAUDE, Richard P. *Comparative rights research: some intersections between law and the social sciences*. En CLAUDE, Richard P. (ed.), *Comparative human rights*. The John Hopkins University Press. Baltimore, 1976. p. 392-93.

¹² Ciertamente este es uno de los problemas detectados en muchas fuentes cuantitativas. Como se examinó en el caso de los informes de Freedom House — por citar un ejemplo— el modelo empleado se basa en el Bill of Rights norteamericano, y atiende casi exclusivamente a la categoría de derechos civiles, sin analizar los derechos sociales y económicos. Por tanto el problema existe, otra cuestión es que el modelo de Scarritt, apoyado en las categorías de Claude, suponga una solución suficiente.

incluyendo el contenido y patrón de desarrollo de las políticas públicas. La propuesta de Scarritt no es mucho más precisa. Lo más destacable es el esfuerzo por unificar criterios metodológicos, en lugar de comenzar una metodología completamente nueva. Y sobre todo puede tenerse en cuenta la propuesta de sustituir los sistemas basados en la agregación de múltiples variables en un único indicador final por un esquema de desarrollo histórico conforme a variables catalogadas previamente en un esquema general de evolución. Si bien él mismo no lo propone, podría llevarse a cabo tal investigación buscando patrones de correlación entre diversos elementos del cambio social (incrementos en las tasas de urbanización, de población escolarizada, de lectura de periódicos, etc.) y variaciones en las categorías concretas de derechos humanos tomadas del esquema de Claude.

5.1.4. John F. McCamant

Al igual que Scarritt, McCamant es también un crítico de los sistemas agregados de indicadores en derechos humanos. Parte de la premisa de que el análisis sociológico debe tener en cuenta que en relación con derechos no existen grados: o son o no son respetados. Lo mismo, evidentemente, es aplicable a los derechos humanos. Por tanto, no puede haber ningún sistema que cuantifique grados intermedios de cumplimiento, y en consecuencia los sistemas de indicadores no pueden aplicarse a esta materia. El corolario que se deduce de lo anterior es la diferencia entre la imputabilidad de delitos tipificados a personas concretas, con lo que nos movemos en un plano estrictamente jurídico, y la posible generalización que pudiese hacerse de este esquema a un Estado tomado en su conjunto. La imputación de delitos implica una calificación precisa; es decir, violación de tal o cual derecho, sin que quepan vaguedades. Pero otra cosa es intentar medir el respeto a los derechos humanos en su conjunto, pues ¿cómo se suman detenciones arbitrarias y genocidios? Por otra parte, y como ya se mencionó con anterioridad, la generalización estadística que supone tratar a un Estado como un todo homogéneo puede servir para

esconder la realidad de profundas diferencias presentes en tal Estado. Recuérdese, en este sentido, la observación de López y Stohl mencionada en el anterior capítulo que tomaba como ejemplo los casos de Guatemala y El Salvador. Puede ocurrir que un gobierno democrático sustituya a otro autoritario, pero que determinados órganos del Estado (como cuerpos policiales) hayan desarrollado entre tanto objetivos propios, y que continúen con prácticas contrarias a los derechos humanos de manera encubierta y contra las órdenes del nuevo gobierno. ¿Podría hacerse una media aritmética entre las conductas imputables al gobierno y a estos cuerpos policiales? Por tanto una medición del grado de cumplimiento de los derechos en un Estado es en sí misma contradictoria¹³. La cuestión es distinta cuando no se trata de imputabilidad en sentido jurídico, sino de medidas políticas de presión.

«Una cosa es pedir a un gobierno que ayude a preparar una causa legal contra personas responsables de violaciones de derechos humanos, y otra completamente distinta procurar políticas no judiciales de sanción hacia gobiernos extranjeros. El uso localizado de la ayuda externa, la ruptura de relaciones diplomáticas o embargos comerciales o de inversiones requieren hacer evaluaciones globales de los gobiernos y los países»¹⁴.

¹³ Igualmente resulta pertinente aquí la objeción ya mencionada que se deduce del cambio de criterios estadísticos del Índice de Desarrollo Humano de la ONU: las diferencias internas dentro de un mismo Estado en cuanto al bienestar o al disfrute de los derechos quedan ocultas en las medias estadísticas. Cuando no se evalúan Estados en su conjunto, sino grupos, aparecen datos tan relevantes como que los negros de Estados Unidos tienen el mismo nivel de desarrollo humano que Trinidad Tobago, mientras los blancos ocupan la primera posición mundial. Sin embargo, cuando no se desagregan los grupos, Japón es el primer país en el índice de desarrollo humano y Estados Unidos el segundo. Tanto este caso como el aludido por López y Stohl indican, en la misma línea que McCamant, que la evaluación de la situación de los derechos humanos tomando a los Estados como unidad de referencia puede desfigurar la realidad subyacente. Se trata de un problema básico de conceptualización operativa de la investigación.

¹⁴ McCAMANT, John F. *A critique of present measures of «Human Rights development» and an alternative*. En NANDA, P., SCARRITT, James R. y SHEPHERD, George W (eds), *Global Human Rights: Public policies, comparative measures and NGO strategies*. Westview Press. Boulder (Colorado), 1981, p. 123. (La traducción es mía).

Pero en tal caso, en el análisis global de la situación ya no se trata propiamente de medir el cumplimiento de derechos. Este análisis debe basarse en otra forma de conceptualización de tales situaciones que las diferencie de la imputabilidad que se refiere a individuos. La consecuencia que McCamant extrae es que no es viable un sistema de indicadores del tipo de los que se han intentado. Y propone un cambio sustancial: abandonar el concepto «derechos humanos» y pasar a conceptos analíticamente diferentes. Esto implica, entre otras cosas, una teoría causal sobre la eficacia de las normas relativas a los derechos humanos. Y además una redefinición conceptual del objeto de la investigación. En este sentido, McCamant propone seis líneas básicas a investigar:

1. Igualdad de condiciones económicas y sociales.
2. Provisión de un mínimo de condiciones económicas y sociales.
3. Respeto por la diversidad cultural.
4. Autodeterminación de los grupos.
5. Participación política.
6. Libertad y represión política.

Estos conceptos generales deberán ser desagregadas en elementos más específicos, pero según opina McCamant, con ellos se cubren la mayoría de los fenómenos incluidos en las previsiones normativas de la Declaración Universal de los Derechos Humanos, y pueden ser vistos como diferentes objetivos sociales que los Estados deberían maximizar¹⁵. Cada uno de ellos merece el diseño de un sistema específico de definición, medida, validación, equivalencia y agregación, dado que se trata de conceptos diferentes.

¹⁵ Esta conceptualización de los derechos humanos parece bastante próxima a la teoría de las necesidades de la Escuela de Budapest, o más concretamente de Agnes Heller, y desde luego se encuentra en una posición completamente distinta del liberalismo individualista de conceptualizaciones como la de Freedom House.

McCamant sugiere estas líneas de trabajo, y aporta el ejemplo de su propia experiencia en la investigación de la represión política, que estaría comprendida en el sexto de estos conceptos.

Por represión política entiende McCamant el uso o amenaza de uso de la coerción por parte de las autoridades contra oponentes o potenciales oponentes en orden a prevenir o debilitar su capacidad de oposición a las autoridades y sus políticas. Pero más relevante que la definición es la traducción operativa en criterios cuantificables. Como señala McCamant utilizando el clásico ejemplo, no pueden sumarse peras y manzanas. Del mismo modo, tampoco pueden sumarse la desaparición de un sindicalista no investigada por la policía, la no concesión de licencia a un periódico crítico con el gobierno o la detención sin cargos de los organizadores de un mitin político. Pero sí pueden sumarse en un agregado las proteínas, azúcares, vitaminas, etc. que contienen diferentes frutas. Paralelamente, pueden desagregarse los componentes fácticos de los diferentes derechos, y así ser reunidos en un agregado. Para ellos, McCamant desarrolla tres criterios de cuantificación: la extensión, arbitrariedad y severidad de los actos de represión dirigidos por autoridades políticas.

La extensión se refiere al tipo de actividades reprimidas. Dado que la capacidad de oposición depende de la capacidad de organización y de la capacidad de comunicación, se consideran ambas variables por separado. En ambas McCamant aplica una escala de cuantificación. La arbitrariedad se refiere a la predecibilidad de las medidas coactivas empleadas por las autoridades. Y la severidad indica el grado de las penas impuestas por diferentes comportamientos y el número de personas afectadas de hecho por el uso de sanciones. Es necesario distinguir estos tres aspectos porque la variación en uno de ellos puede perfectamente ser independiente de las variaciones en los otros. Por ejemplo, un Estado en que la ley prohíba toda organización no oficial tendría un nivel alto de extensión. Pero si las leyes son aplicadas de manera imparcial y regular, su grado de arbitrariedad será bajo. Y si los ciudadanos se han habituado a obedecer tales leyes el grado de severidad será también bajo. No obstante, McCamant observa que la conclusión

de sus estudios empíricos aplicando este método mostraba que las tres variables suelen incrementarse de forma paralela¹⁶. McCamant considera útil la experiencia en esta investigación para tomarla como modelo y desarrollar criterios de cuantificación en los otros cinco componentes en que desagrega los derechos humanos. Su contribución más relevante a la metodología cuantitativa es el abandono de los referentes normativos (al contrario de Humana, Gupta o el informe *Human rights in developing countries*) y su propuesta de desagregación de los derechos en un conjunto de referentes fácticos susceptibles de auténtica cuantificación.

5.1.5. Russell Lawrence Barsh

La propuesta metodológica de Barsh está limitada a un objetivo concreto. Su interés se centra en la categoría específica de investigaciones cuantitativas que han intentado mostrar relaciones causales entre desarrollo socioeconómico y respeto de los derechos humanos. Como habremos de ver, no es un tema menor teniendo en cuenta que es la hipótesis de trabajo más reiterada en aquellos estudios que tratan de poner a prueba modelos teóricos de desarrollo de los derechos humanos. Tras analizar las carencias en los criterios estadísticos de fiabilidad de los mencionados estudios, Barsh concluye que es necesario delimitar los objetivos a aquellos que puedan alcanzarse con mínimas garantías. Utiliza además como premisa una vinculación directa de la metodología con los referentes normativos internacionales sobre derechos humanos, pero en un sentido muy distinto al que siguen Humana, Gupta o el informe *Human rights in developing countries*. Según Barsh, la Declaración del derecho al desarrollo de 1986 ha definido a éste como el logro de los derechos humanos, como el incremento de las

¹⁶ Aunque hay ejemplos de lo contrario. Guatemala a partir de 1966 presentaba índices bajos de extensión, pero muy altos en arbitrariedad y severidad. En el otro extremo, la República Federal Alemana presentaba el grado más bajo de arbitrariedad.

capacidades humanas y de las elecciones individuales. «De acuerdo con esta visión emergente, los derechos humanos y el desarrollo son, por definición, la misma cosa»¹⁷. Esta premisa invita a replantear el objetivo de la investigación: el propósito ya no debe ser demostrar que los derechos humanos favorecen el desarrollo económico o viceversa, dado que esta hipótesis se ha vuelto tautológica. La cuestión ahora es determinar la naturaleza de la relación entre el crecimiento (entendido como el mero incremento del producto nacional bruto) y el desarrollo (en el cual se incluye ya el pleno disfrute de los derechos humanos). Conocida la interdependencia de esta relación, las políticas públicas podrán maximizar los beneficios humanos del desarrollo.

Naturalmente, la utilidad mencionada dependerá de la voluntad política de los gobiernos de incrementar ambos factores. Barsh supone que los regímenes autoritarios no van a cambiar sus prácticas por mucho que se muestre que cooperan a un menor nivel de desarrollo. De ello extrae la consecuencia de que la investigación debe centrarse en dos tipos independientes de datos: los que identifican las instancias de represión, y los que incrementan los beneficios humanos del desarrollo. Para los primeros no hace falta una clasificación de todos los Estados del mundo, procedimiento habitual que Barsh propone abandonar. Y para los segundos, la metodología implica desagregar los derechos humanos en función de su papel en el desarrollo humano. Es decir, en lugar de su consideración unitaria como cuerpo normativo, clasificada internamente en derechos civiles, políticos y económicos, se trata de comprenderlos en función de su cooperación al desarrollo, por lo que Barsh propone reclasificarlos en:

a) Recursos humanos: por ejemplo, tasas de alfabetización, indicadores del estado de salud, etc. Se trata de elementos que son a la vez derechos y componentes del valor factorial del trabajo humano.

¹⁷ BARSH, Russell Lawrence. *Measuring Human Rights: problems of methodology and purpose*. Cit. p. 115

b) Seguridad personal: incluye aspectos como libertad frente a intromisiones en la vida privada, la propiedad, el propio idioma y cultura, la elección del lugar de residencia o la libertad para viajar por el territorio y fuera de él, etc.

c) Poder, tanto su ejercicio (participación, pluralismo, sistema de gobierno, etc.) como en las condiciones para su ejercicio efectivo (por ejemplo: libertad de prensa o de opinión)

Como es lógico, no son categorías mutuamente excluyentes: la alfabetización es a la vez un derecho/factor del valor del trabajo y una condición para el ejercicio efectivo del poder, la propiedad puede ser considerada un elemento de la seguridad pero también una forma de poder, etc. Una vez desagregados los derechos de este modo, sus relaciones ya no vienen dadas por el mero hecho de integrar un cuerpo normativo único. El principal objetivo de la investigación será reconstruir la red de relaciones que forman a partir de la información empírica disponible. Y las relaciones particulares entre estos derechos no deben presuponerse simples. El tipo de estructura política puede determinar la forma de redistribución de la propiedad, pero también a la inversa; la distribución dada de la propiedad limita las posibilidades de modificación de la estructura política. Dado que todos los factores varían conjuntamente, el modelo no puede reducirse a una ecuación lineal. Es necesario desarrollar un modelo complejo en que se contemplen todas las dependencias recíprocas entre variables. Claro que esta exigencia tendría como consecuencia que no podría llevarse a cabo una investigación útil hasta tanto no se hubiesen desagregado todos los derechos humanos y se hubiesen establecido mecanismos de cuantificación para cada uno de ellos, de forma que las relaciones de dependencia pudiesen comprobarse en su totalidad. Consciente de ello, Barsh propone comenzar por partes, dirigiendo la atención en primer lugar a aquellos aspectos de los derechos más objetivamente definibles y cuantificables y que, teóricamente, guardan más correlación con el comportamiento económico. Por ejemplo, la posibilidad de ser asesinado o detenido arbitrariamente. Por el contrario, casos como el derecho a ser oído *con justicia* por un tribunal imparcial se presta a valoraciones

subjetivas y a ambigüedad. Sólo sobre la base de la comprobación empírica de los conceptos más operativos podría avanzarse en las tareas de definición y validación de conceptos que por el momento siguen resultando más difíciles.

Hay un par de precisiones que Barsh añade a su modelo. En primer lugar, considera importante incluir variables que reflejen el estado subjetivo de los actores. Para ello cree que son aprovechables los sistemas de encuestas de opinión que reflejan la percepción que los ciudadanos tienen del grado de bienestar general (expectativas de mejora en el empleo, calidad de los servicios públicos, etc.), a los que se añadirían encuestas sobre la percepción de la seguridad y la libertad de los ciudadanos y sobre los mecanismos de participación y de ejercicio del poder. La segunda precisión es que no necesariamente la investigación debe tomar como unidades «naturales» de análisis a los Estados. Por el contrario, el análisis debe ser desarrollado al nivel en que se adoptan las decisiones relevantes: cuando se refiere a la acción de los gobiernos, la unidad de análisis deben ser los gobiernos, pero cuando se trate del interjuego entre el entorno socio-económico y las actividades económicas, las unidades de análisis deben ser los actores económicos: individuos, empresas o comunidades relativamente autosuficientes. Volvemos así a los criterios de análisis desarrollados también en el Informe sobre desarrollo humano de Naciones Unidas ya mencionado, en que los resultados son desagregados en función de los grupos relevantes, y no necesariamente agrupados por Estados. En resumen, siguiendo la línea de Flathman, Scarritt o McCamant, Barsh propone abandonar el referente normativo y redefinir la conceptualización de los derechos a partir de un modelo teórico de desarrollo, que él centra en la intrínseca identidad entre derechos humanos y desarrollo económico integral (que deduce de la Declaración del derecho al desarrollo de Naciones Unidas de 1986).

«El modelo aquí sugerido describiría las interacciones entre derechos y recursos, tratándolos como causas y efectos los unos de los otros, antes que considerar todos los derechos, bien como causa, bien como efecto del crecimiento. Esto requiere un paso preliminar de tipo conceptual: que abandonemos la terminología de “derechos” y reconozcamos

que este concepto confunde un número cualitativamente distinto de cosas que pueden operar diferentemente: factores de producción, recursos de poder, seguridad frente al ejercicio del poder por otros. Requiere también la inclusión de variables que reflejen el estado subjetivo de los actores, y variables que especifiquen de cuánta prosperidad y poder dispone una sociedad en su conjunto, tanto de fuentes internas como externas»¹⁸.

Ciertamente, se trata de una propuesta poco articulada todavía, en que se sugieren líneas de trabajo que todavía adolecen de indefinición. No obstante, su opinión viene a unirse a otras en una línea que cuenta a su favor con la posibilidad de una definición conceptual de los derechos más operativa y fiable en vistas a la realización de investigaciones cuantitativas. Con la aportación de Barsh cerramos el apartado dedicado a las propuestas metodológicas alternativas. Al final del capítulo se discutirá que conclusiones cabe extraer de este debate y qué modelos realmente sólidos quedan en pie.

5.2. PRINCIPALES HIPÓTESIS DE TRABAJO EN LA INVESTIGACIÓN EMPÍRICA

En la categoría de las investigaciones desarrolladas sobre una hipótesis existe también una aportación metodológica, pero diseñada específicamente con el propósito de poner a prueba un modelo teórico. Es necesario precisar que, en realidad, toda investigación parte de hipótesis. La solución adoptada por investigadores como Humana o los autores del informe anual *Human rights in developing countries*, es decir, sustituir la formulación de un modelo general por el criterio normativo, muestra una disparidad de conceptualizaciones de los derechos que indica que tal sustitución no basta. En otras palabras, que la opción por una hipótesis existe siempre, por más que el investigador concreto no

¹⁸ BARSH, Russell Lawrence. Op. Cit. p. 119

sea consciente de ello y presente su trabajo como carente de este tipo de opciones. Pero en tales casos, la hipótesis de trabajo subyacente se deduce después de observar el modo en que se desarrolla la investigación, por lo que siempre es debatible la hipótesis, sea ésta implícita o explícita. Consideraremos ahora las de este último tipo.

5.2.1. Zehra F. Arat

Democracy and Human Rights in Developing Countries es el estudio de Zehra F. Arat, y se centra en las relaciones causales entre desarrollo socioeconómico y derechos humanos en los países en desarrollo, pero para la cual previamente establece su propio sistema de evaluación estadística de la situación de los derechos humanos a escala internacional. A diferencia de los estudios de Freedom House y Humana, el de Arat no es una publicación periódica, pero hace un seguimiento de la situación internacional entre 1948 y 1982, cubriendo 146 Estados, con lo que cumple el requisito de generalidad geográfica y en el tiempo. También es suficientemente extensivo como para cubrir un espectro de países que se aproxima a la totalidad. No obstante, antes de adentrarse en la metodología y los resultados obtenidos por Arat hay que mencionar algunas características de su conceptualización básica. El título de la obra tiene cierta ambivalencia, pues se refiere a la democracia y a los derechos humanos. Esta ambivalencia no es casual. Arat comienza refiriéndose a la clásica diferenciación de los derechos humanos en derechos civiles, derechos políticos y derechos socioeconómicos, para mencionar a continuación la falta de acuerdo sobre el concepto mismo de los derechos humanos en función de la importancia que se le dé a cada uno de estos aspectos. Particularmente alude a la tradición liberal, que separa los derechos civiles y políticos de los socioeconómicos, por los que se muestra un menor interés (como, según se ha visto, es el caso de los informes de Freedom House). Y después de mencionar tres argumentos invocados por diversos autores contra esta separación entre dos grupos de derechos, añade:

«Insistiendo en que los derechos sociales y económicos son también derechos individuales, presento un cuarto argumento que observa la experiencia de los Estados jóvenes con formas democráticas de gobierno: contrariamente a lo que supone la teoría liberal, los derechos civiles y políticos no pueden prevalecer si los derechos socioeconómicos son ignorados, y la estabilidad de la democracia política (democracia liberal) depende de la medida de equilibrio entre los dos grupos de derechos humanos»¹⁹.

De esta introducción cabría deducir que la autora se propone un tratamiento integral de la noción de los derechos humanos, lo que evidentemente implica la traducción a categorías estadísticamente operativas —en el sentido de cuantificables— de un sistema de derechos humanos no truncado (como ocurre en tantas otras investigaciones cuantitativas). Sin embargo, la tarea a la que se dedica Arat después de esta declaración mencionada es a la «definición y medición de la democracia política», lo que nos devuelve a una conceptualización limitada del objeto de estudio. Porque es indudable que el concepto de democracia política es más restringido que el de derechos humanos, y ambas nociones no pueden ser empleadas como equivalentes, por más que tiendan a coincidir los Estados con mayor desarrollo democrático y aquellos que tienen un más alto nivel de respeto por los derechos humanos. En cualquier caso, esta es una suposición que, si se trata de extraer conclusiones generales a partir de estadísticas (como se propone la autora) debería ser probada.

Pasando entonces a la definición de democracia política empleada por Arat, ésta ha sido tomada en buena medida de la tradición de la ciencia política de Anthony Downs y Robert Dahl, y se basa en tres caracteres definitorios de los sistemas democráticos: la inclusividad del proceso, la competitividad y el nivel de derechos civiles.

La inclusividad del proceso se refiere al nivel de restricciones impuestas a la participación democrática en la toma de decisiones

¹⁹ ARAT, Zehra F. *Democracy and human rights in developing countries*. Lynne Rienner Publishers. Boulder (Colorado), 1991. p. 4. (La traducción es mía)

públicas. Tales restricciones pueden estar relacionadas con la raza, el género, la edad, el nivel educativo y la propiedad. En función del número y grado de restricciones impuestas en un Estado se asignan valores entre 0 y 1, siendo 1 el máximo de participación posible (sugragio universal para los ciudadanos mayores de 18 años) y 0 la puntuación cuando no existe sufragio en absoluto. Las restricciones impuestas se estiman en función del porcentaje de población afectada por la misma. Por ejemplo, si el derecho de sufragio se establece a los 21 años se estima una población afectada del 6% (3 años x 2%, según criterios demográficos), lo que bajaría la puntuación a 0,94. Si el voto es derecho exclusivo de los varones, la población afectada es del 50% (estimativamente), lo que bajaría la puntuación a 0,5, etc.

La competitividad se refiere al nivel de posibilidades de opción del electorado. Esta variable se desglosa en dos elementos: la legitimidad y la competitividad de los partidos políticos. En cuanto a la legitimidad, se asignan 0 puntos cuando no hay partidos políticos, o un único partido dominante, 1 punto cuando algunos partidos son permitidos pero subsiste una significativa exclusión de otros, 2 puntos si sólo se excluyen partidos extremistas y 3 puntos si no hay ninguna exclusión. En relación a la competitividad, se valora con 2 puntos el caso de que el partido mayoritario obtenga menos del 70% del sufragio, mientras que si no se celebran elecciones o el partido mayoritario obtiene más del 70% del sufragio se asigna un sólo punto.

El nivel de derechos civiles es medido a través de un indicador indirecto, y se hace así, según explica Arat, a causa de la falta de información fiable sobre el grado en que los gobiernos reconocen y respetan los derechos civiles. Por esta razón, se evalúa el nivel de coercitividad gubernamental, lo que según ella expresaría en proporción inversa el nivel de respeto por los derechos civiles (dejando ahora aparte la cuestionable afirmación de la imposibilidad de obtener información sobre el indicador directo de tales derechos). Además, y dado que la autora entiende que no existen bases teóricas para poder definir cuál es el nivel óptimo de coercitividad (lo que, nuevamente es discutible, pues el tipo de

sistema de dominación política y su grado de burocratización indican precisamente esto) el procedimiento arbitrado para medir la coercitividad es el siguiente: se hace un análisis regresivo del número de sanciones empleadas por el gobierno en relación con los indicadores de disconformidad e inestabilidad social, de modo que a igual nivel de estos indicadores en dos Estados, y si uno emplea un mayor número de medidas represivas, la ratio indicará su mayor coercitividad. Entre los indicadores de disconformidad e inestabilidad social, Arat incluye las muertes por violencia doméstica, los asesinatos, las huelgas generales, la guerrilla, los motines populares, las manifestaciones antigubernamentales y los ataques armados. Por último, los tres factores mencionados son ponderados para obtener el factor de «democraticidad» mediante la fórmula siguiente:

$$\begin{aligned} & \text{Factor de democraticidad=} \\ & = [(Participación \times (1+\text{inclusividad})) + \text{competitividad}] - \text{coercitividad} \end{aligned}$$

Zehra Arat dice haber sometido los resultados obtenidos con la aplicación de su método a determinadas pruebas estadísticas de fiabilidad, como la de McIver y Carmines para escalas aditivas o el sistema alfa de Cronbach basado en la correlación interitems, y en ambos casos los resultados fueron más que satisfactorios. También ha contrastado sus resultados con investigaciones anteriores, y el resultado ha sido igualmente satisfactorio, encontrándose correlaciones altas con los estudios más recientes (que empleaban técnicas más depuradas) y con aquellos que tenían una cobertura más amplia de Estados. Todo ello no obsta para que la propia autora admita las debilidades de su método, como que en su consideración no entran ni la independencia y eficacia judicial ni la libertad de prensa, que los datos sobre coercitividad estatal se basan en la información sobre casos documentados —se supone que por la prensa—, lo que incluye el hecho de que los indicadores más graves de la coercitividad estatal como la detención arbitraria o las desapariciones son excluidos en muchos casos por la sencilla razón de que tal información es muy raramente documentada. Además hay mecanismos coercitivos más subrepticios, como la intimidación que lleva a la autocensura de la prensa, la disuasión

del voto obligando a los ciudadanos a votar en su lugar de nacimiento u otras medidas semejantes que tampoco pueden ser reflejados en el método diseñado por Arat. Al margen de las limitaciones que la propia autora reconoce en su informe hay otras que se considerarán en la parte destinada a las observaciones críticas sobre los sistemas vigentes de evaluación cuantitativa de los derechos humanos.

Pero el objetivo del informe *Democracy and human rights in developing countries* no es únicamente elaborar una tabulación cuantitativa de los derechos humanos a nivel internacional, como es el caso de otros informes, sino que se propone además extraer algunas conclusiones generalizables a partir del tratamiento estadístico de los datos. Realmente esto es lo que hace interesante a este informe, pues son muy pocos los estudios de tipo cuantitativo que se aventuren a formular hipótesis generales a partir de la información reunida (al margen de la fiabilidad de los datos y su procesamiento con la metodología empleada, lo que será discutido más adelante). Como ya se ha mencionado, la hipótesis en que se centra la investigación es la de la correlación entre desarrollo socioeconómico y nivel de respeto por los derechos humanos. Para ellos, Arat utiliza el indicador ya mencionado de democraticidad y lo contrasta con algunas variables, entre ellas diversos indicadores socioeconómicos, para tratar de probar que precisamente esta variable es significativa.

El método que sigue para ello es el del análisis regresivo, una técnica estadística en la cual se fija una variable como referente y se examina en qué medida evoluciona correlativamente con las demás. Por ejemplo, tomando las variaciones anuales del índice de democraticidad en un determinado Estado, puede observarse qué otras variables guardan una mayor equivalencia de variación. Si se observa, pongamos por caso, que existe una variación correlativa en los índices de incremento de la población urbana y los índices de democraticidad, podría inferirse que hay alguna relación significativa entre ambos, mientras que si variando este segundo índice se mantiene constante el de consumo de calorías en la dieta alimenticia, se hace evidente que esta variable mostraría poca

correlación. Un paso más es la técnica del análisis en regresión por pasos, que permite establecer el orden de correlación de las variables, fijando cuál es la más fuerte a la hora de explicar correlaciones, cuál la segunda, etc., y en qué proporción lo hace cada una en relación con las demás. Empleando este método pueden descartarse las variables que no guardan correlación o lo hacen en una medida no significativa, y por tanto pueden ser eliminadas. Con ello se puede trazar un «mapa» de las relaciones causales más significativas en las variaciones temporales de los indicadores cuantitativos. Más adelante se mencionarán otros casos de aplicación de este análisis al tratamiento de datos sobre derechos humanos.

Un primer análisis regresivo muestra que existe una correlación fuerte entre los Estados con un alto nivel de desigualdades socioeconómicas (evaluadas utilizando un coeficiente Gini, que es el sistema habitualmente empleado para medir la distribución de la renta y sus posibles desigualdades), y la decadencia del indicador de democraticidad. En un esquema genérico, el proceso se inicia cuando la desigualdad socioeconómica favorece la inestabilidad (que se traduce en el indicador de disconformidad e inestabilidad social), y la clase dirigente tiende entonces a incrementar la coercitividad política para mantener el status quo. Pero la hipótesis se concreta más, y Arat prueba otras variables alternativas por si pudiesen tener relevancia causal a la hora de explicar las variaciones en el índice de democraticidad. Concretamente analiza el tipo de sistema electoral y de partidos, la estructura constitucional (más presidencialista o más parlamentarista), la diversidad étnica, lingüística o religiosa, la experiencia colonial —en su caso— anterior a la independencia, el papel político de las fuerzas armadas, el grado de consenso o de polarización entre las opciones políticas, el papel del liderazgo político, la cultura política, el nivel de control estatal de la economía y la influencia extranjera. De todos estos factores, los únicos que muestran alguna correlación significativa con el decline o auge de los indicadores del nivel democrático son el papel político de los militares y la influencia extranjera occidental. Todos los demás, si es que tienen algún peso, es en la medida

en que mantienen una correlación muy directa con las condiciones socioeconómicas de la sociedad en cuestión. Por tanto, la variable de la redistribución social de la riqueza sigue siendo la más significativa.

El número y extensión de las tabulaciones con todas estas correlaciones elaboradas por Arat impiden reproducirlas aquí, lo que ocuparía demasiada extensión. No obstante, lo más destacable de su estudio es el intento de probar con metodología estadística la correlación entre desarrollo socioeconómico y derechos humanos. Y no sólo a un nivel general, sino concretando el desarrollo económico en el tipo de distribución más o menos igualitario de la renta, los cambios en fases de desarrollo económico más rápido y las relaciones que estos factores guardan con medios coercitivos empleados por las autoridades políticas como represión contra tendencias hacia la democratización, tanto política como social y económica.

5.2.2. *James C. Strouse y Richard P. Claude*

Una hipótesis muy semejante a la manejada por Arat ha sido empleada también por Strouse y Claude. Aunque con menor extensión y medios que Arat, avanzan una línea de resultados muy semejante. Los autores parten de la afirmación de que la información cuantitativa disponible como base para un tratamiento estadístico que intentase establecer relaciones causales es todavía muy poco fiable. Por tanto, todo análisis extraído de estos datos debe ser considerado únicamente como una tentativa.

No obstante, aunque sea a nivel de tentativa, añaden que siempre es mejor la construcción cautelosa de una hipótesis sobre esta información, que no la especulación en el vacío. Y como ejemplo de lo último citan una afirmación contenida en el informe *Bienestar social y derechos humanos* de la Comisión de Derechos Humanos de las Naciones Unidas, informe presentado con ocasión de la XXIV Conferencia Internacional sobre Bienestar Social de 1968. En él se afirmaba que algunos derechos humanos (como el

de asociación y manifestación pacífica), por supuesto, son independientes de consideraciones socioeconómicas, pero otros (como el de protección frente al desempleo o a la seguridad social) tienen que depender del nivel de desarrollo socioeconómico. Sin embargo, como prueban Strouse y Claude, el derecho de manifestación pacífica o de asociación no es independiente del desarrollo socioeconómico.

La metodología empleada por estos autores es menos completa que la de Arat porque no diseñan su propio sistema de cuantificación de indicadores, sino que utilizan fuentes de información elaboradas por terceros, y su trabajo consiste propiamente en el análisis estadístico de estos datos. Sus fuentes son, para el campo de los indicadores de desarrollo socioeconómico, el *World Handbook of Political and Social Indicators* (ya mencionado), y para los indicadores sobre derechos humanos, el informe anual de Freedom House, también examinado. En una primera aproximación, el análisis sencillo de las correlaciones entre los datos correspondientes a los 122 Estados contemplados ya muestra relaciones significativas entre el sistema de los derechos legalmente reconocidos y los indicadores de desarrollo socioeconómico. Pasando entonces a un segundo nivel de examen, más profundo y basado en técnicas de análisis por regresión, Strouse y Claude llegan a conclusiones más precisas. La tesis a la que se llega con esta investigación es que «cuanto mayor es el nivel de estabilidad política, el desarrollo social y el desarrollo económico per capita, más derechos civiles y políticos adquieren expresión y se vuelven disponibles; pero cuanto más rápida es la tasa de desarrollo económico, menos derechos civiles y políticos adquieren expresión y se hacen disponibles»²⁰. Es decir, que las situaciones que condicionan más negativamente el nivel de respeto por los derechos humanos no son las de aquellos Estados que tienen indicadores

²⁰ STROUSE, James C. y CLAUDE, Richard P. *Empirical comparative rights research: some preliminary tests of development hypotheses*. En CLAUDE, Richard P. (ed). *Comparative Human Rights*. The John Hopkins University Press. Baltimore, 1976. p. 54.

de desarrollo socioeconómico más bajo, sino más bien las de aquellos que experimentan un incremento más rápido de este desarrollo partiendo de situaciones económicas escasamente desarrolladas.

5.2.3. *Richard P. Claude*

Richard P. Claude, además su cooperación con Strouse en la investigación mencionada anteriormente ha construido una hipótesis sobre la génesis de los derechos humanos tratando de sacar conclusiones a partir de investigaciones estadísticas, lo que no resulta demasiado frecuente en el ámbito de la investigación cuantitativa, y él mismo parte de la premisa de que los estudios comparativos sobre la historia político-legal de los derechos y libertades son escasos y poco sistemáticos. El objetivo que se propone es formular las secuencias características de desarrollo de los derechos y libertades, lo que denomina el «modelo clásico de desarrollo de los derechos humanos», y que pretende formular a partir de la experiencia de este desarrollo en diversos países. «El punto de partida del modelo es una sencilla condición de base; la de un sistema legal seguro y procesalmente regularizado»²¹. En este punto, Claude viene a coincidir plenamente con la premisa básica de la teoría sociológica sobre los derechos humanos formulada en el tercer capítulo de este trabajo, y que está basada en la teoría de Weber de la racionalización del derecho. Claude también reconoce la paternidad weberiana de esta conclusión. «Al comparar el desarrollo del derecho romano y del common law, Weber concluyó que la existencia de un sistema legal seguro y avanzado es una condición necesaria para el progreso hacia el desarrollo de los derechos humanos»²². Según su interpretación de la teoría weberiana, una vez que se logra esa condición entran en juego

²¹ CLAUDE, Richard P. *The classical model of human rights development*. En CLAUDE, Richard P (ed.). *Comparative human rights*. Cit. p. 7.

²² Idem. p. 9.

otros factores en la realización de los derechos humanos. Concretamente, ...

«...asumiendo que se da un sistema legal seguro, cuanto más económicamente evolucionado está el sistema social, mayor es la tendencia a desarrollar estructuras legales institucionalmente diferenciadas y procedimientos de conciliación racionales y flexibles. Cuando prevalecen estas condiciones, la constitución de un sistema de derechos humanos se hace posible a través de la legislación, los procedimientos de conciliación regularizados, y/o la jurisprudencia; los instrumentos institucionales para adaptar un sistema legal al incremento del cambio social y político»²³.

Claude se refiere a continuación a una serie de hechos conocidos en la historia del derecho comparado: el viaje de La Fayette a Estados Unidos y su lectura de Jefferson y de las declaraciones de derechos de varios Estados como factor de influencia en la Declaración francesa posterior, la influencia de ésta en Holanda, Bélgica, Austria o Suiza, y en general la semejanza en los catálogos de derechos a lo largo de muchos Estados.

«Sobre la base de que las formulaciones legales reflejan un contenido social y no sólo la imitación retórica, debería llamarse la atención sobre el hecho de que los modelos francés y americano se han seguido en todas partes también en razón de que las condiciones socioeconómicas asociadas con los esfuerzos por crear nuevas estructuras políticas se han extendido igualmente a otras áreas. Por ejemplo, a medida que los soberanos terminaron dependiendo del apoyo financiero de una creciente clase media, la necesidad de un cambio legal se hizo más apremiante. El contexto político del cambio en Francia —el abandono de una monarquía absoluta intolerable— era diferente del de América —la conquista de la independencia respecto a la metrópoli colonial—. Pero los procesos socioeconómicos subyacentes asociados a estos objetivos tenían una base común. Estaba implicado el problema económico crítico de alterar la distribución de bienes y servicios desde una base de status feudal y privilegio en Francia, y de prerrogativas imperiales y exacciones coloniales en América, hasta un mercado con condiciones más libres buscado por la clase media y los terratenientes independientes de ambos países»²⁴.

²³ Idem. p. 10.

²⁴ Idem. p. 23.

El hecho de que Claude haya invocado la autoridad de Weber para poner de relieve la vinculación entre un sistema jurídico racional y la posibilidad de que comiencen a desarrollarse los derechos humanos podría hacer pensar que también sigue a Weber para justificar su argumento del desarrollo capitalista como condición central en la génesis de los derechos humanos, y que explicaría su vigencia en función de aquellos países en que coincidiese este desarrollo capitalista. Como vimos en la primera parte de esta investigación, lo que Weber mostró es que el desarrollo capitalista sólo puede darse en conjunción con el desarrollo del Estado de derecho, articulado sobre un derecho racionalizado, con una burocratización efectiva de los cuadros que ejercen la dominación y la nivelación social que todo ello comporta, además de ciertas condiciones culturales que en el caso occidental se concretaron en la racionalización de las imágenes del mundo que llevó al iusnaturalismo (si bien entendido éste en sentido sociológico). Reducir esta pluralidad de factores al monocausalismo de la racionalización económica está muy lejos de la opinión weberiana, y más cerca de la de Marx.

5.2.4. Eric Sottas

Eric Sottas plantea la relación entre desarrollo económico y derechos humanos a través de un estudio en el que intenta comprobar la vigencia de estos derechos en los países menos desarrollados²⁵. Este grupo de países está definido en función de tres variables; ingresos muy bajos —medidos en términos de renta per capita anual inferior a 100\$ USA con valor de 1968—, una economía con predominio del sector primario —estimada por debajo de un 10 % de producción industrial sobre el Producto Nacional Bruto— y niveles mínimos de capacitación profesional —medido en tasas de alfabetización inferiores al 20 %—. Aunque

²⁵ SOTTAS, Eric. *Development and human rights*. World Organisation Against Torture, 1991.

se seleccionaron estos tres, Sottas menciona algunos otros factores que tienen en común estos países, como tasas de mortalidad infantil altas, expectativas de vida cortas, escaso nivel de calorías en la dieta, falta de acceso a asistencia sanitaria, etc.

Una vez delimitado el objeto de estudio, Sottas afirma que hay suficiente información estadística sobre desarrollo económico y social, pero que «no existe ningún indicador estadístico que pueda ser usado como medida objetiva del nivel de respeto por los derechos humanos en un determinado país», y que «resulta muy laboriosa toda comparación en términos cuantitativos, pues acarrea complejos problemas de metodología»²⁶. En realidad su estudio no tiene demasiado interés si hay que atender a cómo ha resuelto estos problemas, sino únicamente por la correlación entre desarrollo socioeconómico y derechos humanos que plantea. Su propuesta consiste en examinar el nivel de respeto de los derechos humanos en los 42 países menos desarrollados del mundo a base de comprobar si han ratificado o no 18 instrumentos internacionales relativos a los derechos humanos, entre los que se cuentan los Pactos Internacionales de Derechos Económicos, Sociales y Culturales y de Derechos Civiles y Políticos 1966, la Convención contra la Tortura de 1984, sobre el Estatus de los Refugiados de 1951, de Eliminación de Toda Forma de Discriminación Contra la Mujer de 1979, etc.

Naturalmente, el propio autor admite que la ratificación de un tratado internacional no implica que éste sea respetado, y por tanto las meras ratificaciones no son una prueba. Pero al menos, argumenta, indican que las autoridades públicas del Estado en cuestión admiten esa norma como vinculante, lo que puede ser un indicio únicamente del mínimo de respeto exigible, pero al menos es un indicio observable. Para subsanar el escaso valor probatorio de su método, Sottas lo completa con los datos de los informes anuales de Amnistía Internacional y del Departamento de Estado de EE.UU. El resultado es que, por término medio, cada convenio o

²⁶ SOTTAS, Eric. Op. Cit. p. 28.

tratado internacional ha sido ratificado por un tercio de los países menos desarrollados, frente a un 50% aproximadamente del resto de países²⁷. En la práctica, la metodología de Sottas resulta demasiado simple como para probar ninguna correlación entre desarrollo económico y derechos humanos. No compara las variables económicas con otro tipo de factores (como sí hace Arat) para descartar que los resultados pudieran deberse a otra causa. Además, el estado de las ratificaciones, como ya se ha mencionado, no demuestra que exista un nivel constante de respeto de los derechos humanos, acompañado de procedimientos jurídicos con garantías en manos de los ciudadanos. Sencillamente pueden concluirse dos cosas; que los resultados de Sottas podrían ser útiles como un leve indicio de la dirección en la que habría que buscar más información, y que a la hora de formular hipótesis resulta fácil que la primera en ser mencionada sea la del desarrollo económico.

5.2.5. John Boli-Bennett

La investigación de Boli-Bennett tiene un mérito destacable; es una de las pocas que se plantea un modelo teórico argumentado y que a la vez está apoyado con un análisis sobre hechos, análisis que ha podido llevar a cabo gracias a una metodología planteada a tenor de su hipótesis teórica. De esta forma, y a diferencia de otros estudios, no llega con instrumentos estadísticos sacados del fondo común de la metodología cuantitativa y aparentemente aplicables a cualquier tema, sino que opera simultáneamente desde el plano de la formulación de hipótesis y el del diseño de un sistema de verificación.

Su modelo parte de la observación del incremento correlativo entre la capacidad de control social en manos del Estado y la formulación constitucional de derechos fundamentales. Puede

²⁷ Esto es una síntesis de los resultados de Sottas, expuestos en relación con cada uno de los 18 instrumentos internacionales, pero exponer aquí el listado completo sería tedioso y no serviría para añadir más información.

suponerse, en hipótesis, que en el fondo no se trata de dos procesos, sino de un único patrón de evolución que da lugar a manifestaciones tanto en el control estatal como en los derechos fundamentales constitucionalmente reconocidos. A su vez, este proceso de desarrollo único genera dos corrientes ideológicas enfrentadas (y, como se deduce obviamente, si la hipótesis inicial es verdadera, el enfrentamiento sólo es tal en apariencia). Por una parte se forma la ideología individualista, que refleja el debilitamiento de la cultura tradicional a medida que la industrialización y la urbanización desagregan a los individuos, disolviendo los grupos sociales tradicionales y generando una búsqueda de sentido e identidad subjetivas. En oposición —aparente— a esta ideología se desarrolla la del Estado como principal detentador del poder y autoridad sociales y como la principal fuente de identidad colectiva. Pero el proceso único que subyace a estas dos formas ideológicas de representarlo termina integrando ambos postulados, al trasladar de manera constante las aspiraciones expresivas individualistas reflejadas en los derechos humanos hacia el campo de la acción y el control estatal, a base de convertirlas en derechos de los ciudadanos.

De ahí el proceso paradójico según el cual la imposición de límites al poder estatal únicamente pueda realizarse concediendo poder y autoridad al Estado hasta que, finalmente, se convierte en la única agencia hacia la que pueden volverse los individuos para reclamar la protección de sus expectativas. Así, por ejemplo, los movimientos sociales que reclamaron en el pasado el reconocimiento del derecho al voto llevaban implícita la admisión del monopolio estatal para la toma de decisiones en asuntos públicos, y por tanto la legitimación de las autoridades estatales para resolver aquellos ámbitos que ponían bajo su competencia. Es decir, que la extensión de los derechos políticos significa el cambio desde las anteriores fuentes de lealtad de los individuos (familia, gremio, municipio, etc.) hacia el Estado como nueva fuente de la identidad colectiva²⁸.

²⁸ En este punto Boli-Bennett se adhiere a las tesis de Reinhard Bendix (*Nation-Building and Citizenship*. John Willey and Sons. New York, 1964) y

Para poner a prueba su hipótesis, Boli-Bennett analizó todas las constituciones vigentes en el mundo entre 1870 y 1970, y sintetizó los derechos reconocidos en ellas en un listado que incluía 42 tipos distintos de derechos, divididos en cuatro grupos: civiles, procesales, políticos y económico-sociales. Para cada concepto concreto se asigna un criterio de cuantificación ordinal con dos o tres niveles. El apartado de derechos civiles, por ejemplo, incluye:

- Derecho de reunión pacífica: a) inexistente, b) sólo para personas desarmadas y c) sin restricciones)
- Libertad de asociación: a) inexistente, b) en términos generales, y c) incluyendo el derecho específico a la fundación de partidos políticos, sindicatos, etc.
- La libertad de conciencia: a) inexistente, b) en términos generales, y c) incluyendo objeción al servicio militar
- En relación con los hijos: a) ningún derecho, b) derecho a cuidar de ellos y a la adopción y c) derecho a elegir la educación
- Libertad de movimientos: a) inexistente, b) dentro de los límites del Estado y c) hacia el extranjero
- Inviolabilidad de la persona: a) inexistente, b) reclusión sólo según lo establecido por la ley, y c) prohibición de la esclavitud y la servidumbre
- Derecho a la propiedad privada: a) inexistente, b) derecho a efectos personales y a la herencia y c) derecho a la propiedad de los medios de producción
- Derecho a la autodefensa: a) inexistente, b) en términos generales, y c) derecho a poseer y usar armas

El resto de derechos y libertades civiles incluidos en la lista son enunciados con dos posibilidades únicamente: reconocidos o no. Se incluyen:

- Abolición de la esclavitud
- Abolición del exilio
- Inviolabilidad de las comunicaciones y del domicilio

Richard Rubinson (*The political construction of educational systems*. Tesis doctoral defendida en la Universidad de Stanford, 1974).

- Libertad de prensa
- Libertad para contraer matrimonio
- Derecho a la huelga y al cierre patronal
- Libertad de expresión
- Libertad de enseñanza
- Imposición de cargas fiscales en función de la capacidad de renta

El orden en la lista no implica jerarquía. Como puede verse, la tipología de los derechos civiles no representa una novedad digna de atención frente a otros sistemas de codificación. Los derechos procesales, los políticos y los económico-sociales están también definidos en los términos clásicos de la Declaración de 1948 o de los Pactos Internacionales de 1966, por lo que no hace falta reproducir la lista completa.

Más innovadoras resultan las otras tres codificaciones. En la primera, Boli-Bennett reunió los deberes de los ciudadanos incluidos de manera habitual en las constituciones examinadas, y que sintetizó en los siguientes 10 apartados:

- Criar y educar a los hijos
- Participar en la defensa de la nación
- Educarse (escolarización obligatoria)
- Llevar consigo un documento de identidad
- Lealtad al Estado
- Participación ciudadana
- Pago de impuestos
- Respeto a la propiedad pública
- Respeto a las demás personas
- Trabajo

En la siguiente lista, el autor incluyó las áreas de la vida social sobre las que las constituciones atribuían competencia al Estado de manera explícita. La lista contiene 56 elementos, y su enumeración tiene el interés de que expone en términos legales (y constitucionales) las áreas de expansión de la dominación burocrática pública.

- Agricultura
- Ahorro
- Almacenamiento y mercancías
- Arte y artesanía
- Asociaciones
- Banca y crédito
- Capacitación profesional
- Catástrofes naturales
- Caza y Pesca
- Cementerios
- Ciencia
- Colonias
- Comercio exterior
- Conflictos laborales
- Conservación del patrimonio
- Control de armas y municiones
- Control de las rentas
- Control del mercado laboral
- Cooperativas
- Correos
- Crecimiento de la población
- Cultura y deportes
- Declaración de guerra
- Descanso y tiempo libre
- Drogas y alcohol
- Economía
- Educación
- Energía
- Estadística y censo
- Fuerzas armadas
- Fundaciones
- Herencia cultural
- Hoteles y restaurantes
- Juego y casinos
- Maternidad e infancia (protección)
- Matrimonio y familia
- Meteorología
- Minería
- Normas técnicas
- Prensa y cine
- Precios
- Profesiones
- Radio y televisión
- Recursos naturales
- Relaciones terrateniente/arrendatario
- Religión
- Salarios
- Salud
- Seguros
- Sociedades mercantiles
- Teléfono y telégrafo
- Topografía
- Transportes y puertos
- Turismo
- Pesos y medidas
- Vivienda

La última tabla comprende los mecanismos de que dispone el Estado para incrementar su capacidad de control sobre las relaciones sociales, compuesto por 45 elementos. Incluye un repertorio clásico de las competencias de las diversas autoridades estatales, como la elaboración del presupuesto, la concesión de indultos y amnistías, la legislación, la policía, el control de las posiciones económicas dominantes, la censura, la concesión de extradiciones, las nacionalizaciones, la inversión pública, la recaudación de impuestos, la regulación de los derechos de los extranjeros residentes, la inmigración, etc. Los resultados que cada Estado obtiene en función de las variables de derechos civiles, procesales, políticos y económico-sociales son agregados para obtener lo que Boli-Bennett llama el índice de derechos de los ciudadanos. A los

resultados de la lista de deberes constitucionalmente establecidos los denomina índice de deberes de los ciudadanos. Y la síntesis de los resultados de las dos últimas tablas, la de áreas sociales puestas bajo tutela estatal y la de mecanismos de incremento del control estatal se agregan en el índice de jurisdicción estatal. Es a partir de esta elaboración donde Boli-Bennett comienza su análisis de las variaciones que se han ido produciendo en el último siglo

La asignación de valores numéricos a cada concepto es relativamente sencilla. Por ejemplo, si el derecho de reunión no es mencionado en la constitución, se le asignan 0 puntos. Si es mencionado pero los ciudadanos deben estar desarmados, 1 punto. Y si no hay restricciones, 2 puntos. De todos modos, en realidad este es un trabajo de derecho constitucional comparado, apoyado en el examen de textos constitucionales directamente contrastables. Es decir, que podría sustituirse el criterio de cuantificación por una enumeración pormenorizada de cada caso, con el único problema de la extensión que implicaría. Por tanto, las expresiones numéricas tienen aquí un mero valor utilitario de simplificar tramos de variables, y la cuantificación es una traducción muy directa de hechos, no una forzada interpretación cuantitativa de conceptos de difícil valoración numérica, como si se intentase cuantificar la seguridad jurídica en abstracto. Sobre esta base, Boli-Bennett estudió las variaciones producidas en esos derechos en secuencias de intervalos de 20 años. El resultado obtenido está resumido en la siguiente tabla:

	1870	1890	1910	1930	1950	1970
IDC	16,5	17,0	17,3	19,8	24,6	27,9
IDbC	1,4	1,92	2,07	2,31	3,92	3,82
IJE	25,6	26,7	29,1	35,8	46,2	49,4

IDC expresa el índice de derechos de los ciudadanos, IDbC el índice de deberes e IJE el de jurisdicción estatal. Los valores numéricos expresan la suma de los valores medios de cada uno de los derechos o competencias incluidos en el índice correspondiente.

La variable que proporcionalmente más se incrementa, como puede comprobarse, es la de los deberes de los ciudadanos (un 172%), si bien su valor total es escaso (siete y trece veces menor que el de los otros dos). El índice de jurisdicción estatal se incrementa en un 93% a lo largo del siglo, y el índice de derechos un 69%. No sólo existe entre estas dos variables un incremento bastante próximo, sino que el ritmo de incremento es también muy paralelo. Ahora bien, si se desglosa el índice de derechos en derechos civiles, políticos y socioeconómicos, los resultados merecen atención. El último de los tres se incrementa en un 338%, mientras los derechos civiles lo hacen en un 43% y los derechos políticos en un 127%. Como señala Boli-Bennett comentando estos datos, «el tipo de derechos que más claramente protege a los individuos del Estado es el que menos se expande, en tanto que el tipo que más claramente incorpora a los individuos en los programas estatales es el de mayor expansión»²⁹.

La conclusión evidente a la que llega Boli-Bennett es que los derechos constitucionales de los ciudadanos están más ampliamente reconocidos allí donde la autoridad estatal se amplía para incluir a los individuos en sus programas de acción y, paralelamente, existe una tendencia a definir los derechos como alienables o sujetos a restricciones por parte del Estado.

«La ideología de los derechos de los ciudadanos ayuda a los Estados a expandir su legitimidad en la sociedad y a implicar a los ciudadanos en los programas estatales, tendiendo a reemplazar una forma de subordinación del individuo (coacciones y hábitos tradicionales) por otra (control por el Estado)»³⁰.

A partir de los años 30, cuando se produjo una expansión de la capacidad de control del Estado en todo el mundo, se generalizó también el reconocimiento de los derechos sociales. Y, en términos más generales, no es casual que la mayor expansión de la ideología de los derechos humanos se haya producido en el periodo

²⁹ BOLI-BENNETT, John. Op. Cit. p. 179

³⁰ Idem. p. 182

de la historia en que el poder del Estado se ha incrementado más rápidamente en todas partes³¹. Otro dato destacable de la investigación es que la definición de los derechos como inalienables se ha ido reduciendo con el tiempo. En 1870 el 38% de las constituciones asignaba a los derechos ese carácter, en tanto que en 1970 el número de constituciones que garantizaban la inalienabilidad había bajado al 14%. En síntesis, Boli-Bennett ha conseguido mostrar de forma patente la correlación entre burocratización del sistema político y pretensiones de legitimidad basadas en los derechos humanos. Es importante advertir que con esta hipótesis se consigue dar salida a uno de los problemas básicos de toda teoría sociológica que ponga en relación cambio social y derechos humanos. Se trata de las compatibilidades entre derechos civiles y políticos, por una parte, y derechos económicos, sociales y culturales por otra. No es infrecuente atender exclusivamente a los primeros, ni considerarlos como dos conjuntos independientes de derechos, o directamente ignorar los segundos, como vimos que hacía Luhmann en su obra sobre los derechos fundamentales. Pero la exposición de Boli-Bennett consigue mostrar el desarrollo de los derechos como una línea continua que atraviesa la primera y la segunda generación de derechos humanos sin saltos abruptos. Efectivamente, cuando consideramos el reconocimiento de derechos como el modo que tiene el sistema político de implicar a los ciudadanos en sus crecientes programas de acción tenemos una explicación común, que luego se especifica en función del programa de acción de que se trate. El monopolio de la violencia legítima ejercido por un cuadro administrativo burocratizado necesita vincular a los ciudadanos como titulares de los derechos civiles, especialmente el derecho a la vida y a la seguridad. La burocratización, tanto estatal como económica, en la medida en que supone que los individuos serán capaces de integrarse en complejos entramados de acción que implican por parte de los individuos niveles considerables de autocontrol, tiene que recono-

³¹ Cfr. BOLI-BENNETT, John. Op. Cit. p. 176

cer a tales individuos esta capacidad de autonomía funcional, de la que deriva el reconocimiento de los demás derechos civiles. Igualmente, el sistema político burocratizado requiere implicar a los ciudadanos por medio de los mecanismos de participación política, y así garantizar la legitimidad del control jurídico-normativo monopolizado por el Estado. Finalmente, el mantenimiento del control social en manos del Estado a partir de un determinado momento requiere un incremento de la capacidad burocrática de éste en el ámbito económico, justificada ante los ciudadanos por medio del reconocimiento de los derechos económicos, sociales y culturales.

5.3. RESULTADOS: LA IMAGEN DE LOS DERECHOS HUMANOS EN LA INVESTIGACIÓN CUANTITATIVA

La primera conclusión que convendría establecer de partida es que el estado actual de la investigación empírica sobre los derechos humanos no permite extrapolar resultados para llegar a conclusiones suficientemente fiables. No hay todavía un campo acotado de manera precisa, con métodos y criterios básicos consensuados, y esto conduce a que en lugar de sociología de los derechos humanos casi sería más propio hablar de politólogos que estudian la generalización de los procedimientos democráticos, juristas dedicados al derecho constitucional comparado con especial atención hacia los derechos fundamentales, economistas que elaboran indicadores de desarrollo en los que incluyen variables relacionadas con las libertades públicas o expertos en estadística que experimentan formas diversas de procesamiento de la información cuantitativa sobre todas estas materias. Para unificar esta amalgama de perspectivas sería necesario un marco general explicativo de los derechos humanos, donde se describiesen los procesos de cambio social en que se originan y las condiciones específicas que los hacen posibles o los vuelven ineficaces, y al cual pudiesen incor-

porarse los diversos elementos que hoy forman, dispersamente, el campo de la sociología de los derechos humanos. Es la misma conclusión a la que llega Richard Claude:

«El estudio de la “inhumanidad del hombre hacia el hombre”, la guerra y los abusos sobre los derechos todavía está falto de teorías. Pero la necesidad de aplicación de la destreza a la medición, aunque no esté guiada por una teoría general, se está volviendo crecientemente evidente para quienes tienen capacidad de contribuir de forma útil y práctica a la solución de los problemas de las políticas públicas relativas a los derechos humanos internacionales. La búsqueda de teorías avaladas por datos no ha hecho más que empezar»³².

Esta opinión podría ser partícipe de la herencia intelectual de Thomas S Kuhn, uno de los más influyentes teóricos de la metodología de la ciencia contemporánea, quien refiriéndose a este mismo tipo de problemas comentaba: «he aquí la razón por la cual las mediciones tomadas sin un paradigma conducen tan raramente a una conclusión definitiva»³³. Hay una parte importante de la literatura consagrada a esta materia que trata de resolver el problema de la dispersión metodológica y de puntos de vista a base de intentar afinar las técnicas de investigación cuantitativa, pero de nada sirve medir mejor cuando no se sabe bien qué se está midiendo. Como afirma Barsh³⁴, suponer que todos los problemas a los que se enfrentan los intentos de analizar cuantitativamente los derechos humanos pueden ser resueltos a base de ajustes metodológicos es una falacia. Lo que hace falta es una combinación del esfuerzo por formular paradigmas (siguiendo la terminología de Kuhn) y de la búsqueda de información detallada sobre la generalización de los derechos humanos en el contexto de los procesos de cambio social. Tal ausencia de paradigma es especial-

³² CLAUDE, Richard P. and JABINE, Thomas B. *Exploring human rights issues with statistics*. En CLAUDE, Richard P. and JABINE, Thomas B. *Human rights and statistics. Getting the record straight*. Cit. p. 23

³³ KUHN, Thomas S. *La estructura de las revoluciones científicas*. FCE. Madrid, 1987 (11ª reimp.). p. 210

³⁴ Cfr. BARSH, Russel Lawrence. *Measuring human rights. Problems of methodology and purpose*. Cit. p. 107.

mente relevante en dos aspectos: en primer lugar, en la cuestión preliminar de si las normas pueden ser evaluadas en términos fácticos, y por tanto si los derechos humanos pueden ser objeto de investigaciones empíricas —y particularmente cuantitativas— utilizando para ello como referente directo las normas internacionales sobre los mismos. En segundo lugar está la cuestión del multiculturalismo, pues es obvio que los titulares de los derechos humanos forman una comunidad cruzada de diferencias culturales que determinan las variaciones en el valor que se les atribuye a lo largo del mundo y, en consecuencia, también las variaciones en su eficacia. Pero esto nos lleva a la contradicción de tener que examinar el grado efectivo de vigencia de los derechos humanos utilizando un doble patrón: las normas internacionales sobre los mismos en el plano universal y, en cada lugar, la tradición cultural correspondiente. Resolver esta contradicción no es independiente del tipo de paradigma definido para la investigación, ni por tanto de la cuestión básica sobre si dicho paradigma viene simplemente dado en las normas internacionales sobre los derechos humanos. Volvemos así al modelo del normativismo jurídico, tratado en el capítulo I, pero centrándonos ahora en los problemas que dicho modelo plantea específicamente a la investigación empírica.

Los sistemas normativos no describen hechos, sino cómo debería ser el mundo. Por tanto, no pueden ser evaluados directamente en términos fácticos, ya que es imposible encontrar hechos en una entidad ideal. La pregunta por el grado en que los hechos son consecuencia de las normas jurídicas se resuelve fundamentalmente en el ámbito judicial, dado que la actividad de los jueces consiste precisamente en decidir la correspondencia entre hechos y normas. Pero los jueces no examinan los hechos valiéndose de la metodología de las ciencias sociales, sino de un método que está también normativamente determinado. Efectivamente, el ordenamiento no sólo establece cómo debería ser la acción de los individuos, las autoridades o las organizaciones, sino que incluye además normas sobre el procedimiento que debe seguir un juez para seleccionar qué hechos son relevantes y cuáles deben ser ignorados. Empleando como base las decisiones judiciales sí es

posible realizar valoraciones fácticas de las normas, en forma de estadísticas acerca del número de demandas admitidas a trámite, del incremento de las causas civiles o penales, etc. Pero en este caso hay que tener presentes dos condicionantes: primero, que el objeto de estas estadísticas no son directamente las normas, sino las resoluciones judiciales. Segundo, que la metodología empleada tiene una naturaleza mixta. Por una parte, el análisis cuantitativo se realiza empleando técnicas estadísticas, pero previamente ha sido necesaria una decisión judicial que requiere metodología jurídica. Teniendo en cuenta estas condiciones, sí es posible el análisis fáctico de los sistemas normativos.

No ocurre así, sin embargo, en el caso de las normas internacionales sobre derechos humanos, de las cuales sólo una pequeña parte cuenta con una efectiva tutela judicial³⁵. Falta, por tanto, una administración de justicia unificada, que aplique las normas con un conjunto sistemático de criterios concordantes, y cuyas resoluciones puedan después ser examinadas cuantitativamente. Y esta carencia impide también que pueda establecerse fiablemente el vínculo entre hechos y normas. Supongamos una situación en que el proceso de diferenciación social no haya dado lugar todavía a instituciones políticas y económicas consolidadas, y donde la familia cumpla buena parte de las funciones educativas, políticas o profesionales que después se independizarán de los lazos de parentesco. Sería, por ejemplo, el caso de una sociedad estructurada en clanes, y dentro de ella a un individuo perteneciente a un clan poderoso, capaz de garantizarle suficiente protección de su vida y su integridad física. Consideremos junto a él al ciudadano de un Estado de derecho cuya constitución garantiza el derecho fundamental a la vida y a la integridad física. No puede decirse, sin embargo, que los dos sean casos de aplicación del artículo 3 de la Declaración Universal de los Derechos Humanos. Aunque podamos constatar el *hecho* de que las dos personas disfrutaran de

³⁵ Que se verá reforzada con la acción del Tribunal Penal Internacional, cuya creación se acordó en Roma en Julio de 1998.

seguridad personal, lo que permite diferenciar estas situaciones es el tipo de referente normativo del que los hechos adquieren su significado. Sólo en el segundo caso existe relación entre el hecho y la norma, porque sólo el Estado ha promulgado formalmente una norma y ha instituido una administración de justicia para garantizar que los hechos se correspondan con ella. Si se desdibuja esta diferencia, entonces toda situación en que las personas no sufran atentados contra su vida sería un caso de cumplimiento del artículo 3 de la Declaración Universal, aunque los partícipes en la acción la desconociesen. Otro tanto cabría decir del resto de derechos proclamados en dicha Declaración.

Pero la ausencia de un sistema judicial con competencia y capacidad para hacer cumplir todas las normas internacionales sobre derechos humanos hace que la investigación empírica no pueda emplear como base las resoluciones judiciales. Descartada la vía principal, podría intentarse otra secundaria. La eficacia de los derechos humanos depende de que los Estados conformen su ordenamiento jurídico interno con lo establecido en la Declaración de 1948, y más particularmente a los Pactos Internacionales de 1966, el de Derechos Civiles y Políticos y el de Derechos Económicos, Sociales y Culturales, así como con otras convenciones sectoriales (contra el genocidio, de 1948; contra la discriminación racial, de 1965, etc.). Si existiese verdaderamente esta adaptación, entonces habría otra vía para investigar cuantitativamente los derechos humanos, empleando para ello las resoluciones judiciales internas de cada Estado. Pero esta vía sólo sería posible si las normas internas de todos los Estados y las normas internacionales tuviesen una exacta correspondencia, lo cual a simple vista parece difícil si se tiene en cuenta que cada constitución sigue su propio sistema de positivación de los derechos. Como vimos en el caso de la libertad de expresión en Estados Unidos y Japón, incluso cuando dos constituciones reconocen de manera muy semejante un mismo derecho el resultado puede ser divergente, habiendo un número notablemente inferior de resoluciones judiciales sobre este asunto en Japón que en Estados Unidos. Y no por razón de la norma, sino del contexto cultural en la que ésta se hace efectiva, lo que motiva

que en las expectativas de los ciudadanos este derecho tenga mucha más relevancia en un país que en otro. Häberle insiste en este mismo punto:

«Cualquier Estado constitucional resulta ser un todo individual *sui generis* en el que se desarrolla cada aportación lograda por vía de recepción, es decir, donde lo recibido “debe adaptarse”. Por este motivo un mismo texto constitucional, como podría ser un derecho fundamental concreto, puede ostentar diversos contenidos de un Estado a otro»³⁶.

Ciertamente Häberle se refiere aquí al caso de la recepción de instituciones jurídicas de un país en otro, pero la historia de los derechos humanos antes de su internacionalización se desarrolló precisamente de ese modo, y en cualquier caso la consecuencia es también aplicable al modo en que los Estados incorporan a su ordenamiento las obligaciones contraídas al ratificar alguna de las declaraciones o convenciones mencionados. Lo decisivo es que al integrarse en el sistema constitucional correspondiente, la norma original no permanece inalterada, sino que puede ver modificado su contenido por factores culturales, dado que la Constitución no debe ser entendida exclusivamente en cuanto texto jurídico, sino además como expresión de «cierto estado de desarrollo cultural, como medio de representación cultural de un determinado pueblo»³⁷.

Llegamos así al segundo problema de la investigación empírica, el multiculturalismo, que en este contexto específico no significa la tesis opuesta a la universalidad de los derechos humanos, sino la constatación de que supuestas las normas internacionales sobre derechos humanos, y sin cuestionar su universalidad, pero teniendo presente que no están respaldadas por un sistema de garantías judiciales universales, la vigencia de estas normas en el ordenamiento interno de cada Estado está sujeto a condicionantes

³⁶ HÄBERLE, Peter. *Elementos teóricos de un modelo general de recepción jurídica*. En PÉREZ LUÑO, Antonio Enrique. *Derecho humanos y constitucionalismo ante el tercer milenio*. Marcial Pons, Madrid, 1996. p. 157

³⁷ Idem. p. 153

culturales, lo que en parte explica la diversa efectividad de dichas normas a lo largo del mundo. Como expone Rolando Gaete, tanto la Comisión Europea de Derechos Humanos como el Tribunal Supremo de Estados Unidos han admitido que la interpretación de los derechos humanos (del Convenio Europeo, en el primer caso, y del Bill of Rights en el segundo) puede estar sujeta a variaciones culturales locales. El Tribunal Supremo, en el caso *Miller versus California*, argumentó que las limitaciones que impone la Constitución a los Estados «no varían de comunidad en comunidad, pero esto no significa que haya, o deba o pueda haber patrones nacionales fijos y uniformes»³⁸. Igualmente, la Comisión Europea de Derechos Humanos ha establecido que el lenguaje de los derechos humanos emplea ideas o expresiones que «no tienen un significado exacto y generalmente aceptado, dejando a los Estados un margen de apreciación»³⁹. El caso ejemplar es *Handyside*, en el que el Tribunal Europeo de Derechos Humanos sentenció que la prohibición de una publicación en Inglaterra por atentar contra la moral era conforme al Convenio Europeo, a pesar de que la misma publicación había sido difundida sin problemas en otros países, e incluso en Irlanda del Norte, Gales, Escocia y la Isla de Man. El argumento empleado por el Tribunal fue que las autoridades locales son competentes para apreciar las particulares exigencias de la moral en su territorio⁴⁰. Si bien Gaete cita estos casos para extraer la consecuencia de que los derechos humanos sólo tienen una función retórica, sin necesidad de llegar a tales juicios valorativos y manteniendo la conclusión en términos fácticos, lo cierto es que de estos argumentos se desprende que no existe una única modalidad de ejercicio del derecho de los europeos a la libertad de prensa, sino una expectativa genérica del mismo que luego encontrará un diverso grado de confirmación dependiendo de la estimación que los correspondientes gobiernos hagan de las

³⁸ Citado por GAETE, Rolando. *Human Rights and the limits of critical reason*. Cit. p. 74

³⁹ Citado por GAETE, Rolando. Op. Cit. pp. 73-74.

⁴⁰ Cfr. Idem. P. 74.

tradiciones morales locales. Pero el examen de todos los condicionantes culturales que pueden incidir en las variaciones locales de los derechos humanos hacen inviable cualquier intento de examen empírico de éstos, pues el objeto de la investigación sería inabarcable.

Queda la vía, entonces, de comprobar la correspondencia entre las normas internas y las internacionales sobre derechos a partir de los informes que cada uno de los Estados elabora sobre su propia situación, siguiendo el artículo 40.1 del Pacto Internacional de Derechos Civiles y Políticos de 1966, y los artículos 16 a 19 del Pacto Internacional de Derechos Económicos, Sociales y Culturales⁴¹ por los que los Estados se obligan a remitir dichos informes al Comité de Derechos Humanos y al Consejo Económico y Social, exponiendo las decisiones que han adoptado para dar efecto a los derechos reconocidos en ambos Pactos. Sin embargo, y como también vimos, el cumplimiento de este tipo de obligaciones está lejos de ser satisfactorio, y es difícil que algún Estado se autoinculpe voluntariamente en tales informes.

Todas estas dificultades llevan a autores como Flathman, Scarritt, McCamant o Barsh a proponer que se abandone el referente normativo de los derechos humanos en la investigación empírica y que en su lugar sea redefinida la conceptualización de los derechos a partir de un modelo teórico de desarrollo de los mismos, y en este último aspecto vendrían a coincidir con el intento de Claude de desarrollar un modelo general de los procesos de cambio social en que se originan y consolidan los derechos humanos.

A partir de aquí, los diversos modelos de investigación empírica de los derechos humanos tratan de sortear esta cuestión por

⁴¹ Así como otros artículos equivalentes en los demás convenios o tratados de derechos humanos en que los Estados se obligan a remitir informes específicos sobre la materia correspondiente; eliminación de discriminaciones contra la mujer, derechos de la infancia, etc.

diferentes vías. El informe anual *Human rights in developing countries*, por ejemplo mezcla indiscriminadamente normas con hechos. En su cuestionario se pregunta, entre otras muchas cosas, por la legislación constitucional y ordinaria sobre estados de emergencia y su conformidad con los requisitos previstos en el artículo 4 del Pacto Internacional de Derechos Civiles y Políticos. Es decir, se pregunta directamente si existe una norma vigente o no. Pero cuando se pregunta por el derecho al trabajo bajo condiciones equitativas y satisfactorias, lo que se tiene en cuenta son las tasas de empleo, desempleo y subempleo, tanto agregados como por sectores (mujeres, jóvenes, etc.), o cuando se trata del derecho al disfrute del más alto nivel posible de salud física y mental, la información relevante son los niveles de salud pública (agregados y por grupos sociales), gasto público en sanidad, índices de mortalidad infantil, expectativas de vida, prevención de epidemias, minorías con situación sanitaria por debajo de la media, etc. En este caso lo que se requiere son hechos. Pero no se plantea el problema de cómo comparar después normas y hechos. He aquí un problema capital del marco teórico en que se desarrolla la investigación empírica que muchos estudios desatienden.

Zehra Arat también resuelve este problema a base de una cierta ambigüedad en el uso de los conceptos «democracia» y «derechos humanos», como ya vimos, pero sobre todo a base de una ambigüedad metodológica que probablemente sea más relevante. Arat evalúa el nivel de coercitividad estatal y el de participación, que sintetizados emplea como indicador de democraticidad de un país (indicador que usa como sinónimo de los derechos humanos). De manera inadvertida se han transformado los derechos humanos, que son normas, en un objeto a investigar completamente fáctico. Para ello, aquellos aspectos normativos más difícilmente traducibles a hechos han sido eliminados. Pienso, especialmente, en las garantías procesales, cuya evaluación cuantitativa resulta más difícil, como ya he mencionado. Este problema no debería haberse obviado de forma tan rápida, y es una objeción a la base teórica sobre la que se apoya este estudio que queda sin resolver.

Otro buen ejemplo de esto lo tenemos en los sistemas de indicadores de derechos humanos. Entre los problemas pendientes de solución en este campo está el mencionado en el capítulo cuarto, relativo a los criterios estadísticos de fiabilidad, y particularmente a los criterios de agregación. Como vimos entonces, los indicadores se elaboran sintetizando un conjunto más o menos amplio de variables, pero el peso relativo que se le otorgue a cada una de ellas tiene que estar justificado con razones científicas. Es más, el presupuesto que subyace a las tareas de agregación es que con ellas se reúnen variables que pueden ser alineadas en una misma categoría. El índice de inflación, por ejemplo, se obtiene por agregación de las variaciones de precios de muy diversos productos y servicios, pero está claro que a través de todos ellos estamos evaluando un único concepto; la variación de los precios, por lo que pueden ser sumados en un mismo indicador agregado. Pero si el concepto que se quiere evaluar es «derechos humanos» no está tan claro que puedan sumarse sin más la irretroactividad de las normas penales contemplada en el artículo 11.2 de la Declaración Universal con el derecho a una remuneración satisfactoria que garantice la calidad de vida, como reconoce el artículo 23.3 de la misma Declaración. El conjunto de condiciones que determinan la eficacia de estos derechos es muy heterogéneo, e incluye disponibilidad presupuestaria en manos del poder ejecutivo, institución de una administración de justicia independiente con procedimientos regulares, equilibrio del mercado de trabajo, etc. Incluso que las tres categorías de derechos; civiles, políticos y socioeconómicos puedan cuantificarse con los mismos criterios puede ser discutible. No resulta demasiado difícil definir indicadores para derechos como la educación o la seguridad social (artículos 26 y 22 de la Declaración Universal), pues pueden estimarse a través de la proporción de gasto público reflejada en los presupuestos anuales de cada Estado. Están además disponibles las estadísticas publicadas al respecto por organizaciones internacionales; por ejemplo, las tasas de analfabetismo o el porcentaje de población con estudios superiores, medios y primarios publicados por la UNESCO. Sin embargo, resulta más complejo valorar cuantitativamente el derecho de asilo (artículo 14 de la Declaración Universal), pues

como señala Javier de Lucas⁴², a menudo se concede un poder discrecional muy amplio a las autoridades administrativas en el reconocimiento de la condición de refugiado y la concesión del permiso de residencia y de trabajo, a costa del control judicial y las garantías que ello comporta. Fijar un criterio que permitiese comparar toda la casuística de prácticas administrativas resulta entonces casi imposible.

Que todos estos elementos puedan ser reunidos formando parte de una misma categoría es algo que requiere una buena explicación, y la unidad de todos ellos no puede ser simplemente supuesta. Las posibilidades de desarrollo futuras de los indicadores de derechos humanos necesitan resolver en primer lugar los problemas de conceptualización que hasta ahora han dejado excesivamente desatendidos.

La revisión que Gupta, Jongman y Schmid hacen del indicador diseñado por Humana es otro buen ejemplo de un intento de mejorar los aspectos puramente técnicos sin haber establecido criterios de agregación suficientemente justificados. Los autores elaboraban una tabla con el peso relativo de cada uno de los cuarenta elementos del cuestionario del sistema de Humana por medio de la aplicación de una técnica de discriminación multivariantes. Mencioné en su momento que, si bien los autores empleaban aquella tabla para mostrar la relevancia de su aportación metodológica, de ella se deducían también algunas consecuencias sobre el significado de los derechos humanos en las que ellos mismos no parecían recalar suficientemente. Volvamos ahora sobre esa cuestión.

En primer lugar, Gupta, Jongman y Schmid separan en tres agregados los datos de Humana reprocesados por el análisis discriminativo: un agregado de integridad vital, otro de derechos políticos y otro de derechos civiles. No obstante, pueden ensayarse

⁴² Cfr. LUCAS, Javier de. *El desafío de las fronteras. Derechos humanos y xenofobia frente a una sociedad plural*. Temas de Hoy, Madrid, 1994. P. 219

otras posibilidades. Por ejemplo, podemos diseñar un agregado que contuviese los elementos del cuestionario de Humana más directamente relacionados con las garantías procesales de los derechos humanos. Estos serían DETENTN, NOSEARCH, PRVGUILT, QKTRIAL, PUBTRIAL, INDCOURT y LEGALAIID. Estos elementos ocupan las posiciones 2, 4, 6, 9, 10, 12 y 29 respectivamente. El coeficiente medio de los cuarenta elementos es 0,22742, mientras la media de estos siete elementos referidos a las garantías procesales es de 0,31055. Se deduce, por tanto, un peso relativo considerable de este agregado dentro del conjunto del indicador.

Aun pueden ensayarse otras posibilidades; por ejemplo revisar el agregado formulado por Humana como índice de integridad vital y aceptado como válido por Gupta, Jongman y Schmid. Según hemos visto, Humana suponía que los siete elementos de su cuestionario que componían este índice tendrían una importancia decisiva, y por eso les había asignado un valor triple. Sin embargo, el análisis discriminativo mostró que cuatro de ellos caían a las posiciones 18, 23, 31 y 35 del cuestionario en función del peso relativo de cada uno. Los otros tres eran MURDER, DETENTN y TORTURE. El segundo de ellos tiene una clasificación susceptible de discusión. En cuanto mide el derecho a no ser detenido sin cargos puede ser incluido en el núcleo de los derechos humanos, aquellos cuya falta de respeto implica violaciones de carácter siempre grave. Pero, en sentido estricto, las violaciones del derecho a la libertad no son amenazas a la vida o a la integridad física. Por tanto, podría clasificarse entre las garantías procesales más que en el índice de integridad vital donde lo sitúa Humana.

Dicho de otro modo, lo que aquí se mide no es si en un Estado se detiene o no a las personas, sino si se las detiene con o sin garantías. Cuando la detención se hace por un procedimiento judicialmente supervisado, siendo el detenido sospechoso de cometer un delito previsto en la ley, el caso no puede computarse como violación de un derecho. En cambio, la tortura es siempre tortura, y debería contarse como tal aun en el caso de que las normas vigentes en el Estado correspondiente la contemplasen. Si

mantenemos, por tanto, en el índice de integridad vital sólo los otros dos elementos, encontramos que ocupan las posiciones 1 y 3 del cuestionario ponderado por el análisis discriminativo. Y que tienen una media de coeficiente de 0,41378, que dobla la media de los otros 38 elementos. A su vez, los dos elementos más directamente relacionados con la participación democrática, POLTCOPP y BALLOT, tienen un coeficiente medio de 0,29905, que es un 31,5% superior a la media. Es preciso insistir una vez más que estas cifras no son ninguna expresión directa de la realidad. Por esta razón tampoco pueden extraerse conclusiones inapelables de este mero cruce de datos, lo que presupondría que están aquí cuantitativamente traducidos los derechos humanos de una forma perfectamente fidedigna. Pero sí pueden tratarse las conclusiones como indicios, como sugerencias para contrastar con otras investigaciones, o para diseñar otras metodologías que busquen de manera más específica lo que aquí se apunta. Y sobre todo para poner a prueba hipótesis de trabajo que, en la medida en que se vean avaladas por el contraste con este tipo de información cuantitativa irán dando lugar, como decía Claude, a una teoría sociológica de los derechos humanos respaldada por datos.

En esta misma línea, otro presupuesto revisable en la metodología de Humana es la elección de siete elementos en su cuestionario derivados del artículo 19 de la Declaración Universal de los Derechos Humanos (PAPER, BOOK, TVRADIO, FREEINFO, FREEPRESS, MONITOR y FREEART), que como vimos no tienen un papel demasiado determinante como para explicar que de un sólo artículo se extraigan siete elementos, mientras no se asigna ninguno al derecho a la seguridad social (artículo 22 de la Declaración Universal y artículo 9 del Pacto Internacional de Derechos Económicos, Sociales y Culturales). La misma crítica merece el informe anual de Freedom House, que no por casualidad coloca en primer lugar en el cuestionario de derechos civiles la pregunta por los medios de comunicación, literatura u otras expresiones culturales libres e independientes. Y no sólo se trata de eso; en el informe particular de cada uno de los Estados, los informes de Freedom House siempre dedican un espacio considerable a

enumerar los diarios, emisoras de radio y televisión, editoriales, etc. de propiedad privada que operan en el país en cuestión. Tampoco Freedom House se cuestiona qué grado de cobertura de la seguridad social disfrutaban los ciudadanos de cada Estado. Una vez más se ha construido una imagen a priori de los derechos humanos en que sólo se observan los hechos que puedan avalar un aspecto de ellos al que se concede especial importancia, sin que esté justificado el por qué de esta atribución.

Debe reseñarse que, considerando la generalidad de investigaciones empíricas en materia de derechos humanos, la formalización de hipótesis en relación con las variables más relevantes para explicar su situación a nivel internacional tiende preferentemente a fijarse en cuestiones de desarrollo económico. No existe justificación de este extremo, sino que más bien da la impresión de ser un prejuicio del que los investigadores no son demasiado conscientes. Además de los trabajos que he citado hasta ahora, se ha convertido en un cierto tópico esta correlación⁴³. El problema no se limita a que tal inferencia sea hecha por ciertos investigadores que se apresuran a sacar conclusiones de los estudios analizados más arriba. Lo grave es que este tópico ha tenido incluso eco institucional. El Informe sobre Desarrollo Humano del programa homónimo de Naciones Unidas, en su edición de 1992⁴⁴ estableció que existe una relación estrecha entre los Estados con una alta puntuación en la tabla de derechos políticos del informe de Freedom House y los que tenían alta puntuación en Índice de Desarrollo Humano, que evalúa ingresos per capita junto con

⁴³ Cfr. por ejemplo NEUFVILLE, Judith Ines de. *Social indicators of basic needs: quantitative data for human rights policy*. En *Social Indicators Research*. n° 11 (1981). pp. 391 y ss. En este artículo la autora, tras revisar los informes de Humana y Freedom House, estima que existe una correlación fuerte entre niveles altos de respeto por los derechos humanos y un alto PIB. Pero sólo ha ido a verificar esa hipótesis, y no otras posibles, lo que es indicativo de la tendencia a interpretar los datos exclusivamente a la luz de la relación entre desarrollo económico y derechos humanos.

⁴⁴ United Nations Development Programme. *Human Development Report, 1992*. Oxford University Press. New York, 1992. p. 32.

variables relativas a la salud y la educación⁴⁵. Pero el informe no aportó ninguna información estadística que certificase tal información. Igualmente, el Banco Mundial, apoyándose también en algunos de los informes arriba citados, afirmó la interrelación entre desarrollo y libertades públicas⁴⁶, pero sin confirmar esta tesis con soporte estadístico, o al menos examinando la fiabilidad estadística de los informes citados en su apoyo, fiabilidad que ciertamente era bastante cuestionable.

«Desde finales de los sesenta hasta mediados de los 80 la visión de la mayoría de los miembros de Naciones Unidas era, al menos al nivel retórico, que resultaba necesario conceder cierta prioridad a los derechos económicos, sociales y culturales que habían sido definidos en una serie de documentos como condición para la plena realización de los derechos civiles y políticos»⁴⁷

El hecho de que la formulación de hipótesis tienda preferentemente a buscar correlaciones entre desarrollo económico y derechos humanos resulta tanto más significativa si se tiene en cuenta que, en general, los estudios cuantitativos sobre derechos humanos son muy pobres en soporte teórico. Y esta no es una cuestión menor. No se trata de hacer elucubraciones, sino que todo estudio empírico se basa siempre en una hipótesis de trabajo, incluso cuando el investigador no es consciente de ello, sencillamente porque no puede buscar *todos* los datos, sino aquellos que considera relevantes, y discriminar qué datos son relevantes frente a los que no lo son implica disponer de una teoría que justifique dónde está lo importante. El investigador puede hacer más o menos explícita esta base teórica, y cuanto menos lo haga menos fiabili-

⁴⁵ Más adelante tendremos ocasión de volver sobre este informe de Naciones Unidas, ya que incluye algunos otros aspectos dignos de atención.

⁴⁶ World Bank. *World Development Report, 1991; the challenge of development*. Oxford University Press. New York, 1991. p. 50.

⁴⁷ TÜRK, Danilo (Special rapporteur for the Commission of Human Rights of the United Nations). *The realization of economic, social and cultural rights*. Final report submitted to the 44th session of the Commission of Human Rights. Cit. p. 5

dad tendrá el estudio. Desde la pobreza teórica que parece caracterizar la investigación empírica sobre los derechos humanos, la única línea que parece descollar algo es justamente la de la correlación entre derechos humanos y desarrollo socioeconómico. En buena medida, esta línea de trabajo proviene del recuerdo de la génesis de los derechos humanos en los países occidentales, génesis coincidente con la modernización económica, por lo que ambos aspectos parecen ir necesariamente en relación. Así lo expresa, por ejemplo, Raymond Aron:

«Si no es cierto que las fases iniciales de la industrialización determinan (en el sentido de “hacer necesarios”) los despotismos y menos aun los totalitarismos, sí es verdad que las tensiones inevitables, ligadas a los comienzos de la modernización económica, reducen la probabilidad de un régimen que entrañase a la vez la rivalidad de partidos, participación de las masas y libertades personales. En Europa, en fases comparables de desarrollo, el régimen era a veces parlamentario, pero el parlamentarismo era de esencia aristocrática y las masas populares se hallaban o privadas de derechos políticos o sometidas efectivamente a los privilegiados. Que los países no occidentales sacrifiquen las libertades políticas y personales para satisfacer las necesidades, reales o supuestas, de la modernización es algo de lo que no tenemos por qué asombrarnos»⁴⁸.

Ciertamente la modernización económica europea hasta llegar a la época del reconocimiento constitucional de los derechos fundamentales no fue una situación en que floreciesen las libertades públicas, pero la estabilización económica y la de los derechos fundamentales tampoco guardan una relación de simetría perfecta como para generalizar sus interdependencias. Y, lo que es más importante, también entonces se produjo en Europa la secularización, la urbanización acelerada, el progreso en la escolarización o en las condiciones de la salubridad pública. Por no mencionar factores no tan «cuantificables», como el proceso de codificación jurídica y todo lo que ella entraña, el cambio en las formas de dominación política o el cambio de los paradigmas científicos.

⁴⁸ ARON, Raymond. *Ensayo sobre las libertades*. Alianza Editorial. Madrid, 1974. pp. 93-94

Robert Dahl, quien también se ha tomado la molestia de contrastar sus hipótesis sobre los factores que favorecen la democratización con investigaciones empíricas, señala la complejidad de la correlación entre democracia (o poliarquía, siguiendo su terminología) y desarrollo económico:

«– El moderno capitalismo de mercado no es estrictamente una condición necesaria para la democratización, o al menos no lo fue para las poliarquías masculinas que se desarrollaron en el siglo XIX y a principios del XX: Estados Unidos, Canadá, Nueva Zelanda, Australia, Noruega, Dinamarca, Suecia y Suiza.

– El moderno capitalismo de mercado es definitivamente insuficiente para mantener la democracia: pensemos en la Alemania de los años veinte.

– Sin embargo, el moderno capitalismo de mercado parece generar condiciones muy favorables para la poliarquía.

– Lo que es más, ningún país que no tuviera alguna forma de capitalismo, sea agrario o moderno, ni de propiedad privada, ha evolucionado hacia una democracia estable.

– No obstante, el moderno capitalismo genera también un amplio abanico de desigualdades en la distribución de recursos, incluidos los recursos políticos. En este sentido, es contrario a la igualdad política necesaria para una democracia plena»⁴⁹.

De estas observaciones se deduce, como afirma el propio Dahl de manera un tanto irónica, que el capitalismo es fundamentalmente incongruente con la democracia, pero a la vez es el único sistema económico compatible con ella. O, en otros términos, que el capitalismo es condición necesaria pero no suficiente de la democracia, y en aquello en que no es condición suficiente resulta incluso antidemocrático. De todo ello se deduce que la formación de hipótesis acerca de las variables más significativas en relación con los derechos humanos podría ser bastante más completa que la correlación simple y lineal con el desarrollo económico. En este

⁴⁹ DAHL, Robert. *Desafíos de la teoría democrática*. Seminario impartido el 24 de Mayo de 1996 en el Centro de Estudios Avanzados en Ciencias Sociales del Instituto Juan March. Publicado en *Nueva Revista*, n° 50 (Abril-Mayo de 1997). pp. 172-73.

sentido, la propuesta de Barsh parece más razonable. No se trataría de buscar relaciones causales entre desarrollo y derechos humanos, sino de seguir el modelo de la Declaración del Derecho al Desarrollo: es decir, considerarlos conjuntamente.

Frente a la tendencia a seguir el tópico de la correlación directa y monocausal entre desarrollo económico y derechos humanos, es necesario recordar que la investigación empírica ha llegado a conclusiones más complejas, que toman en consideración más factores, y que incluso demuestran la falsedad del tópico mencionado. En este sentido, y con todas las limitaciones y cautelas del caso, la investigación de Gupta, Jongman y Schmid parece indicar que, en el contexto de los sistemas sociales contemporáneos, los factores más correlacionados con la eficacia de los derechos humanos son la efectividad del monopolio de la violencia física legítima y el ejercicio del poder a través de un sistema normativo racionalizado. Conclusiones ambas que concuerdan plenamente con el modelo sociológico de explicación de los derechos humanos que, como hemos visto en la primera parte, se deduce de la teoría weberiana de racionalización del derecho.

El monopolio de la violencia física legítima significa que quedan excluidos de la lucha social los medios violentos, en el sentido en que Norbert Elias describe la interiorización de la coacción y la progresiva desaparición de la vida cotidiana de la amenaza próxima de perder la vida o la integridad física. Y, como vimos que afirma Weber, esa desaparición confiere al aparato coactivo que respalda al sistema político de control social su legitimidad específica. Claro está que la desaparición real de la amenaza a la vida o la integridad es un hecho que afecta a todos por igual, y por tanto se aplica también a los medios empleados por ese aparato de control social. En definitiva, que las pretensiones de legitimidad del Estado están directamente relacionadas con este factor. Y si las pretensiones de legitimidad se expresan en los derechos humanos, como ocurre hoy en día y examinan Gupta, Jongman y Schmid, entonces su vigencia depende altamente de este factor. Por su parte, el ejercicio del poder a través de un sistema normativo autónomo y racional, es decir, a través de un derecho

racionalizado, también confirma las tesis weberianas de forma suficientemente clara.

Igualmente, la *Liste de descripteurs des droits de l'homme*, de A.C. Kiss, J.P. Massué y K. Vasak, investigadores del Institut International des Droits de L'Homme de Estrasburgo, como vimos, conduce a la conclusión de que el tipo de conceptos más frecuentemente invocados en los textos normativos, jurisprudenciales y doctrinales sobre los derechos humanos son los mismos que sirven para caracterizar al Estado de Derecho, en que todo poder se ejerce a través de normas jurídicas. Como ya se citó en el capítulo anterior, los autores de este estudio consideraban «desalentadora» esta conclusión a la que ellos mismos habían llegado; que los conceptos que con más frecuencia aparecían en los textos no constituían un grupo específico delimitado por su relación con los derechos humanos, sino que pertenecían a «la trama de toda rama jurídica de las ciencias humanas»⁵⁰. La conclusión puede parecer desalentadora desde el punto de vista de un investigador que está tratando de localizar un elemento novedoso, pero no cuando en lugar del significado que tenga el dato aislado lo contemplamos desde el punto de vista de un modelo general del desarrollo de los derechos humanos. Lo que se hace entonces perceptible es la íntima conexión entre el monopolio de la violencia legítima, con el consiguiente establecimiento de un sistema normativo racionalizado, y la eficacia de los derechos humanos. La racionalización del derecho, que es la base de la teoría weberiana, se confirma así como acertada.

Por otra parte, Boli-Bennett, en su análisis de contenido de los derechos constitucionales observa una variación sustancial en la definición de los derechos de los ciudadanos entre países latinoamericanos, musulmanes de oriente medio, asiáticos y europeos. Como ya mencioné en su momento, Boli-Bennett descubrió que los países de tradición católica formulan constitucionalmente los

⁵⁰ KISS, A.C., MASSUÉ, J.P. y VASAK, K. *L'ordinateur au service des droits de l'homme*. Cit., p. 542.

derechos fundamentales de los ciudadanos de manera mucho más extensa de lo que lo hacen los países con una herencia protestante, y que esta extensión es especialmente notable allí donde tuvo arraigo el iusnaturalismo característico del ámbito cultural católico. Este descubrimiento es revelador de cómo el aspecto doctrinal de los derechos humanos tiene una importancia destacada a la hora de explicar su diferente vigencia a lo largo del mundo. Este es un aspecto que también confirma las hipótesis weberianas. No pretendo ser exhaustivo en la comparación entre los resultados de la investigación empírica y la teoría sociológica de los derechos humanos, algo de lo que me ocuparé más ampliamente en las conclusiones. Pero sí me interesa destacar que la pobreza teórica de la investigación empírica que mencionaba Claude, citado al comienzo de este epígrafe, es el problema más destacado de la investigación empírica, pero que el apoyo recíproco de datos que confirman hipótesis y de hipótesis que indican dónde ir a buscar datos es posible, como confirman los ejemplos citados de Gupta, Jongman y Schmid; Kiss, Massué y Vasak o el de John Boli-Bennett.

No quiero acabar sin volver sobre la conclusión a la que llega Richard Claude, uno de los más acreditados especialistas en el campo del derecho comparado de los derechos humanos, entendido en sentido sociológico, sintetizando los resultados de los diversos trabajos que componen el libro colectivo «Comparative Human Rights», del cual fue responsable:

«Los autores que ha contribuido a este volumen comparten las siguientes conclusiones relativas a los derechos humanos. Primera; que para que prevalezca un sistema de derechos humanos se requiere como condición necesaria un sistema legal seguro para dispensar justicia y procedimientos regularizados para resolver disputas. Segunda, que un sistema legal modernizado descansa sobre estructuras legales institucionalmente diferenciadas y procedimientos racionales establecidos para resolver conflictos. Tercera, que cuanto más estructuralmente diferenciado y flexible es el sistema legal (sea a través del desarrollo legislativo, de criterios procesales normativamente establecidos y de la jurisprudencia, o de ambos), tanto más capaz es el sistema de responder al incremento del cambio social y político. Cuarto, que la idea del Estado de Derecho asociada con el establecimiento de límites legales sobre la

acción del gobierno, si bien está sujeta a interpretaciones que varían a lo largo de los diversos países, deriva de los principios, instituciones y procedimientos que la experiencia y tradiciones de los juristas y otros operadores jurídicos ha demostrado ser importante para proteger a los individuos de la arbitrariedad gubernamental. Quinto, que mientras las políticas públicas del Estado del bienestar sobre derechos positivos — típicamente modernas— se amplían codo a codo con la centralización administrativa, las necesidades organizativas del control de programas y de supervisión tienden a entrar en conflicto con ideas anteriores sobre el Estado de Derecho y con ideas anteriores acerca de los derechos negativos, tales como la protección de la propiedad y la privacidad y la prohibición de la discriminación»⁵¹.

De manera más concisa, expone su conclusión de que «el esfuerzo de crear uniformidad en los derechos civiles, políticos y sociales a lo largo de un país podría ser parte de una estrategia de centralización, más que un movimiento humanitario»⁵². Como puede verse, la imagen de los derechos humanos que de aquí se desprende es plenamente coincidente con la teoría weberiana de la racionalización del derecho que sirve como base para la explicación sociológica de los derechos humanos, como expuse en la primera parte de esta investigación.

⁵¹ CLAUDE, Richard P. *Comparative rights research: some intersections between law and the social sciences*. En CLAUDE, Richard P. (ed.), *Comparative human rights*. The John Hopkins University Press. Baltimore, 1976. p. 389.

⁵² CLAUDE, Richard P. *Comparative rights research: some intersections between law and the social sciences*. Cit. p. 388.

Conclusiones

De entre todos los campos de investigación de la sociología jurídica, el de los derechos humanos es posiblemente uno de los menos consolidados, siendo esta la primera conclusión a la que se llega cuando se analizan en detalle los trabajos realizados sobre la materia. Sólo el tiempo y la coordinación de los esfuerzos permitirán asentar premisas y criterios metodológicos comunes, pero hoy en día su signo más característico es todavía la dispersión. Por este motivo, una parte de esta investigación ha consistido en una labor recopilatoria, sobre todo en lo que afecta a la sección dedicada a las investigaciones empíricas. Al revisar de esta manera los diversos modelos de investigación pueden comenzar a destacarse aquellos puntos que merecen el más amplio consenso en cuanto disyuntivas metodológicas ya resueltas, y sobre las que puedan apoyarse ulteriores desarrollos, de forma que se supere aquella dispersión señalada. Trataré de mostrar cuáles son, en mi opinión, esos lugares comunes para investigaciones futuras. Naturalmente, también expondré mis opiniones personales en aquellos puntos debatidos y sobre los que todavía no hay consenso.

Si es la dispersión lo que caracteriza a este campo, el problema se deja notar especialmente en la falta de relación entre la investigación empírica y la teoría sociológica de los derechos humanos. Aun a riesgo de una excesiva simplificación, puede afirmarse que de la primera se ocupan fundamentalmente politólogos y expertos en estadística de origen estadounidense, mientras la teoría sociológica es más bien un fenómeno europeo y llevado a cabo en el ámbito sociológico-jurídico. La distinta naturaleza de los trabajos en uno y otro campo dificultan las posibilidades de comparación directa entre ellos. Por esta razón he tenido que analizar en partes separadas la teoría sociológica y la investigación empírica. Sin embargo, considero que ninguna de las dos se basta a sí misma. El principal propósito de las conclusiones de esta investigación es, en consecuencia, sintetizar los resultados obtenidos separadamente en el análisis de los dos campos para establecer así un terreno

común que pueda servir de base a ulteriores trabajos. La segunda conclusión es que no puede haber una sociología de los derechos humanos sólidamente asentada si no se apoya en la teoría sociológica y en la investigación empírica, reformulando los resultados de una en función de la otra.

El primer problema común a la teoría sociológica y a la investigación empírica es la falta de definición de un paradigma claro de investigación. Y sin resolver este tipo de cuestiones preliminares es difícil que la sociología de los derechos humanos llegue a consolidarse. He reunido todos los modelos empleados en las diversas investigaciones realizadas sobre la materia y he tratado de mostrar sus limitaciones y sus ventajas. Estos paradigmas de investigación de la sociología de los derechos humanos, hasta donde yo he podido ver, son seis: el modelo de la dialéctica sociedad civil-Estado, el del relativismo, el del normativismo jurídico, el del normativismo moral, el funcional-estructuralista de Luhmann y el de la racionalización del derecho de Weber. Sobre la base del modelo weberiano, Habermas ha intentado una variación que podría tal vez convertirse en una séptima alternativa, pero en sentido estricto no se trata de un paradigma sociológico, sino ético, por lo que no se justifica añadirlo a la lista de los seis considerados. La tercera conclusión, y posiblemente la principal, es que el paradigma weberiano ofrece, en mi opinión, la base más sólida para construir una sociología de los derechos humanos con apoyo en la teoría y en la investigación empírica. Avalan esta conclusión, en primer lugar, las debilidades de los otros modelos, y en segundo lugar la confluencia de resultados de la teoría weberiana con las conclusiones principales de los estudios empíricos realizados.

El modelo de la dialéctica sociedad civil-Estado está ejemplarizado en las propuestas de Podgórecki, Treves y Ferrari, todas en el ámbito de la teoría sociológica, mientras sólo James Scarritt ha planteado investigaciones empíricas que en algún aspecto muy puntual coinciden con este modelo. Supone que los derechos humanos se formulan siempre en primer lugar como expectativas que se consolidan en el nivel de las relaciones sociales

no institucionales, y que después encuentran reconocimiento legislativo por parte de la autoridad política, si bien, como afirma Ferrari, esto no es esencial cuando se tiene una idea democrática de la sociedad que se autodetermina. Pero esta conceptualización plantea varios problemas. En primer lugar, resulta difícil extraer una definición operativa de tal concepto de derecho que pueda ser empleada en la investigación. En esa formulación preinstitucional, que Podgórecki llama «intuitive law», propiamente tendríamos que entender por «derechos» ciertos estados de opinión generalizados entre la población. La cuestión es cuál es entonces el factor que unifica tantos estados de conciencia como para que en todos ellos se presenten con el mismo sentido conceptos tan precisos como «habeas corpus» o «derecho a no ser privado arbitrariamente de la propiedad». En realidad, el modelo es débil por la circularidad de su argumentación. El tipo de sociedad que exista, incluido el tipo de expectativas dominantes, depende esencialmente del tipo de sistema político en que se articule esa sociedad, y la generalización uniforme de expectativas no es analíticamente separable de las garantías que las respaldan y que son posibles en función de los sistemas de control social, y especialmente de sistemas de control estatal cuando se trata de derechos. Tocqueville y Peces-Barba, entre otros, muestran cómo el monopolio estatal de la creación del derecho es el factor que permite la formación de consensos socialmente generalizados sobre la igualdad ante la ley, un punto éste en el que Weber profundizará. En general no hay suficiente base histórica para afirmar que los derechos humanos se formulan siempre primero en el nivel de los valores sociales convencionales y sólo en un segundo momento se trasladan al plano jurídico institucional.

El modelo relativista se caracteriza por situar como objetivo de la sociología de los derechos humanos el desvelamiento de la falta de correspondencia entre las pretensiones incondicionales de valor en que se enuncian tales derechos y la realidad social de los mismos. Ejemplos de esta argumentación en el campo teórico son Gaete, Geiger o Fariñas, y en ciertos aspectos también Luhmann. En el ámbito de la investigación empírica, Boli-Bennett sigue

parcialmente este modelo, aunque lo que predomina en su metodología es un paradigma más homologable al weberiano. En todos los casos, aunque en diferentes grados, se supone que la sociología debe discutir los criterios de *justificación* de los derechos humanos, pero la sociología no es una ciencia normativa. Al ceñirla a esta perspectiva se dejan de considerar todos los fenómenos en que ciertos enunciados de valor, sean verdaderos o falsos para el investigador que los observa, inducen efectivamente a los miembros de una determinada sociedad a actuar en correspondencia con ellos. En estos casos, desacreditar los valores por falsos (aunque lo sean) impide observar su eficacia, que es independiente de su verdad o falsedad. Se trata, por tanto, de un modelo incompleto, que no atiende a toda la fenomenología social de los derechos humanos, cuestión aparte de que se admitan los juicios de valor en el método sociológico.

El modelo del normativismo jurídico se corresponde con la concepción de Kelsen sobre la sociología del derecho y su método, es decir, una ciencia sin capacidad para definir su objeto propio, pues sólo la ciencia del derecho puede establecer cuándo una norma es jurídica o no. Consecuentemente, el método de la sociología jurídica se limita al análisis de la eficacia o ineficacia de las normas, entendidas éstas en sentido jurídico formal. Se trata de uno de los modelos más importantes porque la mayor parte de la investigación empírica se asienta sobre él, aunque no hay un desarrollo de teorías sociológicas equivalentes. Es el modelo que, por ejemplo, subyace a las investigaciones de Humana, la realizada por Gupta, Jongman y Schmid, el indicador de Freedom House o el informe anual *Human rights in developing countries*. Ninguno invoca la autoridad de Kelsen, aunque sigan sin saberlo sus presupuestos.

Lo que todos ellos tienen en común es suponer que las normas internacionales sobre derechos humanos pueden ser directamente evaluadas mediante procedimientos cuantitativos. Sin embargo, todo método de cuantificación necesita resolver en primer lugar las tareas de conceptualización de su objeto. El error de estos sistemas está en suponer que este paso puede ser suplido por la referencia

directa a las normas internacionales. Como he mostrado, la ausencia de este primer paso plantea inconvenientes desde los criterios de fiabilidad estadística de medida, precisión, consistencia, validez, equivalencia y agregación. Por otra parte, la metodología seguida deja sin respuesta otro de los problemas básicos de este campo: el del multiculturalismo. A pesar de la existencia de normas internacionales de derechos humanos que, intencionalmente, tienen igual vigencia en todo el mundo, la diversidad de su eficacia que puede comprobarse tiene que ser explicada por factores distintos a que haya sido establecido un sistema normativo como referente común. Se revela así la debilidad metodológica de estos planteamientos. Efectivamente, si se analizan los derechos humanos en cuanto presupuestos normativos fuera de discusión y tratamos de ver cómo operan otros factores sociales en cuanto variables capaces de determinar su eficacia, se introduce ya una hipótesis subrepticamente; la de que no puede ocurrir que sean los derechos humanos los que varíen en función de otros factores sociales. La única conclusión aceptable es que el elemento normativo sea reintegrado como un factor más entre otros, pero entonces hace falta una conceptualización de los derechos humanos que no se base en exclusiva en las normas internacionales sobre los mismos. La solución al problema de diseñar un sistema de evaluación cuantitativa de los derechos humanos capaz de asumir el hecho del multiculturalismo pasa de nuevo por las tareas de definición conceptual, y esto implica una teoría sociológica que aporte hipótesis concretas sobre qué tipo de datos son relevantes y por qué.

Los problemas a que conduce el normativismo en la investigación empírica sobre los derechos humanos han llevado a autores como Flathman, Scarritt, McCamant o Barsh a proponer el abandono de los intentos de cuantificar derechos y sustituirlos por otros conceptos referidos a situaciones fácticas más susceptibles de cuantificación. Así, por ejemplo, Barsh propone evaluar recursos humanos (como tasas de alfabetización o de cobertura sanitaria), seguridad personal y formas de poder político. En cualquier caso, puede establecerse la conclusión de que los sistemas cuantitativos

ensayados hasta ahora necesitan el soporte de teorías explicativas, conclusión en la que coincido plenamente con Richard P. Claude. Ciertamente la experiencia acumulada en estas investigaciones de tipo cuantitativo no puede ser tirada por la borda sin más por la ausencia de las tareas previas de conceptualización y la adscripción al modelo normativista, y particularmente en el campo de los indicadores de derechos humanos se han hecho avances en criterios de medición, en localización de fuentes estadísticas, en el establecimiento de criterios de agregación, o en el peso ponderado que deben tener los principales indicadores. He analizado todos estos puntos, repasando las propuestas de revisión metodológica de Scarritt, Barsh, McCamant o Claude. Aunque existe ya una experiencia acumulada con consensos relativamente firmes sobre algunos de los criterios, en buena medida el debate metodológico sigue abierto. Sea cual sea la dirección a seguir en el futuro, sí cabe concluir, en mi opinión, que la clave pasa por abandonar el modelo del normativismo jurídico y definir un modelo general de desarrollo evolutivo de los derechos humanos sobre el cual se puedan articular variables particulares que expliquen la variabilidad de la eficacia de los derechos humanos en los diferentes contextos culturales. También en esto coincido con las conclusiones a las que llega Claude.

El modelo del normativismo moral trata los derechos humanos desde el punto de vista de las representaciones morales socialmente compartidas en la medida en que sirvan de causa eficaz del comportamiento de los individuos. Se trata de un modelo poco relevante, pues aparte de la teoría de Sánchez de la Torre no hay muchos más casos representativos que hayan intentado su puesta en práctica de manera sistemática. Sin embargo, tiene importancia por ser un argumento recurrente que muy diversos autores emplean como objeción a otros paradigmas de investigación. Básicamente consiste en identificar los derechos humanos con valores que merecen asentimiento por razones éticas. Este es el argumento que está en la base de la crítica de Habermas a Weber, o en muchas de las críticas que se hacen a Luhmann. Pero aunque hay teorías éticas sobre los derechos humanos en cuanto valores, no se han

desarrollado teorías equivalentes de carácter sociológico. De hacerse, se encontrarían con problemas semejantes a los del modelo de la dialéctica sociedad civil-Estado y a los del normativismo jurídico. Con los primeros comparte la falta de capacidad para explicar cómo se generalizan uniformemente patrones éticos llamados a orientar la vida política prescindiendo completamente de la institucionalización del aparato de dominación política. Con los segundos tiene en común el problema de que en la eficacia vinculante de los valores inciden otros factores sociales, pero conceder primacía a los primeros significa introducir en la investigación la suposición de que son esos otros factores los que experimentan variaciones y no los valores, cuando desde un método sociológico no lastrado por ideologías no es descartable lo contrario.

El modelo funcional-estructuralista de Niklas Luhmann es, junto con el del normativismo jurídico, uno de los más relevantes por su consistencia interna. En él se analizan los derechos fundamentales en cuanto elementos estructurales necesarios para mantener la diferenciación sistémica. Las sociedades modernas se caracterizan porque en ellas, a diferencia de la diferenciación segmentaria anterior, se consolida una diferenciación funcional. Cambia también, por tanto, el tipo de factores que estabilizan esta forma de diferenciación. Desde su punto de vista, los derechos humanos no son derechos de las personas, ni tienen relación con los valores socialmente vigentes; el punto de referencia ya no es la personalidad humana, ni la dignidad, ni la libertad, sino la dinámica de los sistemas, las exigencias estructurales de diferenciación de las sociedades contemporáneas. Evidentemente, este modelo elude los problemas propios del normativismo moral o del de la dialéctica sociedad civil-Estado, a la vez que merece frecuentes críticas desde esas posiciones por su distanciamiento de la justificación ética de los derechos. Según Luhmann, un orden social diferenciado debe garantizar la posibilidad de generalización de la comunicación en cuatro ámbitos: la autorrepresentación de la personalidad, la formación de expectativas de comportamiento fiables, la satisfacción de las necesidades económicas y la posibi-

lidad de adoptar decisiones públicas vinculantes. De ellos se deducen, respectivamente, los derechos propios de la libertad de acción individual, los demás derechos civiles, los derechos a la propiedad y a la libre elección de profesión y, por último, los derechos de participación política.

Entre otras consecuencias de la diferenciación funcional de las sociedades modernas, Luhmann destaca la primacía del subsistema económico sobre el político, lo que lo vincula a investigaciones empíricas interesadas por la correlación entre desarrollo económico y derechos humanos. Por otra parte, Luhmann no incluye los derechos sociales, económicos y culturales en su obra dedicada a los derechos fundamentales. Del primado del sistema económico deduce que, si la garantía de los derechos fundamentales está en que la política no invada las otras esferas de acción social, especialmente debe garantizar la independencia del sistema económico, y esta primacía asegura a los derechos fundamentales una base firme. La investigación empírica, no obstante, no parece darle la razón. Zehra Arat ha analizado las pautas típicas de evolución de los derechos humanos en los países en desarrollo y ha demostrado que los derechos sociales, económicos y culturales no son separables de los derechos civiles y políticos, so pena de que el sistema en su conjunto sufra una degradación del segundo tipo de derechos por falta de garantías de los primeros. John Boli-Bennett ha demostrado también que la expansión histórica de los derechos civiles y políticos está íntimamente relacionada con la de los derechos económicos, sociales y culturales. Por el contrario, ninguna investigación empírica ha logrado mostrar que el desarrollo económico por sí mismo sea un factor causal suficiente en la génesis y consolidación de los derechos humanos. Quienes han afirmado tal cosa, como algunos informes del Banco Mundial o el indicador anual de Freedom House, lo han hecho por razones ideológicas, pero no han podido justificarlo con datos.

El esquema funcionalista de Luhmann tiene muchos puntos de coincidencia con el modelo weberiano, no tanto en las premisas generales como en la descripción de los factores sociales que inciden concretamente en cada uno de los derechos fundamenta-

les. También tienen algunas discrepancias que señalaré tras exponer las conclusiones sobre el modelo de Weber.

Por último, el modelo de la racionalización del derecho de Weber permite analizar los derechos humanos como las pretensiones típicas de legitimidad de los sistemas políticos con un ordenamiento formalizado y racionalizado. Junto con el modelo del normativismo jurídico y el de Luhmann es uno de los más relevantes y, en mi opinión, el más sólido y el que más confirmación encuentra en las investigaciones empíricas realizadas.

En primer lugar, Weber establece un concepto sociológico de derecho que permite evitar los problemas de los modelos del normativismo jurídico ni en los del relativismo o del normativismo moral. Las normas, desde la metodología weberiana, deben ser examinadas fácticamente. Si algo normativamente válido se convierte en objeto de estudios empíricos, es decir, cuando es sociológicamente analizado, como tal objeto pierde su carácter de norma: se le trata como un hecho, pero no desde el punto de vista de la validez normativa. Esta es la premisa de lo que Weber desarrollará en la metodología de la *Wertfreiheit*. A la vez, la consideración sociológica no excluye que la acción de los individuos esté orientada por valores, por lo que deberá constatar cuáles son y en qué medida son realmente eficaces para determinar el comportamiento. Nuevamente, lo que se examina es la eficacia fácticamente definida de los valores. Weber denomina a este aspecto metodológico *Wertbeziehung*. Sobre estos presupuestos, define el derecho en sentido sociológico como una de las formas de garantizar externamente un orden legítimo. La legitimidad es atributo de todo un orden social, y no sólo del ordenamiento. Puede estar garantizada interna (por la capacidad de convicción ética que en cada persona despierten sus máximas) o externamente. En el segundo caso, la garantía externa puede ser convencional o jurídica. Ocurre la primera cuando las violaciones del orden legítimo son sancionadas por la generalidad de las personas (mediante boicot, ostracismo, descrédito, etc.) y aparece el derecho cuando de la aplicación de sanciones se ocupa un cuadro de individuos instituidos con la misión de hacer cumplir el orden legítimo. La clave no

está en las normas (cuyo contenido puede ser idéntico en el nivel ético, en el convencional y en el jurídico), sino en la diferenciación social del sistema político que permita instituir un cuadro coactivo garante del orden legítimo.

Los derechos humanos pueden ser definidos, a partir de la teoría weberiana, como la expresión de las pretensiones típicas de legitimidad de los sistemas políticos con un derecho racionalizado, entendiendo aquí por derecho la articulación descrita entre orden legítimo y garantías externas de un cuadro coactivo institucional, y no sólo el sistema de las normas positivas. La legitimidad también es definida en términos fácticos, pues de ella depende la eficacia de la dominación, que es la probabilidad de que mandatos específicos encuentren obediencia. Si un sistema político emplea una determinada pretensión de legitimidad y consigue mantener la dominación de forma estable ha mostrado la eficacia de su legitimación de un modo que es empíricamente verificable. La conclusión es que con los presupuestos generales de la metodología sociojurídica weberiana es posible formular un modelo que evita los problemas de los otros y que además está abierto a la verificación empírica.

El análisis detallado del proceso de racionalización del derecho muestra cómo de él se deduce la aparición de los derechos humanos en cuanto pretensiones típicas de legitimidad. Se trata de un proceso complejo, que implica la concurrencia de muy diversos factores. Me limitaré ahora a sintetizar las principales conclusiones a que conduce su análisis. Siguiendo una indicación de Febbrajo, he expuesto la racionalización del derecho analizando la racionalización de las imágenes del mundo que afectan a la *representación* del orden legítimo, sin consideración todavía al tipo de garantías que éste tenga, y a continuación estudiando la racionalización de la dominación. En relación con lo primero, tanto Weber como Habermas o Luhmann coinciden en señalar que sólo con la secularización se inicia el proceso que conduce a pretensiones de legitimidad modernas, como son los derechos humanos, pues sólo entonces se abre la posibilidad de una diferenciación en que cada uno de los sistemas de acción social se rija por

su racionalidad específica y diferenciada de las demás: la racionalidad económica, la política, la jurídica, la científica, la artística, etc. Ninguna sociedad no diferenciada en este sentido ha conocido fenómenos semejantes a la garantía jurídica de una esfera de acción individual, de la autonomía de las decisiones jurídicas frente al poder político, etc. Se puede sentar como conclusión merecedora de consenso, por tanto, que la secularización es el primer factor a tener en cuenta al analizar el grado de eficacia de los derechos humanos en una sociedad dada.

Desde el punto de vista histórico, la secularización condujo en el ámbito cultural occidental a la reformulación de las pretensiones típicas de legitimidad por obra del iusnaturalismo, la doctrina en cuyo seno se originó la axiomática de los derechos humanos. La racionalización doctrinal fue anterior a la plena racionalización de las formas de dominación, por lo que en esta fase las pretensiones de legitimidad tenían una validez de tipo ético o convencional. Es decir, no había un cuadro administrativo capaz de garantizarlas como derechos. El iusnaturalismo es la doctrina que las revoluciones invocan para legitimar su oposición al derecho vigente. En este sentido, en la génesis sociológica de los derechos humanos tienen una importancia central la actividad de los intelectuales y los procesos revolucionarios. Los intelectuales reelaboran las relaciones internas de sentido de las doctrinas para impedir que conduzcan a contradicciones en los criterios de acción que de ellas se derivan, por lo que refuerzan la consistencia de las pretensiones de legitimidad y su labor puede llegar a tener consecuencias prácticas. Y las revoluciones necesitan una doctrina que oponer al derecho vigente, que necesitan derribar, y que se presente como dotada de una mayor legitimidad que la de éste. Sociológicamente, el iusnaturalismo fue decisivo porque resolvió estas dos necesidades.

El segundo aspecto de la racionalización del derecho es la institución del cuadro coactivo, que se inscribe en la sociología weberiana de la dominación. La existencia de un cuadro administrativo que ejerza las funciones políticas es consecuencia de un proceso de diferenciación social. Tanto Weber como Luhmann coinciden en señalar el papel central de la burocratización en este

proceso. Pero el tipo de organización política que se instituya determina sustancialmente sus pretensiones de legitimidad. Efectivamente, el aparato por medio del cual se ejerce la dominación es una parte, un subsistema diferenciado frente al conjunto del sistema social, pero que a su vez pretende ser competente para gobernar a la totalidad de los ciudadanos. Aquí juega un papel clave el concepto de *representación*. Una masa de individuos sólo se constituye como comunidad política si las decisiones que una parte de los individuos adoptan es percibida por el resto como representativa, es decir, si la acción del cuadro administrativo es interpretada como acción de la colectividad. Pero para eso el cuadro administrativo necesita formular unas pretensiones de legitimidad que tengan éxito, que se presenten como vinculantes para el resto. El modo en que la generalidad de los individuos se perciben a sí mismos en cuanto miembros de la comunidad política está intrínsecamente unido al modo en que ésta organiza su aparato de dominación y establece sus pretensiones de legitimidad. Inversamente, la ineficacia de las pretensiones de legitimidad puede conducir a que sectores específicos de la comunidad política dejen de sentirse representados por el cuadro administrativo, creando sus propios mecanismos de representación y sus propias pretensiones de legitimidad. Es lo que ocurrió con los partidos socialistas y los sindicatos, de orientación revolucionaria, en la época del sufragio censitario, la prohibición de las huelgas y de la sindicación. El sistema político sólo recuperó entonces la eficacia de la dominación modificando sus pretensiones de legitimidad a base de reconocer los derechos sociales y económicos, como ocurrió en la Constitución Mexicana de 1917 o en la de Weimar de 1919.

El proceso de racionalización de la dominación se manifiesta en muy diversas consecuencias, pero en lo que afecta a la aparición de los derechos humanos como pretensiones típicas de legitimidad, las decisivas son cuatro: el monopolio de la violencia legítima, la burocratización del cuadro administrativo, la generalización de los mecanismos de representación y la autonomía del ordenamiento como un sistema social plenamente diferenciado. No son procesos sucesivos sino simultáneos.

El incremento de la socialización de la acción política va de una intervención ocasional frente a amenazas concretas hasta un instituto político de actividad continuada cuando y en la medida en que su cuadro administrativo mantenga con éxito la pretensión al monopolio legítimo de la violencia para el mantenimiento del orden vigente, y esto constituye el consenso específico entre sus miembros acerca de su legitimidad. La prohibición progresiva del uso de la violencia privada para la resolución de los conflictos ocurre únicamente cuando se transforma el deber de venganza en un castigo racionalmente organizado y las querellas y reparaciones en un procedimiento judicial según normas racionales, transformándose la comunidad política en un instituto protector de los derechos. El monopolio de la violencia legítima depende, por tanto, de la racionalización del derecho. Como ha mostrado Norbert Elias, el monopolio de la violencia legítima guarda una relación muy estrecha con la racionalización de la autoacción individual y el incremento de la interdependencia funcional propia de la racionalización económica. La capacidad de prever a largo plazo los efectos de las propias acciones entrelazadas a las acciones de otros en un contexto social altamente complejo requiere la exclusión de la violencia y la interiorización de pautas de autocontrol rigurosas. Ningún sistema de dominación puede suplir ese autocontrol con formas políticas de control social. Se separan entonces el derecho y la moral, y se instituye la autonomía individual como eje articulador de los sistemas sociales altamente diferenciados. La teoría de Elias no es rival sino complemento de la weberiana, y amplía el espectro de ésta, que sólo había considerado la racionalización de la acción individual en las sectas protestantes. De este proceso se deducen los derechos civiles que garantizan una esfera de acción individual independiente de las decisiones políticas y el derecho a la vida y a no sufrir torturas o tratos crueles, inhumanos o degradantes.

La diferenciación social del cuadro administrativo que ejerce la dominación depende en su eficacia de la forma que adopte éste. Evoluciona desde el patriarcado, pasando por sistemas prebendarios y feudales hasta la burocratización, la forma más racional y eficaz

de ejercer una dominación. Las pretensiones típicas de legitimidad de toda organización burocrática conducen a la igualdad de derechos, con independencia del carácter de la dominación. De ahí que el proceso sea el mismo en la burocracia de las comunidades políticas, económicas o en la hierocracia de las comunidades religiosas. Desde este punto de vista, burocratización significa despersonalización de la dominación, e implican siempre un sistema normativo racionalizado. Con ella aparece el Estado moderno, que se funda en la burocracia profesional y en el derecho racional. La igualdad de derechos, uno de los pilares de los derechos humanos como pretensiones típicas de legitimidad, depende funcionalmente de la burocratización.

Las formas de la representación evolucionan en función de la integración social. Inicialmente la representación es ejercida por un individuo, pero el incremento de la interdependencia funcional requiere hacer más complejos los mecanismos de representación: parlamentos nobiliarios o consejos reales. Los pactos feudales (como la Carta Magna) son ejemplo del modo en que la representación que ejerce el cuadro administrativo se hace internamente más compleja e incluye a más individuos que al rey. Cuando la integración social de los diferentes grupos llega a un nivel de interdependencia máxima, en la representación se incluye ya a todos los miembros de la comunidad política. Históricamente, esto se traduce en el principio de la soberanía popular. En consecuencia, el incremento de la integración social lleva a una universalización de los derechos de participación. El sufragio universal no significa desdiferenciación del sistema político, porque como dice Weber, la masa no gobierna. Puesto que la existencia de la comunidad política depende de que en su interior se diferencie un órgano que cumpla la función de representar a la comunidad, éste es siempre otra cosa que la totalidad social. Siempre hay diferencia entre representantes y representados, pues de los primeros dependen realmente las decisiones, mientras los segundos no deciden, sino que se representan que la acción de los primeros es una acción de la totalidad. Cuando el derecho al voto es universal y libre, la falta de adhesión a las decisiones políticas

es privatizada, con lo que se impide la formación de grupos insurgentes y el sistema tiene pretensiones de legitimidad más eficaces.

Por último, ninguno de los procesos conducentes a la formalización de los derechos humanos como pretensiones de legitimidad es posible si no existe un ordenamiento racionalizado y formalizado. Un derecho es racional cuando existe una relación previsible entre los argumentos que fundamentan la legislación o la aplicación judicial y la decisión normativa particular o la consecuencia jurídica que merece el supuesto de hecho concreto, y es formal cuando las decisiones jurídicas no tienen en cuenta más que caracteres generales y unívocos de los hechos, de acuerdo a un sistema conceptual elaborado por la ciencia jurídica. En otras palabras, cuando el ordenamiento se diferencia funcionalmente y se autonomiza frente a consideraciones utilitarias, políticas, éticas, de rentabilidad económica, etc. y sólo justifica las decisiones jurídicas con argumentos jurídicos.

En un desarrollo típico-ideal, el monopolio de la violencia legítima, la burocratización del cuadro administrativo, la generalización de los mecanismos de representación y la racionalización y formalización del derecho conducen al establecimiento de los derechos civiles y políticos como pretensiones típicas de legitimidad. Pero las condiciones típico-ideales no existen en la realidad. La formalización de los derechos fundamentales, por el contrario, es también el resultado de ciertos antagonismos y tensiones que atraviesan el proceso descrito.

En primer lugar, que el iusnaturalismo sea la pretensión típica de legitimidad del derecho revolucionariamente creado significa una contradicción entre el monopolio de la violencia legítima y los derechos fundamentales, porque se debe justificar la violencia para imponerlos. En tales casos es típico que las pretensiones de legitimidad no sean puramente racionales, sino que con ellas aparezcan mezclados motivos afectivos y tradicionales. Este es el caso del nacionalismo, una pretensión de legitimidad que emplean los sistemas políticos como resultado de las luchas colectivas, invocando el destino común en que los individuos ponen en riesgo

su vida en conflictos bélicos. Los gobiernos fomentan la socialización política de sus ciudadanos utilizando esta forma de legitimación. El equilibrio entre pretensiones racionales y afectivas o tradicionales de legitimación depende entonces de la historia política de cada Estado, dando lugar a uno de los factores que explican la diversa eficacia de los derechos humanos en la escena internacional. En este punto, las investigaciones de Edmund Chan e Ivo Duchacek sobre derecho constitucional comparado avalan el esquema weberiano.

En segundo lugar, el monopolio de la violencia legítima significa la posibilidad de una racionalización de la administración, pero ésta puede dirigirse en sentidos distintos. Puede dar lugar a la racionalización *material* de la administración y la justicia realizada por un príncipe patrimonial que procura la felicidad de sus súbditos por motivos utilitarios y ético-sociales, o a la racionalización *formal* mediante la imposición, hecha por juristas profesionales, del dominio de normas jurídicas universales para todos los *ciudadanos del Estado*. Que así ocurra depende de la coincidencia de diversos factores, pero básicamente éstos son el capitalismo y la formación de los juristas. Donde la burocracia no es contrapesada por intereses económicos poderosos o por un estamento de juristas socialmente poderoso, racionaliza el derecho en el aspecto material, pero destruye la técnica jurídica formal, que se presenta como carente de valor frente al sentido material de la justicia. Es el caso de la burocratización en el antiguo Egipto, en la China imperial o, en menor medida, en la Prusia de Federico Guillermo II.

Sin embargo, la racionalización económica conduce a un sistema en el que los criterios de eficiencia tienden a entrar en contradicción con las expectativas de seguridad personal de los individuos, hasta el punto de cuestionarse la legitimidad del sistema político que está en posición de garante de un derecho formal que hace posible la plena autonomía del sistema económico. En estos casos deja insatisfecha a una parte importante de la población la mera igualdad jurídica formal y la justicia y el gobierno racionales, tal como lo exigen los intereses burgueses.

Para tales masas, la legitimidad del derecho y del gobierno dependen de que estén al servicio de la nivelación de las probabilidades de vida económicas y sociales enfrente de los poseedores, lo que implica de nuevo un sentido material de la justicia. Ante la amenaza al monopolio de la violencia legítima que esto puede suponer, el sistema político reformula sus pretensiones de legitimidad introduciendo los derechos económicos, sociales y culturales. Pero en qué grado se reconozcan éstos y qué importancia tengan frente a los derechos civiles y políticos no puede resolverse con criterio general, y depende de las tensiones que en el seno de cada Estado hayan tenido lugar y cómo las luchas colectivas hayan dado lugar a grupos más o menos organizados. De aquí la variabilidad de formas de eficacia de los derechos fundamentales económicos, sociales y culturales en los diversos países.

Por último, otra forma de tensión que afecta a los derechos fundamentales proviene de la democracia de masas. Efectivamente, la racionalización de la dominación conduce a una estructura en que se diferencia la dominación burocratizada de la dominación política. Ninguna decisión en el ámbito de la dominación burocrática es legítima si no está amparada por una norma formalmente establecida, y ningún funcionario es responsable de las consecuencias negativas de su actuación si puede mostrar que ésta ha sido siempre el cumplimiento de lo normativamente establecido. Por el contrario, las decisiones políticas no dependen de normas; los electores no están obligados a justificar sus preferencias mediante ningún tipo de normas, y quienes ocupan cargos políticos no se legitiman ante el electorado por ser meros cumplidores de las normas, sino por ser capaces de marcar la dirección en que debe dirigirse la creación normativa. El equilibrio entre dominación burocrática y dominación política depende de cada caso, pero Weber detecta una tendencia hacia la burocratización excesiva que considera peligrosa porque limitaría las posibilidades de la democracia de elegir destinos alternativos en la vida política.

En síntesis, el modelo weberiano ofrece un eje central que explica la génesis sociológica de los derechos en función de cómo operan las pretensiones típicas de legitimidad en relación al grado

de racionalización del derecho, un proceso complejo que depende de la racionalización de las imágenes del mundo y de las formas de dominación. Pero además aporta factores particulares capaces de dar razón de las variaciones de la eficacia de los derechos humanos a lo largo de los diversos Estados. El modelo de Luhmann, por comparación con el weberiano, sintetiza excesivamente la explicación de los derechos fundamentales al concentrarla en un único factor causal, la diferenciación sistémica, por lo que en su marco teórico apenas cabe otra cosa que un proceso lineal unidireccional, sin tendencias divergentes, ni antagonismos ni fases regresivas. La realidad diversa de los derechos humanos a lo largo del mundo — incluyendo la diversidad entre Estados con sistemas sociales funcionalmente diferenciados muy semejantes entre sí— no parece corresponderse con la imagen de ellos que se deduce del modelo de Luhmann. Por otra parte, esta ausencia de contradicciones es la causa de que no pueda incluir los derechos económicos, sociales y culturales en su explicación de los derechos fundamentales. Por todo ello, la teoría sociológica de Weber tiene una mayor virtualidad explicativa que la hace preferible.

Por su parte, el intento de Habermas de usar la teoría weberiana como base para, reformulándola, buscar una definición distinta de los criterios de legitimación de las sociedades contemporáneas, debe inscribirse en las teorías éticas sobre la legitimidad, por lo que propiamente no debe ser considerado un modelo sociológico alternativo. He mostrado cómo Habermas, que declara tener la intención de continuar la parte que considera inacabada de la teoría weberiana, sólo emplea realmente la racionalización de las imágenes del mundo, pero no la racionalización de la dominación. No se trata de que la teoría esté inacabada sino que se ha tomado una parte por el todo. Por eso su crítica no afecta al núcleo de la teoría sociológica de los derechos humanos que he tratado de formular a partir de la base weberiana.

El paradigma weberiano sobre los derechos humanos representa, además, una base idónea para estructurar la investigación empírica sobre los mismos. Si es preciso abandonar el modelo del normativismo jurídico dominante, como concluía más arriba, y

contar con teorías sociológicas que justifiquen qué datos son relevantes y por qué, el modelo de la racionalización del derecho de Weber resuelve satisfactoriamente la mayor parte de los problemas localizados en la investigación empírica, y se ve confirmada además por las principales conclusiones que hasta hoy se han alcanzado en este campo.

El monopolio de la violencia legítima es un requisito imprescindible para la constitución de un derecho racionalizado, y éste es requisito necesario a su vez para la eficacia de los derechos humanos. La revisión que del indicador de Charles Humana hicieron Gupta, Jongman y Schmid aplicando técnicas estadísticas de discriminación es una prueba de lo acertado de esta tesis. A pesar de todas las cautelas mencionadas en su momento, lo que estos autores han mostrado es que los factores que en mayor proporción determinan la eficacia de los derechos humanos en un determinado Estado son la inexistencia de ejecuciones y desapariciones extrajudiciales, la inexistencia de detenciones sin cargos, la inexistencia de torturas, tratos crueles, inhumanos o degradantes y la existencia de garantías judiciales de la investigación policial. En síntesis, la seguridad de la integridad física y el conjunto de las garantías procesales del Estado de derecho.

Sobre la confirmación de esta conclusión de la teoría sociológica, la investigación empírica debería proseguir en adelante esta orientación para confirmar de manera más sólida sus primeros resultados. Dado que la eficacia del monopolio de la violencia legítima es directamente evaluable en términos fácticos, no ofrece especiales dificultades en cuanto a criterios de cuantificación. En este punto, además, hay ya propuestas de valoración cuantitativa, como las de Barsh o Arat, y especialmente la metodología diseñada para realizar los estudios *Minorities at risk* y *Genocides and politicides since World War II*, de Ted Robert Gurr, Barbara Harff y James R. Scarritt. En este campo, el empleo hasta ahora de un modelo normativista ha impedido resolver con claridad la cuestión de qué datos deben ser buscados. Pero además, si se intenta una verificación de todas las normas, el conjunto de los datos disponibles es siempre necesariamente incompleto, como prueban las

observaciones de George A. López, Michael Stohl, Ted R. Gurr, Barbara Harff y Keneth Bollen al respecto. El objetivo no debe ser mostrar qué eficacia tiene cada norma en particular, sino qué capacidad de mantener una dominación eficaz tiene cada Estado en función de sus pretensiones de legitimidad.

Las variaciones en la eficacia de los derechos humanos que todos los indicadores muestran tienen una de sus causas principales en el multiculturalismo, un problema éste que no puede ser resuelto desde las premisas del normativismo jurídico dominante en la investigación empírica, como ha mostrado especialmente Russell L. Barsh. El paradigma weberiano también ofrece una base para explicar estas variaciones en función del tipo de reelaboración doctrinal llevada a cabo por la clase intelectual en cada contexto cultural, en función de la historia de las luchas políticas en el seno de cada Estado reflejadas en sus respectivas constituciones, como exponen Chan y Duchacek, con el consiguiente uso de criterios nacionalistas de legitimación, y en función de los equilibrios que en cada caso se den entre las tendencias a una legitimación más burocrática o más política.

La investigación de Boli-Bennett demuestra una correlación muy clara entre el incremento en la capacidad de control social de la burocracia estatal y la protección constitucional de los derechos fundamentales. Sus conclusiones son una confirmación muy directa de las tesis weberianas, pues muestran cómo varían las pretensiones típicas de legitimidad formuladas a través del reconocimiento de derechos en la medida en que varían, racionalizándose, las formas de dominación. A pesar de que esta correlación parece una conclusión firmemente asentada, podrían continuarse las investigaciones en esta línea afinando metodológicamente los medios empleados por Boli-Bennett. Por ejemplo, analizando las estadísticas sobre gasto público en la administración general del Estado en comparación con la evolución constitucional ya descrita. O analizando la formalización de los procedimientos de reclutamiento de funcionarios frente a los sistemas basados en el contrato de status anterior, y tomándolo como un indicativo de la evolución de la dominación burocrática. Este tipo de investigacio-

nes no conducirán a conclusiones diferentes, pero permitirán contar con más datos que contrastar con el resto de la información estadísticas sobre los derechos humanos y, de su cruce, extraer nuevas conclusiones. Por otra parte, Boli-Bennett ha mostrado también la relación directa entre derechos civiles y políticos de una parte y derechos económicos, sociales y culturales de otra. Coincide en esto con Arat, y ambos avalan el cuadro evolutivo de los derechos fundamentales que he descrito a partir de la teoría weberiana de la racionalización del derecho.

Kiss, Massué y Vasak, por su parte, han confirmado con su investigación la relación directa entre racionalización y formalización del derecho y pretensiones de legitimidad legales, que se concretan en los derechos humanos. Sobre este punto, Richard Claude, director del *Comparative human rights project* y editor del libro homónimo, al sintetizar las conclusiones compartidas por los partícipes sobre las condiciones sociales de los derechos humanos se refiere de manera muy directa a la necesidad de un sistema legal racionalizado. Reproduzco sus conclusiones, a las que me sumo como cierre de mis propias conclusiones. La primera es que la condición básica para la eficacia de los derechos humanos es la existencia de un sistema legal seguro para dispensar justicia y procedimientos formalizados para la resolución de conflictos. Segunda, que un sistema legal modernizado descansa sobre estructuras legales institucionalmente diferenciadas. Tercera, que cuanto más estructuralmente diferenciado y flexible es el sistema legal (sea a través del desarrollo legislativo, de criterios procesales normativamente establecidos y de la jurisprudencia, o de ambos), tanto más capaz es el sistema de responder al incremento del cambio social y político. Cuarta, que la idea del Estado de derecho asociada con el establecimiento de límites legales sobre la acción del gobierno, si bien está sujeta a interpretaciones que varían a lo largo de los diversos países, deriva de los principios, instituciones y procedimientos que la experiencia y tradiciones de los juristas y otros operadores jurídicos ha demostrado ser importante para proteger a los individuos de la arbitrariedad gubernamental. Quinta, que mientras las políticas públicas del Estado del

bienestar sobre derechos positivos —típicamente modernas— se amplían codo a codo con la centralización administrativa, las necesidades organizativas del control de programas y de supervisión tienden a entrar en conflicto con ideas anteriores sobre el Estado de derecho y con ideas anteriores acerca de los derechos negativos, tales como la protección de la propiedad y la privacidad y la prohibición de la discriminación.

C. EJERCICIOS

1. De acuerdo con lo expuesto por Ignacio Aymerich Ojea, señale por qué es importante contar con una definición operacional de los derechos humanos en las investigaciones empíricas.

2. ¿Cuál es la utilidad de las investigaciones cuantitativas sobre los derechos humanos?

3. Ignacio Aymerich menciona varias fuentes de información sobre derechos humanos. Describa sus principales características:

Publicaciones periódicas	
Estudios monográficos	
Fuentes con caracteres de periodicidad y generalidad	

4. ¿En qué consiste la metodología empleada en los informes de Freedom House?

5. ¿Qué críticas realiza Aymerich a las fuentes de información sobre derechos humanos que están a cargo de los Estados?

6. Mencione las ventajas y desventajas de los siguientes informes:

	Ventajas	Desventajas
<i>Amnesty International Reports</i>		
<i>Human Rights in Developing Countries</i>		
<i>Country Reports on Human Rights Practices</i>		

7. ¿Cuáles son las críticas que Ignacio Aymerich realiza a los informes de Charles Humana?

8. ¿Cuáles son y en qué consisten las objeciones de Ignacio Aymerich a las fuentes cuantitativas de información sobre los derechos humanos?

9. A partir de la lectura de Aymerich Ojea, ¿qué propuestas metodológicas le haría usted a los informes especiales de la Comisión de Derechos Humanos del Distrito Federal?

10. ¿Cuáles son las principales críticas a los estudios monográficos?

D. AUTOEVALUACIÓN

1. ¿De acuerdo con Ignacio Aymerich Ojea, cuáles serían las ventajas de contar con un sistema cuantitativo de evaluación de las normas internacionales sobre derechos humanos?

- a) Permitir la documentación de violaciones a los derechos humanos de manera extensiva y hacer posible la comparación entre países en regiones específicas.
- b) Hacer posible una valoración común de la eficacia de las normas internacionales y permitir comparaciones no sólo sincrónicas sino diacrónicas.
- c) Permitir el intercambio internacional de información y documentar sistemáticamente los casos de incumplimiento de las normas internacionales.
- d) Hacer posible la documentación tanto de las violaciones a los derechos humanos como los progresos en la materia en cada país.

2. Señale uno de los requisitos que, de acuerdo con Aymerich Ojea, es necesario cubrir antes de construir un sistema de indicadores de derechos humanos.

- a) Establecer conceptualmente qué son los derechos humanos.
- b) Elegir los criterios de cuantificación.
- c) Examinar los sistemas de indicadores ya establecidos.
- d) Establecer hipótesis sobre el grado de desarrollo de los derechos humanos.

3. De acuerdo con el texto, ¿cuáles son las principales críticas a las publicaciones periódicas con tratamiento particular de la información?

- a) Echan de menos el análisis complejo y profundo de los casos graves de violaciones a los derechos humanos.
- b) No siguen las reglas de la inferencia sistémica, lo que le resta fiabilidad a los resultados.
- c) Siguen un sistema de indicadores que se concentra fundamentalmente en los derechos civiles y políticos sin considerar los derechos económicos.
- d) No permiten hacer seguimientos sistemáticos y comparativos entre Estados. Además, no cubren todo el espectro de derechos fundamentales.

4. ¿Cuáles son las principales críticas de Aymerich Ojea a los informes de Amnistía Internacional?

- a) La falta de tratamiento sistemático de la información y su interés por analizar cierto tipo de derechos en detrimento de otros.
- b) Sus criterios inadecuados para la definición, categorización y cuantificación de los derechos humanos.
- c) Su estilo narrativo y la falta de un carácter directamente sociológico.
- d) Su limitado ámbito de análisis a algunos países y la carga ideológica de sus conclusiones.

5. ¿A qué se refieren las objeciones a los sistemas de obtención de datos según lo referido por Aymerich Ojea?

- a) A la fiabilidad que puede tener la información cuando se consideran las dificultades de acceso a la misma.
- b) Al sesgo ideológico de las fuentes de información.
- c) A la heterogeneidad de dichos sistemas cuando se aplican en diversos países.
- d) A la rigurosidad de los instrumentos de recolección de información.

6. En relación con el texto responda, ¿cuáles son los criterios de rigor técnico de los sistemas de indicadores sobre derechos humanos?

- a) La efectividad y eficacia, la relevancia, la mensurabilidad, la validez, la equivalencia y la agregación.
- b) La claridad y la conceptualización, la precisión, la relevancia, la pertinencia, la oportunidad y la equivalencia.
- c) La definición y la medida, la precisión, la consistencia, la validez, la equivalencia y agregación.
- d) La economía, la confiabilidad, la oportunidad, la definición y la medida, la precisión y la consistencia.

7. Con base en la lectura de Aymerich Ojea señale a qué se refieren las objeciones del multiculturalismo.

- a) Al carácter arbitrario e ideológico de los derechos humanos.
- b) A la imposibilidad de la universalidad de los derechos humanos en la posmodernidad.
- c) A que los sistemas de indicadores no consideran las particularidades raciales de cada país.
- d) Al problema de la diversidad cultural frente a la univocidad de las normas.

8. Con miras a la generación de un sistema de indicadores, según el texto de Aymerich Ojea, ¿cómo deben conceptualizarse los derechos humanos?

- a) Como normas.
- b) Como valores.
- c) Como hechos.
- d) Como principios.

9. De acuerdo con Aymerich Ojea, ¿cuáles son las desventajas de las fuentes de información limitadas a un determinado país y en un periodo breve?

- a) Que no permiten extraer conclusiones generales ni hacer un seguimiento temporal.
- b) Que no permiten una exposición narrativa de los resultados ni se pueden extraer conclusiones normativas.
- c) Que tienen una fuerte carga ideológica y no profundizan en el detalle de los casos.

- d) Que sus sistemas de indicadores no son replicables y sus conclusiones no son generalizables.

10. Con base en la lectura de Ignacio Aymerich Ojea, señale las características de los informes de Freedom House y Charles Humana.

- a) Generalidad, periodicidad, profundidad en el manejo de la información.
- b) Falta de comparación, sesgo político, definiciones ambiguas.
- c) Generalidad, periodicidad, tratamiento sistemático de la información.
- d) Falta de comparación, sesgo político, particularidad.

Módulo IV.

La estadística en la investigación
en derechos humanos

OBJETIVOS ESPECÍFICOS

- Proporcionar conocimientos específicos para plantear proyectos de investigación, cuyos resultados puedan ser generalizables a todo el Distrito Federal.
- Conocer y utilizar las herramientas necesarias para la elaboración de indicadores en Derechos Humanos.

A. SÍNTESIS

Como se ha visto a lo largo de estos módulos, uno de los objetivos de la investigación cuantitativa es ofrecer resultados que permitan ser generalizables a toda la población. En el caso específico de la investigación en derechos humanos, la estadística es una herramienta fundamental para realizar investigaciones sobre el estado de los derechos humanos más allá de la descripción de casos específicos o de las investigaciones monotemáticas. En este módulo se describirá primero la selección de la muestra y, después veremos otros conceptos básicos de estadística descriptiva.

I. Conceptos básicos de la estadística descriptiva

a. Selección de la muestra

No en todas las investigaciones es necesario seleccionar una muestra; sin embargo, cuando el tamaño de la población es muy grande se requiere hacer esta selección pues ello resulta mucho más económico en cuestiones de tiempo y recursos.

b. Selección de la unidad de análisis

En cualquier investigación social es necesario seleccionar la unidad de análisis. Aquí el interés se centra en los sujetos, objetos, sucesos, instituciones o contextos de estudio. Esto depende del planteamiento inicial de la investigación.⁶ Por ejemplo, si el objetivo es describir la violencia hacia las mujeres en el Distrito Federal, lo más factible sería interrogar a un grupo de mujeres. Además, podría servir interrogar a los esposos de las mujeres. Escoger entre las mujeres o sus esposos depende del diseño de la investigación.

Por tanto, para seleccionar una muestra, lo primero que hay que hacer es definir la unidad de análisis (recuérdese el módulo II). Al respecto, es necesario señalar que la población sobre la que se van a recolectar datos depende del planteamiento del problema investigado y de los alcances del estudio.⁷

c. Tipos de muestra

Las muestras se dividen en dos grandes ramas: las muestras no probabilísticas y las muestras probabilísticas. En las muestras no probabilísticas, la elección de los elementos no

⁶ Hernández, Roberto et al., "Selección de la muestra", en *Metodología de la investigación*, México, McGraw-Hill, 2003, p. 300.

⁷ *Ibid.*, pp. 301-302.

depende de la probabilidad, sino de causas relacionadas con las características de la investigación o del investigador. En las muestras probabilísticas, por el contrario, todos los elementos de la población tienen la misma posibilidad de ser escogidos.⁸ Elegir entre un tipo de muestra u otro depende del planteamiento del problema y de los objetivos de la investigación.

Cuadro iv.1. Comparación entre tipos de muestras

	Muestra probabilística	Muestra no probabilística
Definición	Subgrupo de la población en la que todos los elementos tienen la misma probabilidad de ser escogidos	Subgrupo de la población en la que la elección de los elementos no depende de la probabilidad sino de las características de la población
Tipos	Muestreo aleatorio simple Muestro aleatorio sistemático Muestro aleatorio estratificado Muestro aleatorio por conglomerados	Muestreo de cuotas

II. Tipos de muestreo

a. Muestreo aleatorio simple

En este tipo de muestreo la selección de los elementos de la muestra se hace en una sola etapa. En la práctica equivale a censar o utilizar el censo de la población objeto del estudio, para sacar, después, al azar los elementos que van a formar parte de la muestra (Rodríguez, p. 544).

Así, por ejemplo, si se quiere extraer una muestra de 216 elementos, de un universo formado por el personal de la Comisión de Derechos Humanos del Distrito Federal, habría que obtener el censo de los cerca de 550 empleados y, a partir de aquí, hacer la selección de los individuos.

El muestreo aleatorio simple se aplica fundamentalmente en investigaciones sobre poblaciones pequeñas y plenamente identificables, sobre todo si constituyen universos específicos y diferenciados (Rodríguez, p. 544).

b. Muestreo aleatorio sistemático

El muestreo aleatorio sistemático es una variante del método anterior, en la que difiere la forma de selección de las unidades muestrales pues con este tipo de muestro la selección lleva una secuencia. Con este muestreo se simplifica considerablemente la selección, pero existe el riesgo de introducir sesgos en la muestra al elegir los elementos de forma periódica. Esto ocurre cuando el universo está ordenado en función de determinados criterios que pueden inducir a que la selección sistemática recaiga en elementos que no son representativos de la heterogeneidad del universo (Rodríguez, p. 547). Ello ocurre cuando el censo (base de la selección) lleva un orden específico.

⁸ *Ibid.*, p. 306.

c. Muestreo aleatorio estratificado

La partición o fraccionamiento de la población en subdivisiones constituye la base del muestreo estratificado. En él cada unidad del universo pertenece a una sola subdivisión, y el conjunto de éstas, constituye la urdimbre sobre la que opera el proceso de muestreo (Rodríguez, p. 548).

Por ejemplo, si se quisiera seleccionar una muestra de los empleados de la Comisión de Derecho Humanos del Distrito Federal, se podría estratificar a la población por cargo (fuerza de trabajo, mandos medios, mandos superiores) de este modo se realizaría una selección por cada cargo.

Las principales ventajas de este tipo de muestreo son las siguientes: permite tratar de forma independiente a cada uno de los estratos. Esto facilita la utilización de diferentes métodos de muestreo así como la estimación, por separado, de ciertas subpoblaciones constituidas en dominios de estudio. Además, con este método se pueden reducir las varianzas de las estimaciones de la muestra; aumenta la precisión de las estimaciones y facilita la coordinación de los trabajos de campo (Rodríguez, p. 548).

d. Muestreo por cuotas

Para evitar las rigideces de los sistemas anteriores se recurre muchas veces a los muestreos por cuotas. Es un sistema de muestreo muy utilizado por los institutos dedicados a los estudios de opinión y de mercado. Consiste en facilitar al entrevistador el perfil de las personas que tiene que entrevistar en cada una de las secciones o de las rutas en que se va a hacer la aplicación de las entrevistas. De esta forma, la selección aleatoria de las unidades últimas de muestreo se sustituye por una selección dejada, en parte, al criterio del investigador, con la condición de que se cumplan determinados requisitos, fijados en las cuotas. Éstas, sacadas de los datos del censo, garantizan que las personas entrevistadas, bajo determinados aspectos, reflejen exactamente el universo de la muestra (Rodríguez, p. 564).

Es necesario destacar que *el muestreo de cuotas no es un muestreo probabilístico* y por ello no se pueden hacer estimaciones rigurosas. Debido a esto, cuando se trata de encuestas gubernamentales y, en general, de encuestas que se toman como base para decisiones importantes, hay que recurrir a muestreos totalmente probabilísticos a pesar del costo y del tiempo que se necesita para la realización de un riguroso acopio de la información en el trabajo de campo. Es la garantía de poder medir el grado de precisión de las estimaciones (Rodríguez p. 566). En el siguiente cuadro se resumen las características de los tipos de muestreo.

Cuadro iv.2. Tipos de muestreo

Tipo de muestreo	Características principales
Aleatorio simple	Se obtienen al azar los elementos que formarán parte de la muestra.
Aleatorio sistemático	Es una variante del muestreo aleatorio simple. Los elementos de la muestra se obtienen de manera periódica.
Aleatorio estratificado	La población se divide en estratos, de donde se obtienen los elementos de la muestra.
Muestreo por cuotas	Muestreo no probabilístico. Consiste en facilitar al entrevistador el perfil de las personas que tiene que entrevistar.

e. Tamaño de la muestra

El *tamaño de la muestra* y la *precisión de las estimaciones* son conceptos inseparables, ya que las variaciones en uno de ellos afectan al otro y viceversa. Si aumenta el tamaño de la muestra lo hace el nivel de precisión de las estimaciones y, si se quiere conseguir una mayor precisión, es necesario modificar el tamaño de la muestra (Rodríguez, p. 567). Ésta es una regla general, pero se deben tomar en cuenta las acotaciones señaladas en el módulo II.

Una vez que se ha seleccionado la muestra, se procede a la recolección de los datos. Ello implica diseñar y aplicar instrumentos de medición. Este tema no será abordado en este momento, por ahora consideramos necesario aclarar algunos conceptos básicos de estadística descriptiva para el análisis de los datos.

III. Procedimiento para el análisis cuantitativo de los datos

Una vez que los datos se han codificado, transferido a una matriz y guardado en un archivo, el investigador procede a analizarlos. De acuerdo con Roberto Hernández, el análisis depende de tres factores:

- a) El nivel de medición de las variables
- b) La manera cómo se formularon las hipótesis
- c) El interés del investigador

Por ejemplo, el análisis que se aplica a una variable nominal no puede ser el mismo que el análisis de una variable por intervalos. Al respecto se sugiere recordar los niveles de medición vistos en el módulo II.

Una vez que se cuenta con una base de datos, el investigador busca, primero, describir sus datos y posteriormente realizar análisis estadísticos para relacionar sus variables, es decir, primero se realiza el análisis de *estadística descriptiva* para cada una de sus variables y luego se describe la relación entre éstas. Los tipos o métodos de análisis cuantitativo o estadístico son variados y se comentarán a continuación; pero hay que señalar que el análisis no es indiscriminado, cada método tiene su razón de ser y un propósito específico; no deben hacerse más análisis de los necesarios. La estadística no es un fin en sí misma, sino una herramienta para analizar los datos. Los principales análisis cuantitativos que pueden efectuarse son:

- Estadística descriptiva para las variables tomadas individualmente.
- Puntuaciones Z.
- Razones y tasas.
- Cálculos y razonamientos de estadística inferencial.

En este módulo sólo se tratarán los conceptos básicos de estadística descriptiva; pues ellos son los mínimos requeridos para analizar datos cuantitativos en una base de datos. Estos conceptos son: distribución de frecuencias, medidas de tendencia central (moda, mediana, media) y medidas de variabilidad (rango y desviación estándar).

a. Distribución de frecuencias

La distribución de frecuencias es un conjunto de puntuaciones ordenadas en sus respectivas categorías (Hernández, p. 587). El siguiente cuadro muestra un ejemplo de distribución de frecuencias.

Cuadro iv.3. Distribución de frecuencias

Variable: calificación en el examen del servicio profesional	
Categoría	Frecuencias
98	1
92	3
86	6
85	7
84	5
76	4
73	4
69	2
65	1
Total	33

Las *distribuciones de frecuencias* pueden completarse agregando las *frecuencias relativas* y las *frecuencias acumuladas*. Las frecuencias relativas son los porcentajes de casos en cada categoría y las frecuencias acumuladas son las que se van acumulando en cada categoría, desde la más baja hasta la más alta. Para cada una de las variables de la investigación se obtiene su distribución de frecuencias (Hernández, p. 589). Además de las frecuencias, deben calcularse las *medidas de tendencia central* y de *variabilidad* o *dispersión* (Hernández, pp. 589, 595).

b. Medidas de tendencia central

Las *medidas de tendencia central* son puntos en una distribución, los valores medios o centrales de ésta, y nos ayudan a ubicarla dentro de la escala de medición. Las principales medidas de tendencia central son tres: *moda*, *mediana* y *media*. Es importante destacar que *el nivel de medición de la variable determina cuál es la medida de tendencia central apropiada* (Hernández, p. 595).

La *moda* es la categoría o puntuación que ocurre con mayor frecuencia (Hernández, p. 595) En el cuadro anterior la moda es 85.

La *mediana*, es el valor que divide la distribución por la mitad. Esto es, la mitad de los casos caen por debajo de la mediana y la otra mitad se ubica por encima de la mediana. La mediana refleja la posición intermedia de la distribución. La *mediana es una medida de tendencia central propia de los niveles de medición ordinal, por intervalos o de razón*. No tiene sentido con variables nominales, porque en este nivel no hay jerarquías, ni noción de encima o debajo (Hernández, p. 595). En nuestro ejemplo, la mediana es 84.

La *media* es la medida de tendencia central más utilizada y puede definirse como el promedio aritmético de una distribución. Es la suma de todos los valores dividida entre el número de casos. *Es una medida solamente aplicable a mediciones por intervalo y de razón.* Carece de sentido para variables medidas en un nivel nominal u ordinal (Hernández, p. 597).

Cuadro iv.4. Medidas de tendencia central

Moda	El valor que ocurre con mayor frecuencia.
Mediana	El valor que divide la distribución por la mitad.
Media	El promedio aritmético de una distribución.

c. Medidas de variabilidad

Además de las medidas de tendencia central, el análisis de datos cuantitativos requiere considerar las medidas de variabilidad, éstas, de acuerdo con Hernández Sampieri, *indican la dispersión de los datos en la escala de medición* y responden a la pregunta ¿dónde están diseminadas las puntuaciones o los valores obtenidos?

Las medidas de tendencia central son valores en una distribución y las medidas de la variabilidad son intervalos, designan distancias o un número de unidades en la escala de medición. Las medidas de variabilidad más utilizadas son *rango*, *desviación estándar* y *varianza* (Hernández, p. 599).

El *rango*, también llamado *recorrido*, es la diferencia entre la puntuación mayor y la puntuación menor, e indica el número de unidades en la escala de medición necesario para incluir los valores máximo y mínimo. Cuanto más grande sea el rango, mayor será la dispersión de los datos de una distribución (Hernández, p. 599).

Si retomamos el ejemplo de las calificaciones del servicio profesional:

98 - 92 - 86 - 85 - 84 - 76 - 73 - 69 - 65

el rango será: 98 - 65= 33.

La *desviación estándar* es el promedio de la desviación de las puntuaciones con respecto a la media. Esta medida se expresa en las unidades originales de medición de la distribución. Se interpreta en relación con la media. Cuanto mayor sea la dispersión de los datos alrededor de la media, mayor será la desviación estándar (Hernández, p. 599). Se simboliza con una σ o $\frac{1}{2}$ y su fórmula sintetizada es:

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{N}}$$

La desviación estándar se interpreta como “cuánto se desvía, en promedio, de la media un conjunto de puntuaciones”. La desviación estándar sólo se utiliza en variables medidas por intervalos o de razón (Hernández, p. 602).

Por ejemplo, supongamos que un investigador obtuvo para su muestra de una media de calificación de examen del Servicio Profesional en Derechos Humanos 85 y una desviación estándar de 8. La interpretación es que la calificación de la muestra se desvía, en promedio, 5 puntos respecto de la media.

Finalmente, la *varianza* es la desviación estándar elevada al cuadrado. Es un concepto estadístico muy importante, ya que muchas de las pruebas cuantitativas se fundamentan en él. Sin embargo, con fines descriptivos se utiliza preferentemente la desviación estándar (Hernández, p. 602).

Cuadro iv.5. Medidas de variabilidad

Rango	Es la diferencia entre la puntuación mayor y la puntuación menor.
Desviación estándar	Es el promedio de la desviación de las puntuaciones con respecto a la media.
Varianza	Es la desviación estándar elevada al cuadrado.

B. MATERIAL DE ESTUDIO

Bibliografía obligatoria

Rodríguez, Jacinto, "Tipos de muestreo", en *Métodos de muestreo*, Madrid, Centro de Investigaciones Sociológicas, 2001, pp. 21-46 **[Lectura 1]**.

———, "Tamaño de la muestra", en *Métodos de muestreo*, Madrid, Centro de Investigaciones Sociológicas, 2001, pp. 47-63 **[Lectura 2]**.

Hernández, Roberto, *et al.*, "Análisis de datos", en *Metodología de la investigación*, México, McGraw-Hill, 2003, pp. 493-622 **[Lectura 3]**.

Chernobilsky, Lilia, "El uso de la computadora como auxiliar en el análisis de datos cualitativos", en *Estrategias de investigación cualitativa*, Barcelona, Gedisa, 2006, pp. 239-273 **[Lectura 4]**.

Lectura 1

2

Tipos de muestreo

Las muestras probabilísticas se caracterizan porque en ellas «cada elemento de la población tiene una probabilidad conocida y no nula de ser seleccionado»¹. Dado que la distribución de probabilidades suele ser conocida², en este tipo de muestras es posible proceder a la inferencia estadística, trasladando los datos muestrales a la población. Para ello es necesario hacer las correspondientes estimaciones, en base a los datos muestrales, para proceder, a continuación, al cálculo de su variabilidad, es decir, al cálculo de los errores estándar.

Para que las muestras sean probabilísticas, y, por tanto, reúnan las características señaladas, se han de seguir determinadas normas en el proceso de extracción de la muestra, garantía de la aleatoriedad prevista en el cálculo de probabilidades. Estas normas, ajustadas a las peculiaridades de cada universo, dan origen a los cuatro métodos básicos de muestreo que recubren distintos procesos de extracción de muestras. Estos métodos son los siguientes:

- Muestreo aleatorio simple.
- Muestreo aleatorio sistemático.
- Muestreo aleatorio estratificado.
- Muestreo aleatorio por conglomerados.

De estos métodos los tres primeros suelen ser monoetápicos. En ellos los elementos de la muestra se eligen directamente en una sola etapa. El último es polietápico y las unidades muestrales no son los elementos de la población, sino conglomerados de elementos.

Los cuatro procedimientos de muestreo reseñados son distintos y su elección debe responder a las peculiaridades del universo. Esto quiere decir que, en la mayoría de los casos, no es indiferente utilizar un método u otro, ya que cada uno de ellos suele ser el más indicado para cada situación. No

¹ Leslie KISH, *Muestreo de encuestas*, México D.F., Editorial Trillas, 1972, p. 41.

² Los modelos fundamentales de distribución de probabilidades son el binomial y el normal, ambos bien conocidos.

obstante, en la práctica, a veces hay que utilizar no los métodos más precisos sino los más viables, dada la carencia de información sobre el universo o los elevados costes de determinadas aplicaciones. Así, la utilización del muestreo aleatorio simple para la realización de una encuesta nacional dirigida a la población adulta podría ser, en teoría, la más adecuada, pero, en la práctica, resulta inviable, y lo mismo ocurre en otros muchos casos. Esto quiere decir que, en general, cabe la posibilidad de sustituir unos métodos por otros, mediante la realización de un nuevo diseño y, en especial, la adecuación del tamaño de la muestra al método que se aplique. Esto es así porque, para un error de muestreo dado, varía el tamaño de la muestra en función del tipo de muestreo que se utilice.

Anteriormente se ha señalado que hay cuatro métodos básicos de muestreo. Sin embargo, no se utilizan, necesariamente, de forma aislada. Es más, muy frecuentemente se aplican, simultáneamente, en la extracción de una misma muestra. Esta técnica responde a necesidades de diseño y se concreta en la combinación de métodos por etapas o por submuestras. En el primer caso, se puede utilizar un método distinto en cada fase del muestreo y, en el segundo, en cada submuestra. Para la extracción de una muestra nacional, dirigida a la población en general, se puede proceder, primero, mediante un muestreo estratificado por conglomerados para terminar con la elección de las unidades últimas de muestreo, mediante muestreo aleatorio simple. En una muestra a médicos, después de estratificar por especialidades, se puede extraer la muestra, en las especialidades menos numerosas, mediante muestreo aleatorio simple, y, en el resto, después de estratificar nuevamente, bien por provincias o por lugar donde desarrolla el médico su actividad: hospital, ambulatorio, domicilio.

Hechas las anotaciones anteriores, se va a entrar en el estudio de los diferentes métodos de muestreo. Antes, no obstante, se quiere hacer una ligera referencia al muestreo por cuotas del que se hablará al final del capítulo. Se trata de un muestreo no probabilístico, cuyos resultados son óptimos, en ocasiones. De ahí que se estime necesario introducirlo en este capítulo para explicar en qué consiste y en qué circunstancias sus estimaciones permiten aproximarse a la realidad.

Muestreos probabilísticos

Se describen, a continuación, los diferentes tipos de muestreo señalando sus características básicas y la forma de aplicarlos en muestreos monoetápicos. Más adelante, al hablar del submuestreo y en diversos ejemplos de este trabajo, se explica la forma de utilizarlos en muestreos polietápicos.

Muestreo aleatorio simple

En este tipo de muestreo la selección de los elementos de la muestra se hace en una sola etapa, directamente y sin reemplazamiento³. En la práctica equivale a censar o utilizar el censo de la población objeto del estudio, para sacar, después, al azar, los elementos que van a formar parte de la muestra. Así, si se quiere extraer una muestra de 1.500 elementos, de un universo formado por los médicos en activo en España, habría, en primer lugar, que obtener el censo de los 120.000 médicos, que aproximadamente existen, y, a partir de aquí, se haría la selección de los individuos. Ésta se puede hacer mediante tablas de números aleatorios u otro procedimiento similar, como el bombo, que también garantiza la aleatoriedad.

Las tablas de números aleatorios simplifican considerablemente el proceso. Presentan múltiples combinaciones de números extraídos al azar (véase cuadro 2.1) y, con ellas, a partir de cualquier fila o columna, se toman tantos números consecutivos, como el de elementos de la muestra. Los números sacados al azar indican, sobre un universo numerado, los elementos que se incorporan a la muestra y que, por tanto, deben ser entrevistados.

La utilización del bombo para la selección aleatoria de los elementos de una muestra es prácticamente nula por las dificultades que conlleva. Proviene del mundo de las loterías y queda como ejemplo de la equiprobabilidad y, además, se sigue utilizando cuando se trata de elegir a elementos de universos muy reducidos. En universos con un número relativamente importante de elementos resulta muy costoso hacer la relación individualizada, para su posterior inclusión en el bombo.

El muestreo aleatorio simple se aplica, fundamentalmente, en investigaciones sobre poblaciones pequeñas, plenamente identificables, máxime si constituyen universos específicos y diferenciados. Así una muestra sobre médicos o sobre cualquier otro colectivo diferenciado y pequeño se debería extraer mediante muestreo aleatorio simple. Es más, en muchas ocasiones hay que acudir a este tipo de muestreo empezando por la creación de un censo del universo objeto de estudio⁴.

³ Sin reemplazamiento quiere decir que, una vez elegido un elemento de la población para incluirlo en la muestra, se excluye del universo para que no pueda volver a ser elegido. Las técnicas de la estadística inferencial son adecuadas para el muestreo con reemplazo en el que todos los elementos del universo tienen probabilidad igual. En las muestras sin reemplazo, que son las que se utilizan, las técnicas son igualmente aplicables sin distorsiones cuando se trata de universos grandes «cien veces más grandes que las muestras». (J. L. WEINBERG y K. P. GOLDBERG, *Estadística básica para las ciencias sociales*, México, Interamericana, 1982, p. 220.) En caso contrario, en universos pequeños, la probabilidad de los diferentes elementos varía tanto más cuanto más pequeño es el universo.

⁴ Los estudios sobre las familias pobres de Reims, sobre los grandes accidentados de la carretera y sobre guarderías, citados en la introducción general, recurren a este sistema.

CUADRO 2.1
NÚMEROS ALEATORIOS

10 09 73 25 33	76 52 01 35 86	34 67 35 48 76	80 95 90 91 17	39 29 27 49 45
37 54 20 48 05	64 89 47 42 96	24 80 52 40 37	20 63 61 04 02	00 82 29 16 65
08 42 26 89 53	19 64 50 93 03	23 20 90 25 00	15 95 33 47 64	35 08 03 36 06
99 01 90 25 29	09 37 67 07 15	38 31 13 11 65	88 67 67 43 97	04 43 62 76 59
12 80 79 99 70	80 15 73 61 47	64 03 23 66 53	98 95 11 08 77	12 17 17 68 33
66 06 57 47 17	34 07 27 68 50	36 69 73 61 70	65 81 33 98 85	11 19 92 91 70
31 06 01 08 05	45 57 18 24 06	35 30 34 26 14	86 79 90 74 39	23 40 30 97 32
85 26 97 76 02	02 05 16 56 92	68 66 57 48 18	73 05 38 52 47	18 62 38 85 79
63 57 33 21 35	05 32 54 70 48	90 55 35 75 48	28 46 82 87 09	83 49 12 56 24
73 79 64 57 53	03 52 96 47 78	35 80 83 42 82	60 93 52 03 44	35 27 38 84 35
98 52 01 77 67	14 90 56 86 07	22 10 94 05 58	60 97 09 34 33	50 50 07 39 98
11 80 50 54 31	39 80 82 77 32	50 72 56 82 48	29 40 52 42 01	52 77 56 78 51
83 45 29 96 34	06 28 89 80 83	13 74 67 00 78	18 47 54 06 10	68 71 17 78 17
88 68 54 02 00	86 50 75 84 01	36 76 66 79 51	90 36 47 64 93	29 60 91 10 62
99 59 46 73 48	87 51 76 49 69	91 82 60 89 28	93 78 56 13 68	23 47 83 41 13
65 48 11 76 74	17 46 85 09 50	58 04 77 69 74	73 03 95 71 86	40 21 81 65 44
80 12 43 56 35	17 72 70 80 15	45 31 82 23 74	21 11 57 82 53	14 38 55 37 63
74 35 09 98 17	77 40 27 72 14	43 23 60 02 10	45 52 16 42 37	96 28 60 26 55
69 91 62 68 03	66 25 22 91 48	36 93 68 72 03	76 62 11 39 90	94 40 05 64 18
09 89 32 05 05	14 22 56 85 14	46 42 75 67 88	96 29 77 88 22	54 38 21 45 98
91 49 91 45 23	68 47 92 76 86	46 16 28 35 54	94 75 08 99 23	37 08 92 00 48
80 33 69 45 98	26 94 03 08 58	70 29 73 41 35	53 14 03 33 40	42 05 08 23 41
44 10 48 19 49	85 15 74 79 54	32 97 92 65 75	57 60 04 08 81	22 22 20 64 13
12 55 07 37 42	11 10 00 20 40	12 86 07 46 97	96 64 48 94 39	28 70 72 58 15
63 60 64 93 29	16 50 53 44 84	40 21 95 25 63	43 65 17 70 82	07 20 73 17 90
61 19 69 04 46	26 45 74 77 74	51 92 43 37 29	65 39 45 95 93	42 58 26 05 27
15 47 44 52 66	95 27 07 99 53	59 36 78 38 48	82 39 61 01 18	33 21 15 94 66
94 55 72 85 73	67 89 75 43 87	54 62 24 44 31	91 19 04 25 92	92 92 74 59 73
42 48 11 62 13	97 34 40 87 21	16 86 84 87 67	03 07 11 20 59	25 70 14 66 70
23 52 37 83 17	73 20 88 98 37	68 93 59 14 16	26 25 22 96 63	05 52 28 25 62
04 49 35 24 94	75 24 63 38 24	45 86 25 10 26	61 96 27 93 36	65 33 71 24 72
00 54 99 76 54	64 05 18 81 59	96 11 96 38 96	54 69 28 23 91	23 28 72 95 29
35 96 31 53 07	26 89 90 93 54	33 35 13 54 62	77 97 45 00 24	90 10 33 93 33
59 80 80 83 91	43 42 72 68 42	83 60 94 97 00	13 02 12 48 92	78 56 52 01 06
46 05 88 52 36	01 39 09 22 86	77 28 14 40 77	93 91 08 36 47	70 61 74 29 41
32 17 90 05 97	87 37 92 52 41	05 56 70 70 07	86 74 31 71 57	85 39 41 18 38
69 23 46 14 06	20 11 74 52 04	15 95 66 00 00	18 74 39 24 23	97 11 89 63 38
19 56 54 14 30	01 75 87 53 79	40 41 92 15 85	66 67 43 68 06	84 96 28 25 07
45 15 51 49 38	19 47 60 72 46	43 66 79 45 43	59 04 79 00 33	20 82 66 95 41
94 86 43 19 94	36 16 81 08 51	34 88 88 15 53	01 54 03 54 56	05 01 45 11 76
59 58 00 64 78	75 56 97 88 00	88 83 55 44 86	23 76 80 61 56	04 11 10 80 08
38 50 80 73 41	23 79 34 87 63	90 82 29 70 22	17 71 90 42 07	95 95 44 99 53
30 69 27 06 68	94 68 81 61 27	56 19 68 00 91	82 06 76 34 00	05 46 26 92 00
65 44 39 56 59	18 28 82 74 37	49 63 22 40 41	08 33 76 56 76	96 29 99 08 36
27 26 75 02 64	13 19 27 22 94	07 47 74 46 06	17 98 54 89 11	97 34 13 03 58

En el caso de poblaciones grandes, ofrece mayores dificultades. En primer lugar, es difícil realizar un listado de todo el universo; si ello fuera posible, el proceso para extraer la muestra resultaría tedioso; finalmente, la dispersión de la muestra podría ser muy importante, lo que implicaría aumento considerable de los costes y del tiempo del trabajo de campo. Así, la realización de una muestra de 5.000 elementos dirigida a la población española en edad de trabajar, y utilizando el muestreo aleatorio simple, sería impensable. En primer lugar habría que reconstruir y hacer el listado del universo, con más de 22 millones de unidades. Después, habría que seleccionar, aleatoriamente, los elementos de la muestra para, finalmente, proceder al trabajo de campo, es decir, a la recogida de la información. La primera dificultad radicaría en la reconstrucción del universo, operación prácticamente imposible salvo para el INE que tiene informatizado el censo electoral⁵. Aunque se dispusiera del censo, el proceso de extracción de la muestra sería tedioso por el tamaño de los ficheros, y la dispersión de las entrevistas sería muy grande. Junto a una cierta concentración en las grandes capitales y aglomeraciones se produciría la dispersión en pequeños pueblos, escondidos en los lugares más recónditos. La muestra podría ser, técnicamente, perfecta, pero los costes y el tiempo empleado serían mucho más elevados que por otros procedimientos. De ahí que, cuando hay que hacer muestras sobre poblaciones muy grandes, se recurra a otros métodos de muestreo o a la combinación de varios.

El muestreo aleatorio simple se presenta como el prototipo de muestreo por su sencillez y la facilidad para calcular los errores de muestreo⁶. Esto obedece a que es un muestreo monoetápico y las diferentes unidades del universo tienen la misma probabilidad de formar parte de la muestra⁷. Por otra parte es el muestreo de referencia para calcular lo que KISH llama «efecto del diseño». Éste se define como «la razón de la varianza de la estimación obtenida a partir de la muestra más compleja a la varianza de la estimación obtenida a partir de una muestra aleatoria simple del mismo número de unidades»⁸. En consecuencia, el muestreo aleatorio simple aparece como un patrón para medir la eficacia de los muestreos estratificados y por conglomerados ya que a través del «efecto del diseño» se establece la comparación entre éstos y aquél. En general, y adelantando resultados de

⁵ El INE tiene informatizado el censo electoral. Como la edad de trabajar es a partir de los 16 años, tendría que añadir el grupo 16 y 17 años a dicho censo.

⁶ Esto explica que en la ficha técnica de muchas encuestas se ofrezca el cálculo de errores correspondientes a un muestreo aleatorio simple aunque el tipo de muestreo aplicado tenga poco que ver con éste. Así sucede, con mucha frecuencia, en la ficha técnica que acompaña a las encuestas publicadas en la prensa española.

⁷ Hay que recordar la ligera restricción que supone el tratarse de muestreo sin reemplazamiento, tal como se señala en la nota 3.

⁸ Tomado de William G. COCHRAN, *Técnicas de muestreo*, México D.F., Editorial Continental, 1984, p. 119.

otros capítulos, en el muestreo estratificado la varianza de la estimación suele ser menor que en el muestreo aleatorio simple y, en el muestreo por conglomerados, mayor.

Muestreo aleatorio sistemático

Se puede considerar como una variante del método anterior, en la que difiere la forma de selección de las unidades muestrales. Para ello, se halla el coeficiente de elevación: $\frac{N}{n}$ ⁹ y se elige al azar un número no superior al mismo, que es el que indica el punto de arranque de la selección. A partir de aquí, al número elegido se le suma, sucesivamente, el coeficiente de elevación, dando lugar a los números que pasan a formar parte de la muestra. En la encuesta a médicos, el proceso sería el siguiente: coeficiente de elevación $\frac{N}{n} = \frac{120.000}{1.500} = 80$; primer número al azar, no superior a 80, el 24, por ejemplo. La selección sería 24, $24 + 80 = 104$, $104 + 80 = 184$; y así sucesivamente.

Con este procedimiento se simplifica considerablemente la selección, pero existe el riesgo de introducir sesgos en la muestra al elegir los elementos de forma periódica. Esto ocurre cuando el universo está ordenado en función de determinados criterios que pueden inducir a que la selección sistemática recaiga en elementos que no son representativos de la heterogeneidad del universo. En la selección sistemática de una muestra sobre listas de 20 individuos en que los 10 primeros fueran varones y los 10 últimos mujeres, si el coeficiente de elevación fuera 20, siempre saldrían sólo hombres o sólo mujeres, pero no habría posibilidades de que en la muestra estuvieran representados ambos sexos. SÁNCHEZ CRESPO acusa este peligro con las siguientes palabras: «La selección sistemática tiene las ventajas de extender la muestra sobre toda la población, ser de fácil aplicación, y conseguir un efecto similar al de la estratificación si las unidades se han ordenado previamente siguiendo un cierto criterio como, por ejemplo, colocando al principio de la lista las unidades densamente pobladas y continuando progresivamente hasta llegar a las menos pobladas. Por el contrario, puede introducir sesgos debido al hecho de que cada unidad en la muestra es seleccionada con una periodicidad constante igual a K»¹⁰.

Para evitar este peligro hay que hacer un estudio previo del universo y de los listados que lo contienen, incluso, si fuera necesario, «desordenarlos» *ab absurdo* para preparar la selección de los elementos muestrales. Se hace

⁹ N = Número de unidades del universo; n = número de elementos de la muestra.

¹⁰ J. L. SÁNCHEZ CRESPO, *Muestreo de poblaciones finitas aplicado al diseño de encuestas*, Madrid, INE, 1973, p. 93.

referencia a una técnica utilizada, con frecuencia, cuando se realizan muestreos sistemáticos a partir de ficheros de individuos, almacenados en soporte magnético. En este caso, se puede romper el orden actual del fichero: orden alfabético, de antigüedad, etcétera, a partir de cualquier criterio como, por ejemplo, las 6 últimas cifras del DNI de los individuos que componen el universo, ordenadas al revés. Así, el número del DNI 28.150.824 se quedaría en 150.824 y ordenado al revés sería 428.051. Posteriormente, se procedería al ordenamiento de todo el universo y a la aplicación del muestreo sistemático ¹¹.

Muestreo aleatorio estratificado

La partición o fraccionamiento de la población marco del estudio en subdivisiones constituye la base del muestreo estratificado. En él cada unidad del universo pertenece a una sola subdivisión, estrato, y el conjunto de estratos, subdivisiones del universo, constituyen la urdimbre sobre la que opera el proceso de muestreo.

Las principales ventajas de este tipo de muestreo, ampliamente utilizado, son las siguientes: Permite tratar de forma independiente a cada uno de los estratos. Esto facilita la utilización de diferentes métodos de muestreo así como la estimación, por separado, de ciertas subpoblaciones constituidas en dominios de estudio. Con este método se pueden reducir las varianzas de las estimaciones muestrales. Aumenta la precisión de las estimaciones. Facilita la coordinación de los trabajos de campo.

La estratificación en los procesos de muestreo ofrece una gran flexibilidad ya que posibilita la utilización simultánea, en una misma muestra, de distintos métodos de muestreo en función de las necesidades del diseño y de la información disponible sobre cada estrato. Facilita, además, la estimación por separado de los distintos estratos, siempre que la muestra esté adecuadamente dimensionada. Esto abre el camino a la desagregación del universo en dominios diferenciados, posibilitando el análisis de cada uno de ellos.

Por otra parte, la estratificación, en tanto en cuanto reúna en cada estrato a unidades homogéneas entre sí y heterogéneas en relación con las de los otros estratos, contribuye a reducir la varianza de las estimaciones, lo que se traduce en ganancias en precisión, es decir, en disminución de los errores de muestreo. Esto quiere decir que, si se pudiera hacer la estratificación ideal, este tipo de muestreo sería más preciso que el aleatorio simple, y el «efecto del diseño» sería inferior a 1. En la práctica, esto significa que, para un error muestral dado, es necesario hacer menor número de

¹¹ Existen ficheros de alumnos, de socios de un club, miembros de una sociedad médica, etcétera, informatizados, y, a partir de ellos, se puede aplicar el sistema que se ha reseñado.

entrevistas si se utiliza el muestreo estratificado que si se utiliza el aleatorio simple. Por otra parte, si se conocen las varianzas por estrato¹² se pueden aplicar fracciones de muestreo distintas en cada estrato de acuerdo con el valor de la varianza. En estratos más homogéneos se podrían utilizar fracciones de muestreo más pequeñas que en estratos más heterogéneos, tal como viene haciendo el INE en sus diseños muestrales¹³.

Finalmente, la estratificación puede facilitar la coordinación de los trabajos de campo, ya que permite separar la muestra por estratos. En este caso, cada estrato o cada varios estratos pueden formar áreas operativas distintas bajo la responsabilidad del Jefe de zona que coordina y controla la recogida de la información.

Haciendo referencia ahora a los criterios de estratificación, hay que tener en cuenta las variables que se van a utilizar y el número de estratos resultante. En cuanto a lo primero SÁNCHEZ CRESPO escribe: «las variables utilizadas para la estratificación deberían estar correlacionadas con las variables objeto de la investigación»¹⁴. Es más, si fuera posible, debería recurrirse a variables de alta correlación que garantizaran la homogeneidad de los estratos. Así, para hacer un diseño muestral para el estudio de la pobreza, se podría partir de variables de ingresos, igual que, para hacer un estudio sobre el nivel de conocimientos de la población, habría que utilizar variables referidas a los estudios realizados.

Por lo que respecta al número de estratos, se puede señalar con KISH que «no es aconsejable llevar la división muy lejos», ya que «los estratos muy pequeños contribuyen muy poco a las ganancias de la estratificación» y «la formación de sólo unos cuantos estratos producirá típicamente la mayoría de las ganancias posibles a partir de una variable»¹⁵.

Aunque la teoría, en este campo, es conocida, sin embargo, resulta difícil estratificar por variables altamente correlacionadas, puesto que no se suelen conocer las varianzas. En general, y como criterio más viable, se suele recurrir a variables espaciales: comunidades autónomas, provincias, municipios, etcétera, o a subdivisiones inherentes al universo en estudio. Así, el universo alumnos de universidad, se puede estratificar por universidades, facultades, especialidades, cursos, etcétera; a los profesores de EGB, por tipo de enseñanza —pública o privada—, comunidades autónomas, tamaño del centro, etcétera. En estos casos también es posible controlar el número

¹² Es difícil conocer de antemano la varianza por estratos, referida a variables concretas. No obstante, se puede disponer de indicadores que permitan determinar, en general, cuáles son los estratos más homogéneos y cuáles los más heterogéneos desde distintas perspectivas: distribución de rentas, cohesión social, etcétera.

¹³ El INE considera que las zonas rurales son más homogéneas que las urbanas y, por eso, carga la muestra en estas últimas.

¹⁴ J. L. SÁNCHEZ CRESPO, *Curso intensivo de muestreo en poblaciones finitas*, 3.ª edición, Madrid, INE, 1984, p. 41.

¹⁵ Leslie KISH, *Muestreo de encuestas*, México D.F., Editorial Trillas, 1972, p. 131.

de estratos, siendo el investigador el que tiene que decidir las variables que va a utilizar para estratificar y el número de estratos resultantes. En cualquiera de los casos, y como norma general, hay que dejar claro que la estratificación tiene un límite y que, incluso operativamente, la multiplicación excesiva de estratos puede complicar el diseño, ya que aparecerán estratos vacíos o casi vacíos.

Finalmente, y antes de pasar a explicar con un ejemplo cómo se opera en el caso del muestreo estratificado, se quiere señalar que este tipo de muestreo es el más utilizado tanto en muestreos monoetápicos como en polietápicos, en los que suele acompañar a otras formas de muestreo. De ahí que se debiera dedicar, en cada caso, una especial atención a la estratificación, clave de las ventajas de este sistema de muestreo.

Volviendo sobre el desarrollo de una muestra estratificada se va a explicar, a continuación, cómo se opera en el muestreo estratificado tipo, el monoetápico, para hablar, en otro epígrafe, de la utilización de la estratificación en los muestreos polietápicos. Tomando el proceso de muestreo desde el principio, se puede concretar en el siguiente ejemplo ¹⁶:

Universo: Trabajadores de Iberia.

Estratificación: Dada la gran heterogeneidad de funciones que existen dentro de la compañía se pensó en subdividir al personal en estratos homogéneos. Para ello, y a partir del catálogo de puestos de trabajo de la empresa, un grupo de personas conocedoras de la misma hizo los siguientes estratos:

- Directivos y técnicos de grado superior.
- Técnicos de grado medio y técnicos auxiliares.
- Administrativos.
- Especialistas y servicios auxiliares.
- Pilotos y operadores de radio.
- Auxiliares de vuelo.

Se hicieron, por lo tanto, seis estratos, cada uno de los cuales reunía a personas de intereses profesionales parecidos y distintos de los intereses de los otros grupos. Aparte de las ventajas de esta estratificación para mejorar la precisión de las estimaciones ¹⁷, también permitió la realización de seis submuestras, tal como aparece en el cuadro 2.2. Dichas submuestras son representativas de cada una de las poblaciones de las que fueron extraídas.

Tamaño de la muestra: El tamaño de la muestra se fijó en 1.500 entrevistas, buscando un compromiso entre representatividad del universo y de

¹⁶ Se trata de una encuesta realizada a los trabajadores de Iberia en septiembre-octubre de 1987 (núm. 1.692 del Banco de Datos del CIS).

¹⁷ Tal como se ve en el capítulo 4 la estratificación cuando se hace con variables que correlacionan con el objeto de la investigación disminuye los errores muestrales para un tamaño dado de la muestra.

CUADRO 2.2
ESTUDIO DE IBERIA. DISTRIBUCIÓN DE LAS ENTREVISTAS POR ESTRATOS

Submuestra	Universo ‰	Entrevistas		Coeficiente de ponderación
		‰	Número	
10. Directivos y Técnicos de Grado Superior.....	35	133	200	0,263
11. Directivos	5	19	28	
12. Téc. Grado Superior	30	114	172	
20. TGM y Téc. Auxiliares	75	167	250	0,449
21. TGM	22	49	73	
22. Téc. Auxil.....	53	118	177	
30. Administrativos.....	255	200	300	1,275
31. Cuerp. Sup. Admtvo.	11	8	12	
32. Cuerpo Gral. Admtvo.....	244	192	288	
40. Esp. y Serv. Aux.....	413	167	250	2,473
50. Pilotos y OTV	71	133	200	0,755
51. Pilotos	52	97	146	
52. OTV	19	36	54	
60. Auxiliares de vuelo	151	200	300	0,755
Total.....	1.000	1.000	1.500	1.500

FUENTE: CIS.

cada uno de los grupos. La muestra es plenamente representativa del universo trabajadores de Iberia y, con menor precisión, de cada uno de los grupos en que se dividió el universo.

*Afijación*¹⁸: Tal como se ve en el cuadro 2.2, la afijación no fue proporcional porque se buscaban submuestras representativas. De ahí que se fijara

¹⁸ Este concepto se estudia en el capítulo 3 al que nos remitimos.

el número mínimo de entrevistas por estrato en 200 y el máximo en 300, dada la limitación de 1.500 entrevistas en total.

*Ponderación*¹⁹: Cada una de las submuestras se puede tabular por separado. Si se desea tabular la muestra conjuntamente es necesario proceder a la ponderación porque las tasas de muestreo son distintas en cada submuestra. Se incluyen los coeficientes de ponderación en la última columna del cuadro 2.2.

Selección de los elementos de la muestra: La selección se hizo en una sola etapa utilizando la técnica del muestreo sistemático. Para ello se partió de los listados nominales de la población de cada estrato, dentro del cual se hizo la selección de la submuestra correspondiente. Primero se eligió, al azar, un número no superior al coeficiente de elevación y, posteriormente, se eligieron los números siguientes, con sólo añadir al anterior el valor del coeficiente de elevación. Una alternativa a esta solución hubiera sido extraer los elementos de cada submuestra de la lista nominal de individuos, utilizando el muestreo aleatorio simple.

Muestreo aleatorio por conglomerados

En los tipos de muestreo analizados la unidad muestral la componen los elementos de la población objeto de la investigación. Existe otro procedimiento de muestreo, también aleatorio, en el que la unidad muestral no son los individuos, sino un conjunto de individuos que, bajo determinados aspectos, se puede considerar que forman una unidad. A este tipo de muestreo se le denomina muestreo por conglomerados. Las unidades hospitalarias, los departamentos universitarios, una caja de productos terminados, constituyen ejemplos de conglomerados naturales. Existen otros tipos de conglomerados no naturales formados también por conjuntos de elementos, que pueden recibir el mismo tratamiento muestral que los casos anteriores como, por ejemplo, las urnas electorales. Hay que destacar, además, por su gran utilización, los conglomerados definidos «como áreas o partes bien delimitadas del terreno, de modo que todas las unidades últimas correspondientes al área sean las que constituyen el conglomerado»²⁰. En estos casos, en vez de hablar de muestreo por conglomerados, se suele hablar de muestreo por áreas, aunque la técnica a utilizar es la misma.

Definidas las unidades muestrales, los conglomerados, la forma de operar para extraer la muestra depende, entre otras cosas, del tamaño del mismo. Cuando éste es pequeño, se puede proceder de forma similar a la señalada en los epígrafes anteriores, pero entrevistando a todos los elementos del conglomerado. Si los conglomerados son muy grandes resulta im-

¹⁹ *Ibid.*

²⁰ F. AZORIN POCH, *Curso de muestreo y aplicaciones*, Madrid, Aguilar, 1972, p. 151.

posible la aplicación de las entrevistas a todos y cada uno de sus elementos por lo que hay que recurrir al submuestreo.

El muestreo por conglomerados monoetápico, que es el que ahora interesa, se reduce en la práctica a la aplicación de los muestreos aleatorio simple, sistemático o estratificado, pero tomando como unidad de muestreo los conglomerados y no los individuos, tal como se puede ver con un ejemplo tomado de la realidad ²¹. La muestra que sirvió de base para este estudio se diseñó mediante muestreo aleatorio por conglomerados con las siguientes características:

Universo: Las unidades —clases— de primero, segundo y tercero de BUP de los centros escolares del municipio de Madrid.

Estratificación: Los centros se estratificaron según su tamaño y su pertenencia al sector público o privado. Del cruce de estas dos variables se obtuvieron los 9 estratos que aparecen en el cuadro 2.3.

Tamaño de la muestra: Se fija en 133 unidades —clases—. Éstas se distribuyeron de forma proporcional al peso de cada estrato. Las 133 unidades responden a unos 4.000 alumnos.

Selección de los elementos de la muestra: A partir de un listado de unidades —clases— por centros escolares y estratos se extrajeron las unidades que deberían formar parte de la muestra, mediante muestreo aleatorio simple. Los cuestionarios se aplicaron a todos los alumnos de las clases seleccionadas.

Las muestras por conglomerados están más concentradas frente a una dispersión mucho mayor en el resto de los métodos de muestreo. Por eso

CUADRO 2.3
DISTRIBUCIÓN DE LOS ALUMNOS DE BUP DEL MUNICIPIO DE MADRID POR ESTRATOS

<i>Tamaño</i>	<i>Tipo de centro</i>			<i>Total</i>
	<i>Público</i>	<i>Privado religioso</i>	<i>Privado no religioso</i>	
Pequeño	7	154	159	320
Mediano	90	94	120	304
Grande	350	9	17	376
Total	447	257	296	1.000

NOTA: La distribución está expresada en tantos por mil.

FUENTE: Explotación de datos facilitados por el profesor Juan Luis RECIO.

²¹ Se trata de un estudio realizado en 1988, con la financiación de la Cruz Roja Española, sobre la prevención del uso de drogas. El estudio lo dirigió el profesor Juan Luis RECIO.

los muestreos por conglomerados simplifican los procesos de extracción de la muestra, reduciendo costes y tiempo en su ejecución. En este tipo de muestreo el conocimiento del universo objeto del estudio se simplifica, dado que basta con tener un listado de conglomerados, del cual se van a seleccionar los que van a formar parte de la muestra. Son éstos, los seleccionados, los que van a servir de base para recoger la información que se pretende obtener mediante el muestreo. En resumen, se puede decir que en el muestreo por conglomerados la unidad muestral no son los individuos sino un conjunto de ellos; la muestra resultante está mucho más concentrada y se opera con un nivel de información sensiblemente inferior.

Muestreo polietápico

El submuestreo es una técnica que consiste en dividir en un cierto número de unidades más pequeñas las unidades de población utilizadas en el diseño muestral. Es, por tanto, un caso especial del muestreo aleatorio por conglomerados, en el que la unidad final de muestreo no van a ser los conglomerados sino subdivisiones de éstos.

Se utiliza, como norma general, cuando el número de elementos del conglomerado es elevado y, no sólo por la dificultad de entrevistar a todos y cada uno de los elementos del mismo ²², sino, y sobre todo, por la necesidad de conseguir estimaciones más precisas. Como dice KISH: «Cuando se compara una muestra de conglomerados con una muestra de elementos del mismo tamaño n , podemos esperar una varianza mayor pero un costo menor, en la muestra de conglomerados. En general, la mayor distribución de una muestra de elementos en la población produce mayor precisión, pero es más costosa. Aunque esta situación no es universal, es común y puede llevarnos al submuestreo. El submuestreo resulta de la búsqueda de un punto de equilibrio entre los dos efectos conflictivos de la conglomeración de los elementos en la economía del diseño; la disminución en el grado de conglomeración puede disminuir la varianza notablemente, sin incurrir en un aumento proporcional en el costo» ²³.

El submuestreo, por tanto, afecta al grado de conglomeración, dispersa la muestra y, aunque puede repercutir negativamente en los costes, influye positivamente en la precisión de las estimaciones ya que puede hacer disminuir la varianza total.

²² En un muestreo por áreas, en el que se tome como unidad la sección censal, que tiene en torno a los 2.000 elementos —individuos—, es impensable que hubiera que entrevistar a todos y cada uno de los individuos.

²³ Leslie KISH, *Muestreo de encuestas*, México D.F., Editorial Trillas, 1972, p. 189.

Etapas del muestreo

En este tipo de muestreo, la selección de las unidades últimas, a las que se ha de aplicar la muestra, se realiza por etapas sucesivas. «La selección de conglomerados en primer lugar y de elementos a continuación, requiere de dos etapas de selección. Este método se puede extender rápidamente a más etapas. El muestreo polietápico consiste en una jerarquía de diferentes tipos de unidades; cada unidad de primera etapa se divide, o es potencialmente divisible, en unidades de segunda etapa, etcétera»²⁴.

Las unidades de primera etapa se llaman *unidades primarias de muestreo* y a las unidades subsecuentes, *unidades de segunda etapa*, etcétera; a las unidades finales se las suele llamar *unidades últimas de muestreo*. Aunque en este tipo de muestreo las unidades seleccionadas en las etapas sucesivas son de menor tamaño, el proceso no llega, necesariamente, hasta la selección de elementos individuales. La selección se puede detener en la extracción de conglomerados tales como hogares, clases de alumnos, unidades hospitalarias, etcétera, que, en este caso, serían las unidades últimas de muestreo.

En cuanto a la información requerida para el desarrollo de cada una de las etapas hay que señalar que ésta varía y se reduce a medida que el proceso avanza. «Se requerirá un marco en cada etapa para las unidades que han sido seleccionadas en ella. Inicialmente, se requiere un marco mediante el cual se puedan definir y seleccionar las unidades de primera etapa. Para la segunda etapa de selección se requiere un marco mediante el cual las unidades secundarias se puedan definir dentro de las de primera etapa que han sido seleccionadas»²⁵. Esto representa una gran ventaja, desde el punto de vista operativo, ya que se puede trabajar con información reducida y, en cualquiera de los casos, referida sólo a las unidades que intervienen en cada etapa del proceso de selección.

Avanzando en el estudio de los muestreos polietápicos de conglomerados, hay que destacar la importancia que, para el desarrollo de la muestra, tiene la estructura del conglomerado. Se hace referencia, en concreto, a su homogeneidad o heterogeneidad y a su tamaño. Lo primero influye en el número de conglomerados que se han de seleccionar, mientras que lo segundo, el tamaño, en el método de selección de los mismos.

Cuando los conglomerados son homogéneos basta con coger, en primera etapa, pocas unidades de cada uno de ellos y, por lo tanto, hay que aumentar el número de conglomerados que se seleccionan. Si, por el contrario, los conglomerados de primera etapa son heterogéneos, se pueden tomar más elementos por conglomerado y, consiguientemente, hay que seleccionar

²⁴ *Ibid.*

²⁵ *Ibid.*, p. 190.

un número menor. Guardando estos principios, también se podría abordar el problema operando con conglomerados relativamente grandes que, previsiblemente, serían más heterogéneos. Ello permitiría seleccionar muchos elementos por conglomerado y, por tanto, pocos conglomerados. Dado que esta situación no siempre es viable habrá que mantenerla como una nueva posibilidad de actuar, pero sin renunciar a los principios antes reseñados. ABAD DE SERVÍN resume así ambas soluciones: «Desde el punto de vista de la precisión, en un esquema de submuestreo interesa que las unidades primarias sean altamente heterogéneas respecto a la(s) característica(s) en estudio. Una manera de lograr esta heterogeneidad consiste en definir a la unidad primaria de manera que resulte grande relativamente, para así usar el criterio que es válido en muchas ocasiones y que establece que, a mayor lejanía, mayor diferencia o asemejanza entre unidades, familias pobres y familias ricas. Si ocurre esto, pueden elegirse para la muestra a pocas unidades primarias. Si por el contrario, los conglomerados primarios resultan ser muy homogéneos, se hace necesario aumentar la fracción de muestreo de las primarias o, en otras palabras, aumentar el tamaño de muestra de ellas, para percatarse y tomar en consideración la alta variabilidad entre primarias. En esta situación, dentro de las unidades primarias muestrales se elegirán pocas unidades secundarias, ya que los elementos en la misma primaria tenderán a parecerse»²⁶.

Tamaño de los conglomerados y proporcionalidad

Los conglomerados no siempre tienen un tamaño similar y esto afecta al proceso de selección de la muestra, ya que no deben tener la misma probabilidad de formar parte de la misma conglomerados con pesos poblacionales distintos. En estos casos se suele recurrir a otras técnicas de las que se hablará más adelante. Dado que el problema se plantea, frecuentemente, en el muestreo por áreas, se va a estudiar en este contexto, al filo del diseño de una muestra de estas características²⁷. Se incluyen las fases del muestreo y las soluciones que se suelen dar, cuando los conglomerados son de distinto tamaño. El ejemplo que se va a utilizar es el siguiente:

Universo: Población de la provincia de Madrid de 18 años y más.

Tamaño: 1.000 entrevistas.

Afijación: Proporcional.

Procedimiento de muestreo: Estratificado por conglomerados con selección de las unidades de primera etapa de forma proporcional, de segunda

²⁶ A. ABAD DE SERVÍN y L. A. SERVIN ANDRADE, *Introducción al muestreo*, México D.F., Limusa, 1978, p. 190.

²⁷ Se trata sólo de explicar las técnicas que se utilizan cuando los conglomerados son de distinto tamaño. Para ello, se hace una breve referencia a las fases del muestreo.

etapa de forma aleatoria simple y, las de última etapa, por rutas y tablas aleatorias.

Estratificación: En primer lugar se procede a la estratificación, operación prácticamente inseparable de los muestreos por conglomerados. En este caso se hace por tamaño de hábitat y presenta los resultados del cuadro 2.4.

CUADRO 2.4
POBLACIÓN DE LA PROVINCIA DE MADRID ESTRATIFICADA POR TAMAÑO DE HÁBITAT

<i>Estratos: tamaños del hábitat</i>	<i>Población en tanto por mil</i>
1. Menos de 2.000 habitantes	17,8
2. De 2.001 a 10.000 habitantes	32,2
3. De 10.001 a 50.000 habitantes	58,1
4. De 50.001 a 100.000 habitantes	68,8
5. De 100.001 a 400.000 habitantes	183,4
6. Más de un millón	639,7
	1.000,0

FUENTE: Explotación de datos del censo de población.

Selección de las unidades de muestreo: A cada estrato le corresponde un número de entrevistas igual al tanto por mil que representa sobre la población total, dado que se trata de un muestreo proporcional y de una muestra de 1.000 entrevistas. Fijándonos en el estrato 3, la extracción de la muestra seguiría estos pasos.

La primera etapa consistiría en la selección de los municipios en que deben realizarse las 58 entrevistas del estrato. Para que exista una adecuada dispersión de la muestra se podrían extraer 3 municipios, haciendo en el primero 20 entrevistas y 19 en los restantes. La selección de estos tres municipios se ha de hacer con probabilidad proporcional al peso de la población de cada municipio, mediante el empleo de una tabla de números aleatorios y operando sobre los datos agrupados de forma similar a como aparecen en el cuadro 2.5.

En la segunda etapa se elegirían las áreas de menor tamaño, de las que se van a extraer, en una tercera etapa, los elementos individuales que se van a entrevistar. Aquí las técnicas pueden ser varias y para exponerlas se va a continuar desarrollando el ejemplo anterior.

Se supone que el último número que salió en la tabla aleatoria, al hacer la selección de primera etapa, fue el 912, y, por tanto, fue seleccionado Valdemoro (véase cuadro 2.5). Como en este municipio hay que hacer 19 entrevistas, éstas se pueden repartir en dos bloques de 10 y de 9 entrevistas,

respectivamente. Para elegir las dos áreas en que se van a hacer estas entrevistas lo más correcto parece recurrir a las secciones censales, unidades territoriales utilizadas en casi todos los países y que, en teoría, en España, tienen alrededor de 2.000 habitantes. A partir de aquí, la selección de estas dos secciones se puede hacer de distinta forma, pero para ello es necesario conocer la distribución de las secciones censales del municipio (véase cuadro 2.6).

CUADRO 2.5

MUNICIPIOS DEL ESTRATO 3. CIFRAS ABSOLUTAS, POBLACIÓN ACUMULADA Y TANTO POR MIL

<i>Código municipio</i>	<i>Municipios</i>	<i>Población</i>	<i>Población acumulada</i>	<i>Tantos por mil</i>
13	Aranjuez	36.687	36.687	232,43
14	Arganda	23.872	60.559	383,67
40	Ciempozuelos	10.076	70.635	447,51
45	Colmenar Viejo	29.495	100.130	634,38
47	Collado Villalba	20.396	120.526	763,60
127	Rozas de Madrid (Las)	20.818	141.344	895,49
161	Valdemoro	16.496	157.840	1.000,00

FUENTE: Explotación de datos del censo de población.

CUADRO 2.6

SECCIONES CENSALES DE VALDEMORO. CIFRAS ABSOLUTAS, POBLACIÓN ACUMULADA Y TANTO POR MIL

<i>Secciones</i>	<i>Población</i>	<i>Población acumulada</i>	<i>Tantos por mil</i>
Distrito 1			
1.....	982	982	60
2.....	1.021	2.003	121
3.....	1.982	3.985	242
4.....	1.062	5.047	306
5.....	1.102	6.149	373
6.....	1.995	8.144	494
Distrito 2			
1.....	1.256	9.400	570
2.....	1.875	11.275	683
3.....	1.497	12.772	774
4.....	634	13.406	812
5.....	1.462	14.868	901
6.....	1.628	16.496	1.000

FUENTE: Explotación de datos del INE.

Una primera técnica de extracción de las dos secciones sería la selección aleatoria simple, en la que a todas las secciones se da la misma probabilidad de formar parte de la muestra. Este sistema no exige cálculos complementarios, pero no tiene en cuenta que, de hecho, las secciones no suelen ser de igual tamaño. Una segunda técnica sería similar a la utilizada para la selección de las unidades de primera etapa. Se trata de la selección aleatoria proporcional, que podría realizarse con los datos acumulados en la última columna del cuadro 2.6. Finalmente, cabe la posibilidad de actuar sobre el marco muestral en un intento de formar conglomerados —áreas— de tamaño similar. En nuestro caso se podrían refundir en una las secciones 1 y 2 del distrito 1; la 4 y 5 del mismo distrito y la 1 y 4 del distrito 2. Con ello se habrían reducido a 9 las áreas —secciones— y, además, se habrían homogeneizado considerablemente los tamaños de las mismas.

De las técnicas aplicadas para la selección de las unidades de segunda etapa, la última, la que hace referencia a la reestructuración de los conglomerados —áreas— para convertirlos a tamaños similares, resulta muy costosa. Por eso no se utiliza y porque, además, hay soluciones alternativas que ofrecen buenos resultados. Por su parte, la selección aleatoria simple ofrece garantías de equiprobabilidad si las secciones tienen tamaños similares, pero no en caso contrario. De ahí que su utilización debiera reservarse a los casos en que las diferencias de tamaño no sean notables. Finalmente, la extracción de los conglomerados mediante selección aleatoria proporcional responde a una técnica desarrollada por HANSEN y HURWITZ²⁸ en la que cada conglomerado —sección, área, etcétera— tiene probabilidad proporcional a su tamaño de formar parte de la muestra. Por eso es el más utilizado de los tres sistemas reseñados.

Dada la importancia que tiene el empleo de la técnica adecuada en el diseño de muestras, el Servicio de Muestreo del CIS está realizando diferentes investigaciones sobre la incidencia de diseños alternativos en la precisión de las estimaciones. Se trata de introducir modificaciones en un diseño base y probar cuál de ellas permite ajustar mejor las estimaciones.

Haciendo referencia al tema que aquí interesa, se trabajó con una muestra de 500 elementos, referida a la provincia de Álava. El diseño fue similar al que aquí se ha desarrollado para la provincia de Madrid, pero a la hora de elegir las unidades de segunda etapa, se utilizaron dos métodos; en un primer diseño, el aleatorio proporcional, y, en un segundo, el aleatorio simple. Después, y con cada uno de los diseños, se estimó la estructura de edades de la población de la provincia, fácilmente contrastable puesto que existen los datos reales, deducidos del padrón. El resultado de la experiencia se ofrece en el cuadro 2.7.

La columna C refleja, en términos relativos, la diferencia entre las esti-

²⁸ M. H. HANSEN y W. N. HURWITZ, «On the theory of sampling from finite populations», *Ann. Math. Stat.*, 14, 1943, pp. 333-362.

CUADRO 2.7

ESTRUCTURA DE LA POBLACIÓN DE LA PROVINCIA DE ÁLAVA Y ERRORES RELATIVOS RESULTANTES DE LA APLICACIÓN DE UNA MUESTRA CON DISTINTOS DISEÑOS

Estructura de edades (años)	Estructura real	Estimaciones		Errores relativos	
		A	B	C	D
16-19	9,08	9,34	9,33	2,86	2,75
20-24	11,00	11,54	11,55	4,91	5,00
25-29	10,60	10,89	10,41	2,74	-1,79
30-34	10,00	9,52	9,20	-4,70	-7,91
35-39	10,00	9,70	9,45	-3,00	-5,5
40-44	8,88	8,66	8,69	-2,48	-2,14
45-49	7,45	7,43	7,44	-0,27	-0,13
50-54	7,98	7,89	7,87	-1,13	-1,38
55-59	6,79	6,93	6,99	2,06	2,95
60-64	5,55	5,52	5,67	-0,54	2,16
65-69	4,12	4,19	4,35	1,70	5,58
70-74	3,48	3,49	3,71	0,29	6,61
75 y más	5,07	4,83	5,25	-4,73	3,55
	100,00				

NOTA: Estructura real: es la que se deduce de la explotación del Padrón de 1986.

- A. Estimaciones que se deducen de una muestra con selección aleatoria proporcional.
- B. Estimaciones resultantes de una muestra con selección aleatoria simple.
- C. Diferencias entre las estimaciones en A y la estructura real, expresadas en porcentajes de la estructura real.
- D. Diferencias entre las estimaciones en B y la estructura real, expresadas en porcentajes de la estructura real.

FUENTE: Investigación sobre validación de muestras. CIS.

maciones obtenidas con la muestra en que se aplicó la selección aleatoria proporcional y la estructura real. La columna D refleja, también en términos relativos, la diferencia entre las estimaciones obtenidas a partir de la selección aleatoria simple y la estructura real. Hechas estas anotaciones hay que resaltar que tanto las estimaciones que corresponden a la columna A como las de la columna B son bastante precisas²⁹, a pesar del alto número de categorías en que se ha dividido la estructura de edades. No obstante, las estimaciones de la columna A son más precisas ya que como se observa

²⁹ La precisión de las estimaciones es mayor cuando el error relativo es más pequeño. Para dar una idea de la precisión se recuerda que, hasta hace poco, el INE, en la introducción a la publicación de la EPA, decía que se podrían considerar como estimaciones válidas aquellas cuyo error relativo no superara el 10 %.

en C, hay cinco valores con un error relativo inferior al 2 % y, en ninguno de los casos, el error relativo supera el 5 %. Las diferencias en la columna D son bastante más elevadas, en conjunto, ya que sólo hay tres valores con error relativo inferior al 2 % y hay cuatro en que se supera el 5 %.

En consecuencia, se confirma, una vez más, que el tema de la proporcionalidad no es baladí y, por lo tanto, hay que tenerlo en cuenta a la hora de realizar los correspondientes diseños.

Selección de las unidades últimas de muestreo

En el ejemplo desarrollado, realización de una muestra en Madrid, se ha llegado hasta la selección de las unidades de segunda etapa, las secciones censales. Falta por elegir, aleatoriamente, y con probabilidades iguales, a los individuos dentro de las secciones seleccionadas. Aquí, los sistemas utilizados, con mayor o menor fortuna, se multiplican, por lo que parece necesario proceder a una ordenación de los mismos ³⁰.

En primer lugar, hay que hablar del sistema utilizado por el Instituto Nacional de Estadística en sus encuestas dirigidas a hogares: EPA, Presupuestos familiares, etcétera. En ellas, una vez elegida la sección, el Instituto procede a la selección de las unidades últimas de muestreo, a partir de listados sobre la composición de la sección: viviendas, hogares, individuos. Esto permite utilizar métodos de selección más rigurosos, ya que se tiene más información sobre el marco muestral gracias a los datos facilitados por los censos de habitantes. Por otra parte, sus muestras actúan sólo sobre 3.000 secciones de las 33.000 existentes, lo que permite una continua actualización de las mismas para tener al día el marco muestral. Este sistema sólo es utilizado por el INE y por alguno de los institutos de estadística de las Comunidades Autónomas porque sólo ellos tienen acceso a la información completa del censo de habitantes.

Otros institutos de investigación, entre ellos el CIS, se basan en las secciones censales como unidades de segunda etapa y, a partir de ellas, se hace la elección de los elementos de la muestra. Ésta se puede hacer de múltiples formas pero, como sistemas más generalizados, se utilizan dos. En ambos se hacen alrededor de 10 entrevistas por sección, por lo que hay que seleccionar un número alto de secciones; las entrevistas se realizan necesariamente en las secciones seleccionadas y solamente en las viviendas.

El primer sistema a que se hace alusión es el de las rutas aleatorias dentro de la sección. Elegido el comienzo de la ruta, el entrevistador ha de

³⁰ El estudio se centra en el muestreo por áreas porque las técnicas son específicas y variadas y no se abordan en otro apartado. Para la selección de las unidades últimas, en otro tipo de muestreo por conglomerados, se suele recurrir a alguno de los sistemas reseñados al hablar de los muestreos monoetápicos.

seguir las calles, según normas muy concretas que indican los giros a derecha e izquierda, eligiendo solamente los edificios terminados en los números que se seleccionan para cada estudio. Una vez en el edificio, si tuviera más de una vivienda, la elección de la escalera, del piso, de la vivienda en el piso y de la persona a entrevistar dentro de la vivienda se hace siguiendo hojas de contacto, como la que se adjunta. En ésta, el cruce de la primera columna con la primera fila, va marcando la selección aleatoria que corresponde. Por ejemplo si el edificio tiene 10 plantas y se trata de la cuarta entrevista de la ruta, hay que elegir la planta 5; si en la vivienda hay 8 personas de dieciocho años y más y es la entrevista cuarta, hay que elegir la 6, etcétera.

El segundo sistema consiste en el estudio detallado y previo de las secciones antes de aplicar las entrevistas. El coordinador de la zona levanta un croquis de la sección, calcula aproximadamente el número de viviendas y las refleja en el croquis y, a partir de ahí, fija la ruta y la frecuencia de selección de los entrevistados: una de cada 100 viviendas, por ejemplo. Posteriormente, saca, al azar, un número aleatorio y es éste el que marca el comienzo de la ruta que se ha de seguir. Una vez en la vivienda, la selección de los individuos se hace como en el caso anterior.

En este método hay un estudio previo de la sección, realizado por pocas personas —los coordinadores—, lo que permite unificar criterios y, además, se aporta una hoja de ruta, muy estudiada, que facilita la inspección de los trabajos de campo. El resultado de su aplicación es bueno, tal como se ha constatado en diferentes estudios sobre predicción de voto ³¹.

Estos sistemas de elección son aleatorios y no necesitan un conocimiento del marco muestral tan amplio como en el caso del INE. Sin embargo, su aplicación también ofrece dificultades porque la persona elegida, aleatoriamente, para ser entrevistada impone a la muestra una gran rigidez: no está en casa y volverá tarde, está de viaje, no quiere contestar, no se abre la puerta de la vivienda, etcétera. El éxito radica en la rigurosa aplicación de las normas referentes a la aleatoriedad de la selección y en evitar el recurso fácil a las sustituciones. Para ello será necesario hacer, en los casos más difíciles, hasta tres visitas a la vivienda con lo que se garantiza un alto porcentaje de respuestas.

Finalmente, se quiere hacer alusión a un método de selección utilizado por muchas consultoras y que parte de marcos muestrales más incompletos. En estos casos, una vez elegidos los municipios donde se van a realizar las entrevistas, se pasa a definir el número de rutas, tantas como sean necesarias para hacer en cada una de ellas de 8 a 10 entrevistas, aproximadamen-

³¹ El autor de este trabajo realizó, para la Academia de Humanismo Cristiano de Santiago de Chile, un diseño muestral basado en esta técnica. Se viene aplicando en Chile desde febrero de 1988 y ha dado resultados óptimos tanto en los sondeos previos al Plebiscito como en los realizados con motivo de las elecciones generales de 1989.

Muestreo por cuotas

Para evitar las rigideces de los sistemas anteriores, se recurre muchas veces a los muestreos por cuotas. Es un sistema de muestreo ampliamente utilizado por los institutos dedicados a los estudios de opinión y de mercado. Se aplica en la última etapa y consiste en facilitar al entrevistador el perfil de las personas que tiene que entrevistar en cada una de las secciones o de las rutas en que se va a hacer la aplicación de las entrevistas. De esta forma, la selección aleatoria de las unidades últimas de muestreo se sustituye por una selección dejada, en parte, al criterio del entrevistador, con la condición de que se cumplan determinados requisitos, fijados en las cuotas. Éstas, sacadas de los datos del censo, garantizan que las personas entrevistadas, bajo determinados aspectos, reflejan exactamente el universo de la muestra.

Los criterios a tener en cuenta a la hora de diseñar una muestra de cuotas hacen referencia, básicamente, a las variables que se van a utilizar. Deben ser pocas, de fácil aplicación en el conjunto de la muestra y en cada una de las rutas o secciones en que se van a hacer las entrevistas, y, sobre todo, deben relacionarse fuertemente con las variables de la encuesta. Sin embargo, estos criterios se suelen cumplir sólo parcialmente por la dificultad de llevarlos a cabo. En la práctica sólo se utilizan cuotas de sexo y edad que, en la mayoría de los casos, están poco relacionadas con las variables de la encuesta³², pero que son de más fácil obtención. Algunas veces se suele introducir una tercera variable, referida a nivel de estudios, ocupación, ingresos, etcétera, pero esta práctica es poco frecuente porque introduce cierta rigidez en la aplicación de las entrevistas, justamente lo contrario de lo que se pretende con esta técnica.

Para explicar con un ejemplo el desarrollo de un muestreo de cuotas se señalan, a continuación, las fases principales de un muestreo de estas características, partiendo del supuesto de que se trata de una muestra dirigida a la población en general.

En primer lugar, se ha de tener en cuenta que el diseño de la muestra y el proceso de selección, en sus primeras fases, siguen el mismo camino que cualquier otra muestra, ya que las cuotas sólo afectan a la selección de las unidades últimas de muestreo. En cuanto a las variables que se van a utilizar para las cuotas no caben muchas alternativas ya que suele existir información escasa sobre el universo objeto de estudio. Por eso casi siempre se introducen las cuotas de sexo y edad, y, raras veces, las de nivel educativo. Sobre estas tres variables hay información suficiente en los censos de población, con desagregación a nivel provincial, aunque con un desfase de

³² KISH, *Muestreo de encuestas*, op. cit., p. 650, señala, por ejemplo, que las variables sexo y edad intervienen sólo en un 1 ó 2 % en la conducta electoral.

dos o tres años desde el momento de su recogida³³. A partir de la información del censo o similar se reconstruye la estructura que debería tener la muestra en cuanto a las variables elegidas, teniendo en cuenta la distribución geográfica de las entrevistas. Esta estructura se traslada a las hojas de ruta que presentan un formato similar al que se ofrece en la hoja adjunta. Una vez la hoja en manos del entrevistador, éste tiene que seleccionar solamente a personas que cumplan las condiciones impuestas.

CUOTA DE EDAD, SEXO Y ESTUDIOS

Provincia..... Municipio.....
Sección..... Número de entrevistas 10

<i>Edades</i>	<i>Varones</i>						<i>Mujeres</i>					
18-24.....	①	1	1				1	1	1			
25-34.....		2	2	2			②	2	2			
35-44.....		③	3	3			③	3	3			
45-54.....		④	4	4			④	4	4			
55-64.....		⑤	5	5			⑤	⑤	5			
Más de 64 años.....		⑥	6	6			⑥	6	6			
<i>Estudios</i>												
Menos de primarios.....	1	2	3	4	5	⑥	1	2	3	4	⑤	⑥
Primarios completos.....	1	2	3	4	⑤	6	1	②	3	4	5	6
Bachiller elemental y equivalentes.....	1	2	3	④	5	6	1	2	③	4	5	6
Bachiller superior.....	①	2	3	4	5	6	1	2	3	④	5	6
Grado medio.....	1	2	③	4	5	6	1	2	3	4	5	6
Universidad.....	1	2	3	4	5	6	1	2	3	4	5	6

NOTA. El código de la edad identifica a los individuos en el nivel de estudios. En la parte superior se indica el sexo y la edad. Se ha señalado el círculo en la primera columna porque se ha elegido un solo individuo de esa edad y sexo. El número redondeado se traslada a la parte inferior del cuadro e indica el nivel de estudios que debe tener la persona elegida de cada sexo y edad: el número 1 de varones debe tener estudios de bachiller superior, etc.

Para que el proceso se acerque más a la aleatoriedad se suelen dar instrucciones muy concretas a los entrevistadores al objeto de evitar rutinas en la aplicación de las entrevistas. Estas instrucciones pueden ser las siguientes:

- Las entrevistas se harán sólo en las viviendas.
- Cuando en una vivienda haya más de una persona que reúna las condiciones exigidas por la cuota, se entrevista al más joven.

³³ Se está haciendo referencia a la publicación de los datos por parte del INE.

- El sistema no exige segunda visita para entrevistar a los ausentes. No obstante, si se tiene garantía de que la persona ausente volverá más tarde, debe repetirse la visita.
- Las entrevistas sólo pueden realizarse en las secciones o áreas previstas en la muestra.
- Finalmente, sería deseable que antes de iniciar la realización de las entrevistas se hiciera un croquis de la sección y, sobre él, se distribuyeran las entrevistas por toda el área.

El muestreo de cuotas es muy utilizado porque agiliza y simplifica el trabajo de campo, ya que la selección aleatoria de los elementos de la muestra —las unidades últimas de muestreo— plantea muchas dificultades a la hora de hacer las entrevistas. Es más, en muestreos probabilísticos hay que recurrir, muchas veces, a sustituciones, sin garantías de no deformar el diseño teórico inicial. Hay que añadir a lo anterior que, de hecho, la estructura de la población entrevistada se aleja, con frecuencia, de la del universo, debido a las sustituciones, que pueden introducir importantes desviaciones³⁴. Por otra parte, en sondeos de opinión, el muestreo de cuotas suele dar resultados muy aceptables, comparables con los de sondeos totalmente probabilísticos, siempre y cuando el diseño muestral sea riguroso en todas sus fases. Entre las pocas investigaciones que se han hecho sobre el particular, en España, hay dos realizadas por el CIS. En ambos casos, se hicieron muestras paralelas sobre intención de voto en el área de Cataluña, con diseños similares, salvo la elección de las unidades últimas de muestreo que en un diseño fue probabilística y en el otro por cuotas. Los resultados de estas encuestas fueron similares en las dos experiencias realizadas y, en ambos casos, ajustados a los de las elecciones que se realizaron pocos días después.

A pesar de lo anterior, y de los buenos resultados que se obtienen en muchas ocasiones, el muestreo de cuotas no es un muestreo probabilístico y por ello no se pueden hacer estimaciones rigurosas, con medición de errores muestrales e intervalos de confianza. De ahí que, como dice DEMING³⁵, el muestreo probabilístico y el de cuotas son «mercancías diferentes que no son intercambiables». Por ello cuando se trata de encuestas estatales y, en general, de encuestas que se toman como base para decisiones importantes, hay que recurrir a muestreos totalmente probabilísticos a pesar de su coste y del tiempo que se necesita para la realización de una rigurosa recogida de la información en el trabajo de campo. Es la garantía para poder medir el grado de precisión de las estimaciones.

³⁴ Es práctica frecuente tener que acudir a ponderaciones de la encuesta para darle la estructura real del universo, no conseguida en las entrevistas.

³⁵ W. E. DEMING, *Some Theory of Sampling*, Nueva York, Wiley, 1950.

Lectura 2

3

Tamaño de la muestra

Tamaño de la muestra y precisión de las estimaciones son conceptos inseparables, ya que las variaciones en uno de ellos afectan al otro y viceversa. Si aumenta el tamaño de la muestra lo hace el nivel de precisión de las estimaciones y, si se quiere conseguir una mayor precisión, es necesario modificar el tamaño de la muestra, salvo las acotaciones que se introducirán más adelante.

El tamaño de la muestra hace referencia al número de elementos del universo que se seleccionan para extraer de ellos la información que, después, se va a generalizar. La precisión de las estimaciones refleja la concentración de la distribución del estimador o, si se quiere, la magnitud de las desviaciones con respecto a la media obtenida en el muestreo. A medida que aumenta la precisión del estimador el intervalo de confianza se hace menor y, en ausencia de sesgos, las posibles diferencias entre los valores de los parámetros poblacionales y los estimados en el muestreo se hacen menores, tanto más cuanto mayor sea el nivel de precisión. Éste se mide por el error de muestreo, desviación típica del estimador, que refleja la dispersión de su distribución. En consecuencia, los errores de muestreo son concreción numérica del nivel de precisión de la estimación a que se refieren. Por eso hay que incluirlos en el estudio del tamaño de la muestra, que se calcula para variables y errores de muestreo determinados.

Antes de dejar esta incursión en el campo de la estimación y errores de muestreo, objeto de estudio del capítulo siguiente, hay que señalar que los errores de muestreo, que se fijan para el cálculo del tamaño de la muestra, deben variar según el tema de que se trate, puesto que los niveles de precisión requeridos son distintos según los objetivos de la investigación. La Encuesta de población activa, por ejemplo, dada su enorme trascendencia económica y social, necesita un nivel de precisión muy distinto al que se requiere para estimar el consumo de bebidas refrescantes.

Variables que intervienen

Volviendo al tema central de este capítulo, el estudio del tamaño de la muestra, hay que resaltar que si bien el nivel de precisión requerido —el

error de muestreo— es el elemento más importante en este campo, puesto que condiciona y, en gran parte, determina la dimensión de la muestra, hay, también, otra serie de factores relevantes que, necesariamente, se han de tener en cuenta al dimensionar una muestra. Por eso se dedican a su estudio las páginas que siguen.

Varianza poblacional

En una bolsa se pueden incluir 500 bolas blancas y 500 bolas negras o se pueden incluir, entre otras muchas posibilidades, 100 bolas blancas y 900 bolas negras. La probabilidad de extraer, aleatoriamente, una bola blanca o una bola negra es, en el primer caso, de 500/1.000 y, en el segundo, de 100/1.000 para el caso de las bolas blancas y de 900/1.000 para el caso de las bolas negras. Como la suma de todas las probabilidades es igual a 1, también se puede decir que, en el primer caso, la relación entre las probabilidades es de 0,50-0,50, y, en segundo, de 0,10-0,90. Dicho de otra forma, la primera bolsa es más heterogénea que la segunda, ya que en esta última, las bolas negras, por sí solas, representan el 90 % del total del contenido, mientras que en la primera sólo representan el 50 %.

Esta distinta distribución del universo real incide en el tamaño de la muestra, de tal forma que, en el primer caso, sería necesario extraer un número mayor de elementos de la muestra que en el segundo. Ello quiere decir que, cuando la relación entre variables es muy desigual —10 % y 90 %, por ejemplo—, la población es más homogénea y se necesita una muestra de tamaño menor que cuando la relación es del 50 %. Dicho de otra forma, para un tamaño dado de la muestra, el error será menor cuando la relación entre variables sea más desigual.

El planteamiento anterior responde a un concepto central en el campo de la estadística inferencial. Se está haciendo referencia al análisis de la varianza que, como se sabe, es una medida de dispersión. Por eso, cuando una población es más homogénea, la varianza es menor y, consecuentemente, el número de entrevistas necesarias, para construir un modelo reducido del universo, será más pequeño. Puesto que el producto de dos probabilidades — P y $(1-P)$ —, en el caso de sólo dos categorías, equivale a la varianza, en la medida en que cada una de ellas se aleje de 0,50 la varianza será menor. Por eso AZORÍN POCH afirma: «Obsérvese que al disminuir P disminuye n (el número de entrevistas), lo mismo que ocurre si P aumenta. Esto quiere decir que, a medida que se hace más uniforme la población, esto es, que la proporción tiende a ser cero o uno, será necesario un menor tamaño de la muestra para alcanzar determinada precisión»¹. Al límite, y en determinadas situaciones, bastaría con una muestra de tamaño 1 para estimar

¹ FRANCISCO AZORÍN POCH, *Curso de muestreo y aplicaciones*, Madrid, Aguilar, 1972, p. 65.

determinadas características de una población, siempre y cuando ésta fuese totalmente homogénea con respecto a la variable de análisis. Así, si se quiere estimar el color de las bolas de un bombo y se sabe que todas tienen el mismo color, es decir, que bajo este aspecto tienen homogeneidad total, varianza 0, es suficiente extraer una sola bola del bombo para estimar el color con total precisión.

Lo anterior permite resaltar que el conocimiento de la homogeneidad o heterogeneidad del universo, en el aspecto que se desea analizar, es de gran trascendencia ya que, cuando la homogeneidad aumenta, aumenta también el grado de precisión de la estimación para un número dado de entrevistas. Por tanto, en las operaciones previas al diseño muestral debería intentarse el conocimiento aproximado de la varianza, a efecto de tenerla en cuenta a la hora de decidir el tamaño de la muestra. Para ello se suele acudir a dos técnicas muy utilizadas en la investigación empírica. La primera proviene de la experiencia, el caso de estudios que se repiten periódicamente, y la segunda de prospecciones directas a través de estudios piloto. Los Eurobarómetros mensuales de las Comunidades Europeas y la Encuesta de población activa del INE obedecen a estudios periódicos cuyos universos son, en parte, conocidos y, en consecuencia, se dispone de una aproximación a las varianzas de distintas variables. Este conocimiento es suficiente a efectos del diseño de la muestra.

Los estudios piloto, previos a la realización de la investigación, son una práctica habitual en las ciencias sociales y de gran utilidad desde distintas facetas. En el caso que nos ocupa se trata de aproximarse al conocimiento de la varianza de la variable que interesa, para diseñar, después, el estudio. Baste recordar, como aplicación de esta técnica, el estudio piloto que el CIS y distintos organismos oficiales realizaron en 1985, previo a la encuesta sobre «condiciones de vida y trabajo de los españoles». En aquella ocasión se realizó un estudio piloto de estas características. (Véase cuadro 3.1.)

Como se observa, el estudio consta de cinco submuestras y, además, se

CUADRO 3.1
ENCUESTA PILOTO. NÚMERO DE ENTREVISTAS DE CADA SUBMUESTRA

<i>Áreas</i>	<i>Número de entrevistas</i>
Metropolitana de Barcelona	800
Metropolitana de Madrid.....	800
Provincia de Cádiz.....	500
Provincia de Ávila.....	300
Provincia de Alicante	500
Total	2.900

FUENTE: CIS.

extiende a áreas plenamente diferenciadas. Se realizan entrevistas en áreas metropolitanas, industriales, turísticas y agrícolas al objeto de tomar muestras de situaciones muy dispares que sirvan de referencia para el diseño del estudio definitivo. En concreto, el estudio piloto sirvió para probar el borrador del cuestionario y examinar las diferencias fundamentales entre zonas rurales y urbanas y entre áreas industriales, turísticas y agrícolas. Los datos facilitados sirvieron para calcular la varianza en relación a la actividad y, consecuentemente, para dimensionar la muestra y preparar una afijación no proporcional por provincias.

El estudio piloto reseñado no responde a los estudios que se suelen hacer ordinariamente. Se hizo un estudio ampliamente dimensionado y mediante un proceso de muestreo real. Se pretendía, con ello, tomar el máximo de garantías técnicas para la posterior realización de un estudio de 60.000 entrevistas con representatividad provincial. En la práctica, los estudios piloto suelen tener un número pequeño de entrevistas —entre 100 y 200— distribuidas entre la población que va a ser objeto de investigación. Ese número reducido es suficiente para orientar al investigador y facilitar el diseño muestral.

Hasta aquí se ha hablado del cálculo de la varianza con respecto a una variable, pero en la realidad son muchas las cuestiones que interesan y que deberían tenerse en cuenta a la hora de calcular la muestra. De ahí que, como dice PULIDO SAN ROMÁN, para el cálculo del tamaño de la muestra se deberían tener en cuenta las preguntas más relevantes o la fracción más importante del cuestionario: «Quiero indicar, por último, en relación con el tema de determinación del tamaño de la muestra, que, generalmente, una investigación mediante encuestas no se referirá a una única cuestión, sino que, como posteriormente veremos, recogerá datos sobre los múltiples aspectos que puede incluir un cuestionario. De esta forma, los resultados de cada pregunta exigirían un tamaño de muestra diferente para una misma fiabilidad. Ante la complejidad del problema, en la práctica se opta, generalmente, por determinar un tamaño de muestra tal que garantice una fiabilidad dada, bien en las preguntas clave o bien en una fracción importante del cuestionario. "A posteriori" deberá indicarse el margen de error que cabe esperar para los diferentes tipos de cuestiones base de la encuesta»².

Esto se traduce en que, una vez hecho el estudio de la varianza de las variables que más interesan, debería tomarse como punto de referencia la más heterogénea, la de varianza más alta y, por tanto, la que exige mayor tamaño de la muestra. En la práctica del muestreo, cuando no se realizan estudios previos de varianza o cuando se pretende estimar parámetros referidos a múltiples cuestiones, que se supone tienen diferencias notorias en su homogeneidad/heterogeneidad —varianza—, se toma la proporción

² A. PULIDO SAN ROMÁN, *Estadística y técnicas de investigación social*, Madrid, Pirámide, 1976, p. 184.

$P = (1 - P)$, es decir, las probabilidades 0,50 y 0,50. Esto exige un mayor tamaño de la muestra, para un nivel de precisión dado, encarece los costes, pero da garantías de no equivocarse por defecto al determinar el tamaño de una muestra, que se va a utilizar para estimar parámetros de variables muy dispares.

Tipo de muestreo

El tipo de muestreo utilizado tiene también su incidencia en el tamaño de la muestra. En general, tomando como referencia el muestreo aleatorio simple, suele suceder que el muestreo por conglomerados es menos preciso y el estratificado más³. En este último, que es el que permite obtener mayores ganancias en precisión, se trata de acotar, con la estratificación, subpoblaciones homogéneas en su composición y heterogéneas entre sí, con lo que se habrá conseguido disminuir la varianza total y, consiguientemente, serán necesarias menos entrevistas para el mismo grado de precisión. Esto hace que el estudio de la estratificación, aunque sólo fuera por las ganancias que aporta en precisión, sea relevante y haya que tenerlo en cuenta en los diseños muestrales. La influencia real de la estratificación sobre el nivel de precisión se cuantifica, en el capítulo 4, con un ejemplo real.

Nivel de confianza

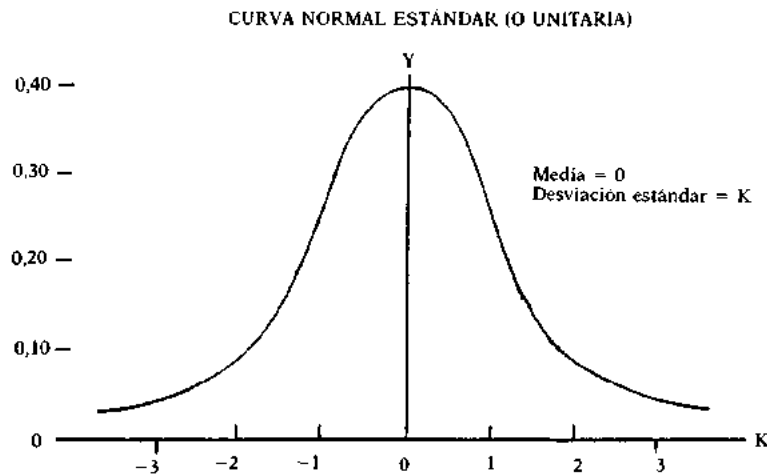
Cuando se habla en estadística de nivel de confianza se hace referencia a la probabilidad de acertar, es decir, a la probabilidad de que una estimación, en ausencia de sesgos, se ajuste a la realidad. En consecuencia, el nivel de confianza hay que interpretarlo como una acotación de la función de densidad de probabilidades, o simplemente distribución de probabilidad. De ahí que el estudio de esta función sea de especial importancia en el campo de la estadística inferencial.

Entrando en este análisis, hay que anotar, en primer lugar, que «para muchas variables aleatorias, la distribución de probabilidad es una curva específica y bien delineada que recibe el nombre de curva normal o de Gauss...; es la distribución de probabilidad más útil en estadística»⁴. «Además, muchas distribuciones de las cuales deseamos encontrar las áreas bajo sus curvas de probabilidad correspondientes son aproximadamente normales. Por consiguiente, muchos problemas de estadística matemática se re-

³ No se insiste en este tema porque se ha hablado de él en el capítulo 2 y se volverá a hablar en el capítulo 4, incluyendo los respectivos cálculos.

⁴ Thomas H. WONNACOTT y Ronald J. WONNACOTT, *Introducción a la estadística*, México D.F., Limusa, 1979, p. 92.

suelven o se pueden resolver mediante una curva de distribución normal»⁵. Es más, cuando se trata de muestras relativamente grandes se puede asumir que siguen una distribución aproximadamente normal, como señalan todos los autores. En estos casos, cuando se da la normalidad, las distribuciones quedan perfectamente caracterizadas a partir de la media y la varianza y, por lo tanto, se puede calcular el área que queda bajo la curva, dadas las siguientes propiedades: «(1). La curva normal estándar se centra (tiene la media) en 0 y es simétrica respecto al 0. Esto significa que es exactamente igual a la derecha o a la izquierda de 0. (2). La curva normal estándar se extiende en forma indefinida hacia la derecha y hacia la izquierda, aproximándose cada vez más al eje horizontal pero sin tocarlo. Las curvas de distribución normal son curvas matemáticas. Por ello se supone que son posibles todos los valores a lo largo del eje horizontal. (3). El área total bajo la curva normal estándar es de 1,00. Por simetría, esto significa que existe un área de 0,5 a la derecha de la media 0, o un área de 0,5 a la izquierda»⁶.



Con la ayuda del cuadro 3.2 se puede obtener la probabilidad correspondiente a cada una de las múltiples áreas que se pueden acotar en la curva normal con sólo fijar la desviación estándar⁷. Ésta delimita el nivel

⁵ S. L. WEINBERG y K. P. GOLDBERG, *Estadística básica para las ciencias sociales*, México D.F., Interamericana, 1982, p. 167.

⁶ S. L. WEINBERG y K. P. GOLDBERG, *ibid.*, p. 169.

⁷ La tabla ofrece resultados para las desviaciones a la derecha o a la izquierda de la media. Cuando el área acotada por K desviaciones viene precedida del signo \pm , significa que se acotan dos áreas simétricas, una a la izquierda y otra a la derecha de la media 0. En consecuencia, la probabilidad que aparece en la tabla, correspondiente al valor de K , hay que multiplicarla por 2.

CUADRO 3.2

ÁREAS BAJO LA CURVA NORMAL ESTÁNDAR (ÁREA ENTRE LA MEDIA 0 Y EL PUNTAJE K)
(Segundo lugar decimal en k)

K	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998									
4,0	0,49997									
4,5	0,499997									
5,0	0,4999997									

de confianza con que se desea trabajar y, por tanto, la probabilidad de ajustar a la realidad las estimaciones derivadas del diseño muestral.

Así, para la desviación estándar igual a ± 1 , la probabilidad es de $0,3413 \times 2 = 0,6826$; para ± 2 desviaciones estándar, es de $0,4772 \times 2 = 0,9544$, etcétera. De acuerdo con la propiedad 2, *se pueden calcular las probabilidades no sólo para 1, 2 ó 3 desviaciones estándar, sino para cualquier valor, dentro de los límites de las abscisas*. Para el valor $\pm 1,5$ desviaciones, por ejemplo, la probabilidad sería de $0,4332 \times 2 = 0,8664$, etcétera.

Todo ello quiere decir que, a la hora de diseñar la muestra, se debe fijar el nivel de confianza —la probabilidad de acertar— de acuerdo con los objetivos de la investigación. Si no se fija, la probabilidad será del 68 %, valor de la desviación igual a ± 1 y, si se fija, la probabilidad será la que se corresponda con \pm el valor de la desviación (K) que se acote en el eje de abscisas. (Véase cuadro 3.2.)

En resumen, los distintos niveles de confianza corresponden a distintas probabilidades y la elección de un nivel mayor ofrece mayores garantías de acertar en la estimación. Como contrapartida, el tamaño de la muestra deberá ser mayor. En general, en las ciencias sociales, se suele trabajar con un nivel de confianza del 0,9544, que es el que corresponde a la probabilidad de la curva acotada por ± 2 desviaciones estándar. Ello significa que de 100 muestras que se diseñaran adecuadamente, la probabilidad de acertar en las estimaciones se daría en el 95,44 % de los casos, mientras que el riesgo estadístico de equivocarse se reduciría al 4,56 %.

Afijación de la muestra

La distribución de las entrevistas entre los diferentes subconjuntos en que se puede dividir el universo objeto del estudio se llama, en la terminología estadística, afijación de la muestra. Dicha distribución puede realizarse de muy distintas formas, en función de los fines que se persiguen en el estudio y del tipo de muestreo que se utilice.

Criterios básicos

Los criterios básicos para hacer la afijación son tres: afijación simple, afijación proporcional y afijación óptima. La afijación simple consiste en asignar a cada estrato un número igual de entrevistas; en la asignación proporcional la distribución se hace de acuerdo con el peso relativo de la población de cada estrato y, en la afijación óptima, se tiene en cuenta la

homogeneidad o heterogeneidad⁸ de la población bajo determinados aspectos.

La aplicación de las distintas afijaciones a una muestra para la provincia de Ávila, y a otra dirigida a funcionarios ofrece los resultados de los cuadros 3.3 y 3.4.

CUADRO 3.3

APLICACIÓN DE DIFERENTES TIPOS DE AFIJACIÓN A UNA MUESTRA DE LA PROVINCIA DE ÁVILA

Estratos	Porcentaje población	Desviación típica	Afijación		
			1	2	3
1. Menos de 2.000 h.....	50,2	10	333	502	209
2. De 2.001 a 10.000 h.....	27,9	30	333	279	344
3. De 10.001 a 50.000 h.....	21,9	50	333	219	450
	100,0		999	1.000	1.000

NOTAS: Afijación: 1) simple; 2) proporcional; 3) óptima.

La Desviación Típica no se conoce, se supone para desarrollar el ejemplo. Hace referencia, en este caso, a niveles de ingresos.

FUENTE: Tomado de Jacinto RODRÍGUEZ OSUNA: «La muestra, teoría y aplicación» en GARCÍA FERRANDO y otros, *El análisis de la realidad social*, Madrid, Alianza Editorial, 1986, p. 277.

CUADRO 3.4

APLICACIÓN DE DIFERENTES TIPOS DE AFIJACIÓN A UNA MUESTRA DE FUNCIONARIOS

Estratos	Porcentaje funcionarios	Desviación típica	Afijación		
			1	2	3
1. Grupo A.....	10	100	250	100	260
2. Grupo B.....	15	60	250	150	240
3. Grupo C.....	20	40	250	200	210
4. Grupo D.....	55	20	250	550	290
Total.....	100		1.000	1.000	1.000

NOTAS: Afijación: 1) simple; 2) proporcional; 3) óptima.

Los grupos se corresponden con la titulación exigida para formar parte de los Cuerpos del Estado. Grupo A, titulados superiores; grupo B, titulados de grado medio; grupo C, bachillerato superior; grupo D, EGB completa.

La Desviación Típica no se conoce, se supone para desarrollar el ejemplo. Hace referencia, en este caso, a niveles de ingresos.

FUENTE: Tomado de Jacinto RODRÍGUEZ OSUNA: «La muestra, teoría y aplicación» en GARCÍA FERRANDO y otros, *El análisis de la realidad social*, Madrid, Alianza Editorial, 1986, p. 277.

⁸ La homogeneidad o heterogeneidad se mide por la varianza o su raíz cuadrada, la desviación típica.

En ambos ejemplos se ha partido de una muestra de 1.000 elementos para facilitar los cálculos. En la realidad habría que haber empezado por discutir el tamaño de la muestra, pero esto se hace en otra parte de este trabajo.

La primera afijación es la simple, en la que a todos los estratos se les ha asignado el mismo número de entrevistas. Este tipo de afijación tiene la finalidad de que la muestra, con un determinado error, sea representativa de cada uno de los estratos y, por lo tanto, se puedan sacar conclusiones a este nivel. En la práctica, y para conseguir este objetivo, se suele aplicar la distribución proporcional, cargando la muestra en los estratos en los que interesa obtener estimaciones más próximas a la realidad. En el ejemplo de los funcionarios, con 100 entrevistas en la afijación proporcional, difícilmente se pueden hacer estimaciones para el grupo A. En este caso se puede cargar la muestra con 200 entrevistas más, lo que supondría 300 entrevistas para este estrato y 1.200 en total.

La afijación proporcional no necesita ningún tipo de aclaración —véanse los ejemplos—, pero la óptima sí. La filosofía que preside el muestreo de la afijación óptima es la siguiente: cuando los estratos son muy homogéneos⁹, desviación típica muy pequeña, con una muestra reducida se pueden obtener buenas estimaciones, y cuando el estrato es muy heterogéneo, desviación típica elevada, es necesario sacar una muestra mucho más grande para conseguir estimaciones que se aproximen a la realidad. En el ejemplo de los funcionarios, la desviación típica hace referencia a los sueldos. Quiere decir que la banda de fluctuación de los mismos se va haciendo mayor a medida que se habla de grupos de titulación más alta. Los cuerpos que exigen titulación superior son múltiples y los salarios, hasta época reciente, variaban considerablemente de unos a otros. La forma de extraer la muestra bajo el supuesto de la afijación óptima se presenta el cuadro 3.5.

Conocida la desviación típica del estrato, se multiplica por el porcentaje que éste representa sobre el universo. Se repite la operación con los diferentes estratos hasta crear un nuevo universo (columna A × B del cuadro), expresado en unidades de desviación típica. Posteriormente se desarrolla la muestra como en el caso de la afijación proporcional.

Nivel de desagregación

Aparte de los sistemas de afijación, señalados anteriormente, se da, con frecuencia, la necesidad de diseñar muestras para obtener estimaciones no sólo de la población tomada como un todo, sino también de diferentes sub-

⁹ Cuando se habla de estratos homogéneos o heterogéneos se está haciendo referencia a la desviación típica. Ésta hay que calcularla en función de variables concretas. En nuestros ejemplos se está haciendo referencia a los niveles de ingresos de la población del estrato.

CUADRO 3.5
APLICACIÓN DE LA AFIJACIÓN ÓPTIMA A LA MUESTRA DE FUNCIONARIOS

<i>Estrato/Grupos</i>	<i>A. Porcentaje funcionarios</i>	<i>B. Desviación Típica</i>	<i>A × B</i>	<i>C. Porcen- taje</i>	<i>Distribución de acuerdo a C</i>
1. Grupo A.....	10	100	1.000	26	260
2. Grupo B.....	15	60	900	24	240
3. Grupo C.....	20	40	800	21	210
4. Grupo D.....	55	20	1.100	29	290
	100		3.800	100	1.000

FUENTE: Tomado de Jacinto RODRÍGUEZ OSUNA: «La muestra, teoría y aplicación» en GARCÍA FER-
 RRANDO y otros, *El análisis de la realidad social*, Madrid, Alianza Editorial, 1986, p. 279.

divisiones en las que se puede acotar el universo objeto de estudio. En el primer caso se trataría de diseñar una muestra de tamaño suficiente para poder hacer estimaciones globales, mientras que, en el segundo, lo que se desea es hacer estimaciones, además, de parámetros pertenecientes a subpoblaciones o dominios de estudio ¹⁰.

Para centrar el análisis se va a recurrir a los datos del cuadro 3.6, extraídos de una muestra real, en el que se estima el hábito de fumar de los ciudadanos a nivel nacional, por grupos de edades, tamaño de municipios, nivel de educación, etcétera. Centrando el análisis en los datos globales y en los obtenidos cuando se introduce la variable educación, se observa que, en el primer caso, la estimación se hace en base a 2.498 entrevistas (véase el cuadro) y en el segundo en base a un número muy inferior: 762 con menos de estudios primarios; 899 con estudios primarios; 578 con bachiller y 242 con estudios superiores.

El nivel de precisión es diferente en cada uno de estos casos y los errores de muestreo aumentan a medida que disminuye el número de entrevistas. En concreto, cuando el 47 % de la población total dice que nunca ha fumado, el error de muestreo, para un nivel de confianza del 95,44 % es de 0,02, mientras que cuando el 27 % de la subpoblación con estudios superiores afirma lo mismo, el error de la estimación es de 0,057 ¹¹ es decir, 2,85 veces mayor. El error de muestreo en este último caso representa el 21 % del valor de la estimación —0,057/0,27— lo que pone en evidencia la escasa precisión y aproximación al parámetro poblacional. En consecuencia, con

¹⁰ Nótese que el planteamiento que aquí se hace es distinto al que se hizo en el epígrafe anterior al hablar de la afijación simple, en la que se asignaba el mismo número de entrevistas a cada estrato.

¹¹ Las fórmulas para calcular los errores se explican más adelante.

CUADRO 3.6
EL HÁBITO DE FUMAR

PREGUNTA: *Para empezar, ¿podría usted decirme si fuma o ha fumado alguna vez en su vida de forma habitual?*

	<i>N.º de entrevistas</i>	<i>Fuma</i>	<i>Ahora no fuma pero ha fumado</i>	<i>Nunca ha fumado</i>	<i>No sabe No contesta</i>
NACIONAL	2.498	37	16	47	1
SEXO					
Varones	1.185	55	23	21	1
Mujeres	1.313	21	9	70	1
EDAD					
De 18 a 25 años.....	489	57	9	34	0
De 26 a 40 años.....	656	50	14	35	1
De 41 a 50 años.....	404	32	16	51	—
De 51 a 60 años.....	398	26	16	58	—
Más de 60 años.....	541	15	22	61	2
No contesta	10	30	60	10	—
TAMAÑO DEL MUNICIPIO					
Menos de 10.000 hab.....	641	37	14	49	1
De 10.000 a 100.000 hab....	525	36	16	48	—
De 100.001 a 1.000.000 hab.	718	37	16	46	1
Más de 1.000.000 hab.....	614	39	17	44	1
EDUCACIÓN					
Menos de Primarios.....	762	22	16	61	1
Primarios	899	36	15	48	1
Bachiller	578	52	14	33	1
Superiores	242	55	18	27	0
No contesta	17	29	18	53	—
SITUACIÓN LABORAL					
Trabaja	1.021	54	18	28	0
Parado.....	182	64	13	23	1
Jubilado.....	310	25	36	38	1
Estudiante	152	47	7	45	1
Sus labores	804	12	7	80	1
No contesta	29	41	21	38	—
OCUPACIÓN ACTIVOS					
Emp. Directivos.....	73	55	21	23	1
Peq. Emp. Artes. Vend.....	269	48	20	32	—
Emp. Cuadros Medios.....	211	48	22	30	—
Labradores	187	36	26	37	1
Obreros	734	52	20	27	1
No procede	1.024	20	8	71	1

NOTA: Se trata de una muestra nacional de 2.500 entrevistas con afijación proporcional.

FUENTE: *Revista Española de Investigaciones Sociológicas*, núm. 44, octubre-diciembre de 1988, p. 193.

diseños muestrales de este tipo no se puede lograr la suficiente precisión para estimar parámetros de las subpoblaciones.

Para resolver estas situaciones se suele recurrir a dos técnicas bien diferentes. La primera consiste en aumentar el tamaño de la muestra total hasta que el número de entrevistas en la subpoblación menor sea el adecuado. En el caso que nos ocupa, si para realizar las estimaciones correspondientes a la categoría «estudios superiores» se necesitan 600 entrevistas, habría que elevar el tamaño de la muestra, de forma proporcional, hasta las 6.151 entrevistas. Es decir, la muestra inicial de 2.498 elementos se ha multiplicado por 2,462 y, por tanto, también cada una de las subpoblaciones. En la nueva muestra aparecerán 1.889 elementos con estudios inferiores a primarios, 2.229 con estudios primarios, 1.433 con bachiller y 600 con estudios superiores.

La solución tomada garantiza los niveles de precisión que se requieren, pero a costa de una considerable elevación del número total de entrevistas y de los costes consiguientes. Además, sería necesario conocer cuáles son las proporciones de las diferentes subpoblaciones a efectos de, una vez fijadas las entrevistas para la categoría menos frecuente, poder fijar, proporcionalmente, las que corresponden a los restantes.

La segunda técnica, y la más habitual, consiste en hacer una afijación no proporcional, tratando a cada subdivisión o dominio como si fuera un universo diferente. Una investigación nacional que necesitara estimaciones a nivel de alguna Comunidad Autónoma, una investigación regional que necesitara estimaciones a nivel de alguna provincia, o una investigación en la universidad, que necesitara estimaciones por facultades, son algunos ejemplos que habría que tratar mediante afijación no proporcional. Como la precisión del estimador depende, en general, del tamaño de la muestra y no de la fracción del muestreo esto hace que, «en ciertos casos pueda necesitarse una muestra prácticamente del mismo tamaño para obtener datos de una provincia, y datos para la totalidad de la nación»¹². Ello quiere decir que, en estos casos, para cada una de las subpoblaciones que se necesitan estudiar habría que realizar las respectivas muestras por lo que, en rigor, habría que hablar de un diseño con varias submuestras. Con esta técnica, al contrario de lo que ocurría con la anterior, sólo se aumenta la muestra en los dominios que interesa estudiar y, por tanto, se produce un ahorro de entrevistas y de costes sin que afecte los objetivos de la investigación.

Puesto que el tema de los dominios de análisis es muy importante, se quiere volver sobre el mismo, trayendo como ejemplo la muestra de la Encuesta de población activa. Se trata de una muestra nacional de 60.000 hogares pensada no sólo para ofrecer resultados a nivel de todo el país sino

¹² J. L. SÁNCHEZ CRESPO, *Muestreo de poblaciones finitas aplicado al diseño de encuestas*, Madrid, INE, 1973, p. 12.

también de cada una de las provincias. Si la distribución de la muestra fuera proporcional, con igual fracción de muestreo en todas las provincias y, a la vez, fuera representativa de cada una de ellas, habría que aumentar considerablemente el número de entrevistas¹³, con costes prohibitivos y bastantes dificultades operativas, en la recogida de la información. Por eso, la Encuesta de población activa opta por otra solución, más próxima a la que, anteriormente, se ha llamado «segunda técnica». En la EPA se utiliza una afijación de compromiso entre la uniforme y la proporcional. Consiste en que cada una de las provincias es tratada como un universo diferente, a efectos de muestreo, aunque el número de entrevistas sólo varía cuando el tamaño de la provincia rebasa un determinado número de habitantes. Para aplicar el sistema se establecen muestras de diferentes tamaños, que se asignan a las provincias según el número de habitantes de las mismas. En concreto, los módulos de la EPA son los que aparecen en el cuadro 3.7.

CUADRO 3.7
MÓDULOS —TAMAÑOS— DE LA MUESTRA EPA QUE SE APLICAN A LAS DIFERENTES PROVINCIAS

<i>Módulos</i>	<i>N.º secciones en la muestra</i>	<i>N.º viviendas en la muestra</i>	<i>N.º Provincias en cada módulo</i>
1.º	36	720	19
2.º	48	960	1
3.º	72	1.440	25
4.º	108	2.200	3
5.º	144	2.880	2
			50

FUENTE: INE.

Los módulos más utilizados son el 1.º y el 3.º, lo que quiere decir que en un gran número de provincias, las pequeñas, se entrevistan a 720 hogares y, en las provincias de tamaño intermedio, a 1.440, el doble. El resto de los módulos tienen muy poca aplicación y se reservan para Madrid, Barcelona, Valencia, etcétera.

Lo dicho hasta aquí resuelve la mayor parte de las situaciones que se suelen plantear al diseñar una muestra. Sin embargo, hay otros casos de mucha más difícil solución. Se está haciendo referencia a la estimación de parámetros raros, cuya existencia es muy poco frecuente. Así, es prácticamente imposible estimar, por muestreo, el porcentaje de población accidentada en la carretera, la que pasa las vacaciones en un lugar poco fre-

¹³ El número de hogares a entrevistar sería superior a 150.000.

cuentado, la que ha contraído matrimonio 4 veces, etcétera. En estos casos la estimación por muestreo se hace imposible ¹⁴ y, si se necesitara estimar no el montante de población que pertenece a estos dominios sino algunas características de la misma, habría que crear ficheros ¹⁵ de estas subpoblaciones para, posteriormente, muestrearlos.

Ponderación de la muestra

Las técnicas de afijación no proporcional hacen que las fracciones de muestreo sean distintas y, por lo tanto, si se quieren tabular conjuntamente las distintas submuestras ¹⁶, hay que proceder a su ponderación, si no se quieren deformar las estimaciones. Se trata de devolver a cada subpoblación, estrato o dominio la proporcionalidad que tiene en la realidad al objeto de poder agregarla según interese para el estudio. Dados los diferentes aspectos que presenta el tema, se va a explicar con la ayuda de los datos del cuadro 3.8.

Este cuadro recoge los datos básicos de un estudio electoral realizado en una comunidad autónoma con dos provincias, de las cuales, en el ejemplo, una tiene 500.000 electores, población de 18 años y más, y la otra, un millón. Dado que se necesitan estimaciones a nivel provincial, se ha diseñado una muestra de 1.000 entrevistas, dividida en dos submuestras de 500 entrevistas, correspondiendo cada submuestra a una provincia.

De los datos del cuadro se deduce que, en la provincia A, piensan votar al partido C el 60 %, es decir 300 de los 500 entrevistados, — $300/500 = 60\%$ — y, en la provincia B, el 20 %, es decir 100 de los 500 entrevistados, — $100/500 = 20\%$. Sin embargo, si se quiere saber el porcentaje de votantes del partido C en el conjunto de la comunidad no se puede proceder por simple agregación porque las fracciones de muestreo son distintas en una y otra provincia. En consecuencia, la operación de agregar los datos de las dos submuestras, 400 votantes del partido C y dividirlos por los 1.000 elementos de las submuestras, igual a 40 %, no es posible y el resultado es erróneo. La única posibilidad de agregar los datos en estos casos es introducir una operación intermedia previa a la tabulación, que iguale las tasas de muestreo. Esto se consigue bien con la ayuda de coeficientes de ponderación bien con el recurso a elevadores.

Los coeficientes de ponderación, que se calculan en el apartado 4 del cuadro, se obtienen dividiendo el porcentaje que representa la población de

¹⁴ El tamaño de la muestra sería elevadísimo. Véase el último epígrafe de este capítulo.

¹⁵ La policía tiene ficheros de los accidentes; el juzgado, de los matrimonios; las agencias de viajes o las compañías aéreas, de los viajes, etcétera. Cada vez se cuenta con más medios estadísticos para reconstruir ficheros de los universos poco frecuentes.

¹⁶ Si cada una de las submuestras se tabula por separado no es necesario proceder a la ponderación.

CUADRO 3.8
CÁLCULO DEL COEFICIENTE DE PONDERACIÓN Y DE ELEVADORES

	Provincias		Total
	A	B	
1. Población de 18 años y más			
1.1. Número.....	500.000	1.000.000	1.500.000
1.2. Porcentaje.....	33,3	66,6	99,9
2. Muestra			
2.1. Elementos.....	500	500	1.000
2.2. Porcentaje sobre el total.....	50	50	100
3. Votantes del partido C en la muestra			
3.1. Número.....	300	100	400
3.2. Porcentaje sobre la muestra.....	60	20	40
4. Ponderación			
1.2 ÷ 2.2.....	0,666	1,332	
5. Elevadores			
1.1 ÷ 2.1.....	1.000	2.000	

FUENTE: Elaboración propia.

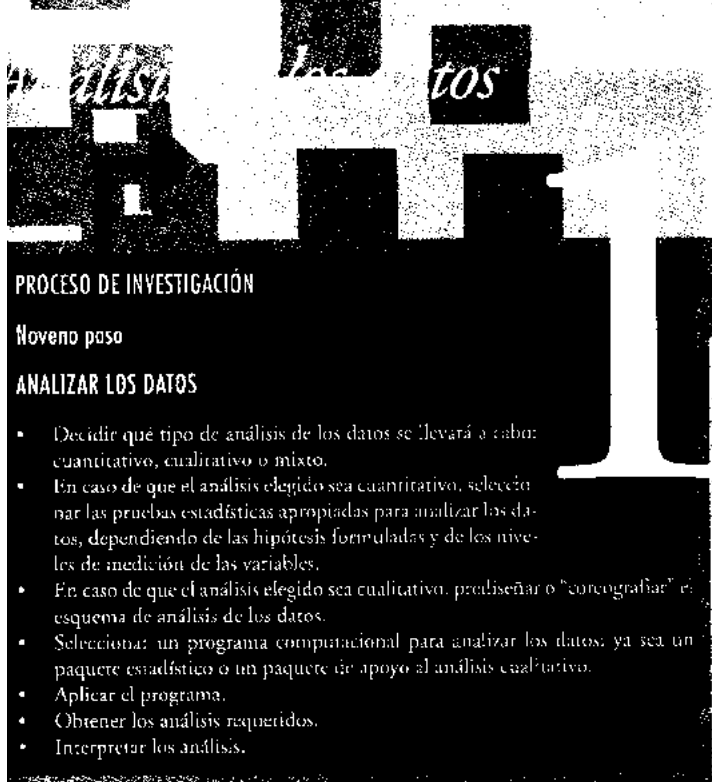
cada submuestra sobre el universo total, por el porcentaje que representa en la muestra. En el ejemplo, y para la provincia A, las operaciones son estas: 33,3 dividido por 50 igual a 0,666 y, para la provincia B, 66,6 dividido por 50 igual a 1,331. Si los 300 votantes del partido C en la submuestra de la provincia A se ponderan por 0,666 y al resultado se agregan los 100 votantes al mismo partido en la submuestra de la provincia B, ponderados por 1,332, se obtienen 333 votantes que, divididos entre los 1.000 elementos de la muestra, arrojan el porcentaje real del 33,3 %. A este mismo resultado se puede llegar por otros caminos. El 60 % de votantes del partido C en la provincia A equivale a 300.000 votantes de los 500.000 que hay en la provincia y el 20 % de votantes al mismo partido en la provincia B equivale a 200.000. Por tanto, hay 500.000 electores que piensan votar al partido C, de entre el millón y medio que hay en la Comunidad, es decir, el 33,3 %.

Otro sistema de homogeneizar las submuestras, para poder agregarlas, es recurrir al sistema de elevadores —la inversa de la fracción de mues-

treo— (véase núm. 5 del cuadro 3.8), con lo que se consigue, además, trasladar los resultados de la muestra al universo total. En nuestro caso, la muestra de la provincia A se multiplica por 1.000 y la de la provincia B por 2.000 con lo que, en el nuevo fichero creado con los elevadores, todos los datos aparecerán referidos a datos reales del universo.

A estas notas sobre la ponderación de la muestra hay que hacer dos salvedades del máximo interés. Para utilizar elevadores hay que conocer el volumen real del universo, pues de lo contrario se pueden falsear los datos por deficiencias de información. Por otra parte, tanto aquí como al utilizar los coeficientes de ponderación, hay que tener en cuenta que el soporte de la información y, por lo tanto, de la inferencia que se pueda hacer, lo constituye el número de entrevistas realmente aplicadas. En consecuencia, los errores de muestreo y, en definitiva, la precisión de las estimaciones hay que calcularlos sobre la muestra real, independientemente del número de casos que puedan aparecer en la tabulación como consecuencia de la aplicación de coeficientes de ponderación o de elevadores.

Lectura 3



Capítulo 10

PROCESO DE INVESTIGACIÓN

Noveno paso

ANALIZAR LOS DATOS

- Decidir qué tipo de análisis de los datos se llevará a cabo: cuantitativo, cualitativo o mixto.
- En caso de que el análisis elegido sea cuantitativo, seleccionar las pruebas estadísticas apropiadas para analizar los datos, dependiendo de las hipótesis formuladas y de los niveles de medición de las variables.
- En caso de que el análisis elegido sea cualitativo, predefinir o "concepcionar" el esquema de análisis de los datos.
- Seleccionar un programa computacional para analizar los datos: ya sea un paquete estadístico o un paquete de apoyo al análisis cualitativo.
- Aplicar el programa.
- Obtener los análisis requeridos.
- Interpretar los análisis.

SÍNTESIS

El capítulo presenta los esquemas fundamentales para efectuar análisis estadístico y análisis cualitativo. Asimismo, se comentan, analizan y ejemplifican las pruebas y los análisis estadísticos más utilizados en ciencias sociales, incluyendo estadísticas descriptivas, análisis paramétricos, no paramétricos y multivariados. En la mayoría de estos análisis el enfoque del capítulo se centra en los usos y la interpretación de la prueba, más que en el procedimiento de calcular estadísticas. Debido a que en la actualidad los análisis se realizan con ayuda de computadora y no manualmente, las fórmulas ayudan a entender los conceptos estadísticos, pero no a calcular estadísticas hoy en día.

El capítulo también proporciona una introducción general a los análisis multivariados.

OBJETIVOS DE APRENDIZAJE

Al terminar este capítulo, el alumno será capaz de:

- Conocer los principales métodos para realizar análisis cuantitativo y cualitativo de los datos.
- Visualizar cómo el análisis cuantitativo y el análisis cualitativo de los datos son complementarios.
- Diferenciar la estadística descriptiva y la inferencial, la paramétrica y la no paramétrica.
- Conocer la secuencia que se sigue regularmente para analizar en forma cuantitativa los datos.
- Comprender las principales pruebas estadísticas desarrolladas, así como sus aplicaciones, situaciones en las que se utiliza cada una y la forma de interpretarlas.

Tomada de: Roberto Hernández *et al.*, "Análisis de datos", en *Metodología de la investigación*, México, McGraw-Hill, 2003, pp. 493-622.

OBJETIVOS DE APRENDIZAJE

(continuación)

- Analizar la interrelación entre distintas pruebas estadísticas.
- Conocer una propuesta para efectuar análisis cualitativo de los datos.
- Identificar técnicas, métodos y procedimientos para analizar cualitativamente los datos.

Por otra parte, se propone un proceso para realizar análisis cualitativo, que va desde la revisión del material y el desarrollo de un plan o una coreografía de análisis hasta la interpretación de los datos en su contexto y la valoración del estudio. Se explica la codificación en profundidad y se comentan brevemente los principales programas computacionales de análisis cualitativo, así como algunas técnicas para establecer relaciones de causalidad.

¿ANÁLISIS CUANTITATIVO O ANÁLISIS CUALITATIVO?

La elección del tipo de análisis que habremos de realizar depende de los datos que hayamos recolectado. Si obtuvimos datos cuantitativos, éstos ya fueron codificados y preparados para el análisis en la matriz de los datos; ahora lo que sigue es efectuar el análisis cuantitativo. Si los datos recolectados fueron cualitativos, el análisis, en consecuencia, es cualitativo, por lo que puede tener algún análisis cuantitativo (en particular, conteo de frecuencias de aparición de categorías). En el caso de que hayamos obtenido datos cuantitativos y cualitativos, a cada tipo de datos le aplicamos el análisis correspondiente.

ANÁLISIS CUANTITATIVO DE LOS DATOS (ANÁLISIS ESTADÍSTICO)

¿QUÉ PROCEDIMIENTO SE SIGUE PARA ANALIZAR CUANTITATIVAMENTE LOS DATOS?

Una vez que los datos se han codificado, transferido a una matriz y guardado en un archivo, el investigador procede a analizarlos.

En la actualidad, el *análisis cuantitativo de los datos* se lleva a cabo *por computadora*. Ya nadie lo hace de forma manual, en especial si hay un volumen considerable de datos. Por otra parte, casi en todas las instituciones de educación superior, centros de investigación, empresas y sindicatos se dispone de sistemas de cómputo para archivar y analizar datos. De esta suposición parte el presente

capítulo. Por ello, el *énfasis* se centra en la *interpretación de los métodos de análisis cuantitativo* y no en los procedimientos de cálculo.¹

El análisis de los datos se efectúa sobre la *matriz de datos* utilizando un *programa computacional*. El procedimiento de análisis se esquematiza en la figura 10.1.

Veamos paso a paso el procedimiento mencionado.

¿QUÉ ANÁLISIS PUEDEN EFECTUARSE EN LOS DATOS?

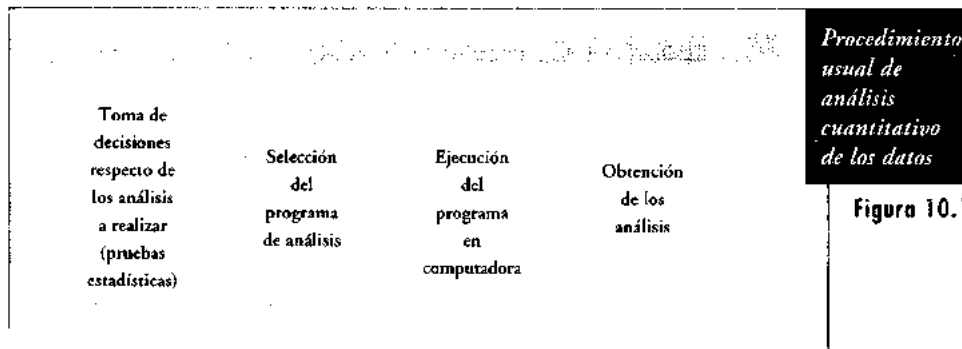
Los análisis dependen de tres factores:

- El *nivel de medición* de las variables.
- La manera como se hayan formulado las *hipótesis*.
- El *interés del investigador*.

Estadística: procedimientos para clasificar, calcular, analizar y resumir información numérica que se obtiene de manera sistemática.

Por ejemplo, no son los mismos los análisis cuantitativos que se aplican a una variable nominal que a una por intervalos. Se sugiere recordar los niveles de medición vistos en el capítulo anterior.

El investigador busca, en primer término, describir sus datos y posteriormente efectuar análisis estadísticos para relacionar sus variables. Es decir, realiza análisis de *estadística descriptiva* para cada una de sus variables y luego describe la



¹ Para aquellos lectores que deseen conocer los procedimientos de cálculo de los métodos de análisis cuantitativo se recomienda Wright (1979), Nie *et al.* (1975), Levin (1979), Kerlinger y Pedhazur (1997), Ferrán (2001), Camacho (2001) y los diferentes volúmenes de la serie "Quantitative Applications in the Social Sciences" editados por Sage Publications, Inc. Además, cualquier libro de estadística social contiene dichos procedimientos de cálculo.

relación entre éstas. Los tipos o métodos de análisis cuantitativo o estadístico son variados y se comentarán a continuación; pero cabe señalar que el análisis no es indiscriminado, cada método tiene su razón de ser y un propósito específico; no deben hacerse más análisis de los necesarios. La estadística no es un fin en sí misma, sino una herramienta para analizar los datos.

Los principales análisis cuantitativos o estadísticos que pueden efectuarse son:

- Estadística descriptiva para las variables, tomadas individualmente.
- Puntuaciones *z*.
- Razones y tasas.
- Cálculos y razonamientos de estadística inferencial.
- Pruebas paramétricas.
- Pruebas no paramétricas.
- Análisis multivariados.

A continuación hablaremos de estos distintos análisis cuantitativos.

ESTADÍSTICA DESCRIPTIVA PARA CADA VARIABLE

La primera tarea es describir los datos, los valores o las puntuaciones obtenidos para cada variable. Por ejemplo, si aplicamos a 2 048 niños el cuestionario sobre los usos y las graficaciones que la televisión tiene para ellos (Fernández-Collado, Baptista y Elkes, 1998), ¿cómo pueden describirse estos datos? Describiendo la distribución de las puntuaciones o frecuencias.

¿Qué es una distribución de frecuencias?

Una *distribución de frecuencias* es un conjunto de puntuaciones ordenadas en sus respectivas categorías. El cuadro 10.1 muestra un ejemplo de una distribución de frecuencias.

Ejemplo de una distribución de frecuencias	VARIABLES: CONDUCTOR PREFERIDO		
	CATEGORÍAS	CÓDIGOS	FRECUENCIAS
Cuadro 10.1	AMT	1	50
	LEM	2	88
	FGI	3	12
	MML	4	3
	TOTAL		153

VARIABLE: CALIFICACIÓN EN LA PRUEBA DE MOTIVACIÓN	
CATEGORÍAS	FRECUENCIAS
48	1
55	2
56	3
57	5
58	7
60	1
61	1
62	2
63	3
64	2
65	1
66	1
68	1
69	1
73	2
74	1
75	4
76	3
78	2
80	4
82	2
83	1
84	1
86	5
87	2
89	1
90	3
92	1
TOTAL	63

Ejemplo de una distribución que necesita resumirse

Cuadro 10.2

A veces, las *categorías* de las distribuciones de frecuencias son tantas que es necesario resumirlas. Por ejemplo, examinaremos detenidamente la distribución del cuadro 10.2.

Esta distribución podría resumirse o compendiarse como en el cuadro 10.3.

¿Qué otros elementos contiene una distribución de frecuencias?

Las distribuciones de frecuencias pueden completarse agregando las *frecuencias relativas* y las *frecuencias acumuladas*. Las *frecuencias relativas* son los porcentajes de casos en cada categoría, y las *frecuencias acumuladas* son lo que se va acumulando en cada categoría, desde la más baja hasta la más alta. El cuadro 10.4 muestra un ejemplo con las *frecuencias relativas* y *acumuladas*.

Las frecuencias acumuladas constituyen lo que se acumula en cada categoría. En la categoría “sí se ha obtenido la cooperación”, se han acumulado 91. En la categoría “no se ha obtenido la cooperación”, se acumulan 96 (91 de la categoría anterior y 5 de la categoría en cuestión). En la última categoría siempre se acumula el total. Las *frecuencias acumuladas* también pueden expresarse en *porcentajes* (entonces lo que se va acumulando son porcentajes). En el ejemplo del cuadro 10.4 tendríamos, respectivamente:

<i>Categoría</i>	<i>Códigos</i>	<i>Frecuencias acumuladas relativas (%)</i>
— sí	1	74.6%
— no	2	78.7%
— no respondieron	3	100.00%

Ejemplo de una distribución resumida	VARIABLE: CALIFICACIÓN EN LA PRUEBA DE MOTIVACIÓN	
	CATEGORÍAS	FRECUENCIAS
Cuadro 10.3	55 o menos	3
	56-60	16
	61-65	9
	66-70	3
	71-75	7
	76-80	9
	81-85	4
	86-90	11
	91-96	1
	TOTAL	63

VARIABLE: COOPERACIÓN DEL PERSONAL PARA EL PROYECTO DE CALIDAD DE LA EMPRESA					Ejemplo de una distribución de frecuencias con todos sus elementos
CATEGORÍAS	CÓDIGOS	FRECUENCIAS ABSOLUTAS	FRECUENCIAS RELATIVAS (PORCENTAJES)	FRECUENCIAS ACUMULADAS	
– Sí se ha obtenido la cooperación	1	91	74.6%	91	
– No se ha obtenido la cooperación	2	5	4.1%	96	
– No respondieron	3	26	21.3%	122	
TOTAL		122	100.0%		

Cuadro 10.4

Las frecuencias relativas o los porcentajes se calculan así:

$$\text{Porcentaje} = \frac{N_C}{N_T} (100)$$

Donde N_C es el número de casos o frecuencias absolutas en la categoría y N_T es el total de casos. En el ejemplo del cuadro 10.4, tendríamos:

$$\text{Porcentaje} = \frac{91}{122} = 74.59 = 74.6\%$$

$$\text{Porcentaje} = \frac{5}{122} = 4.09 = 4.1\%$$

$$\text{Porcentaje} = \frac{26}{122} = 21.31 = 21.3\%$$

Son resultados que corresponden a los porcentajes del cuadro 10.4.

Al elaborar el reporte de resultados, una distribución se presenta con los elementos más informativos para el lector y la verbalización de los resultados o un comentario, tal como se muestra en el cuadro 10.5.

En el cuadro 10.5 pudieron haberse incluido solamente los porcentajes y eliminarse las frecuencias.

Distribución de frecuencias: conjunto de puntuaciones ordenadas en sus respectivas categorías.

Ejemplo de una distribución para presentar a un usuario

¿SE HA OBTENIDO LA COOPERACIÓN DEL PERSONAL PARA EL PROYECTO DE CALIDAD?

OBTENCIÓN	NÚM. DE ORGANIZACIONES	PORCENTAJES
Sí	91	74.6
No	5	4.1
No respondieron	26	21.3
TOTAL	122	100.0

Cuadro 10.5

COMENTARIO. Prácticamente tres cuartas partes de las organizaciones sí han obtenido la cooperación del personal. Llama la atención que poco más de una quinta parte no quiso comprometerse con su respuesta. Las organizaciones que no han logrado la cooperación del personal mencionaron como factores ausentismo, rechazo al cambio y conformismo.

En los comentarios sobre las distribuciones de frecuencias se utilizan frases como “la mitad de los entrevistados prefiere la marca X” (con 50%), “poco menos de la mitad” de la población mencionó que votarán por el candidato X (por ejemplo, con 48.7%), “casi la tercera parte...” (por ejemplo, con 32.8%), “cuatro de cada diez señoras...” (40%), “solamente uno de cada diez...” (10%), “la mayoría...” (96.7%), etcétera.

¿De qué otra manera pueden presentarse las distribuciones de frecuencias?

Las distribuciones de frecuencias, especialmente cuando utilizamos las frecuencias relativas, pueden presentarse en forma de histogramas o gráficas de otro tipo. Algunos ejemplos se muestran en la figura 10.2.

Actualmente se dispone de gran variedad de programas y paquetes computacionales que elaboran cualquier gráfica, a colores, utilizando efectos de movimiento y en tercera dimensión.

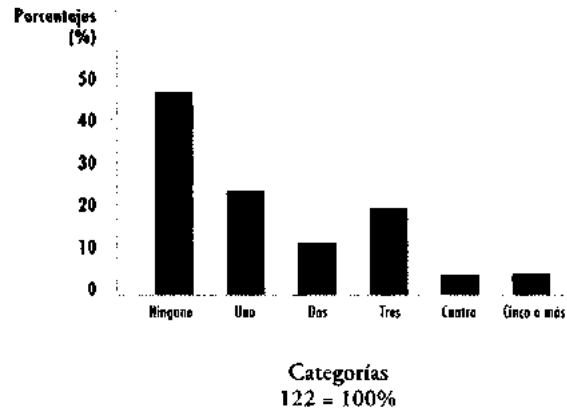
Las distribuciones de frecuencias también se pueden graficar como polígonos de frecuencias

Los *polígonos de frecuencias* relacionan las puntuaciones con sus respectivas frecuencias. Es propio de un nivel de medición por intervalos o razón. La forma de construir un polígono de frecuencias es la siguiente:

- En el eje horizontal (X), se colocan las categorías o los intervalos.
- En el eje vertical (Y), se colocan las frecuencias, dependiendo de cuál es el mayor número posible de frecuencia.

HISTOGRAMAS

Cursos, seminarios o talleres sobre calidad y áreas relacionadas en que han participado los niveles directivos y gerenciales (122 = 100%)



En casi la mitad de las empresas (48.4%), los niveles directivos y gerenciales no han participado en cursos, talleres o seminarios sobre calidad y áreas relacionadas.

GRÁFICAS CIRCULARES

Cooperación de todo el personal (o la mayoría) para el proyecto de calidad (122 = 100%)

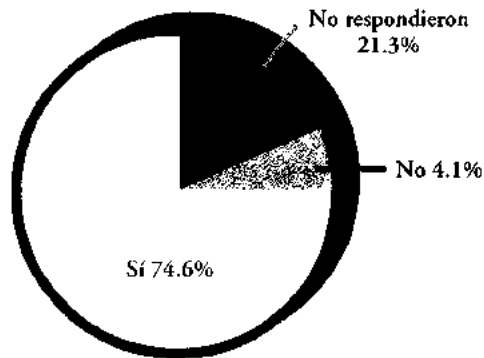
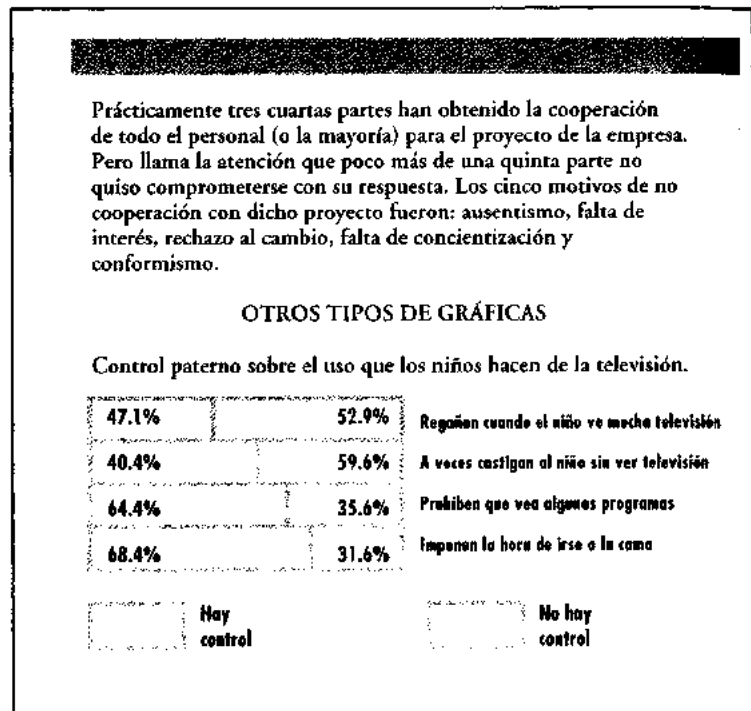


Figura 10.2

Figura 10.2
(continuación)



- c) Se determinan los puntos medios de cada categoría o intervalo. Por ejemplo, si los intervalos fueran 25-29, 30-34, 35-39, etcétera; los puntos medios serían 27, 32, 37, etcétera.
- d) Se ve cuántas frecuencias tiene cada categoría y se traza un punto en la intersección de las frecuencias y los puntos medios de las categorías o los intervalos.
- e) Se unen los puntos trazados en las intersecciones.

La mayoría de los paquetes estadísticos realiza esta labor en forma automática.

Un ejemplo de la elaboración de un polígono de frecuencias se muestra en la figura 10.3.

El polígono de frecuencias obedece a la siguiente distribución:

<i>Categorías/intervalos</i>	<i>Frecuencias absolutas</i>
20-24.9	10
25-29.9	20
30-34.9	35

Ejemplo de la elaboración de un polígono de frecuencias

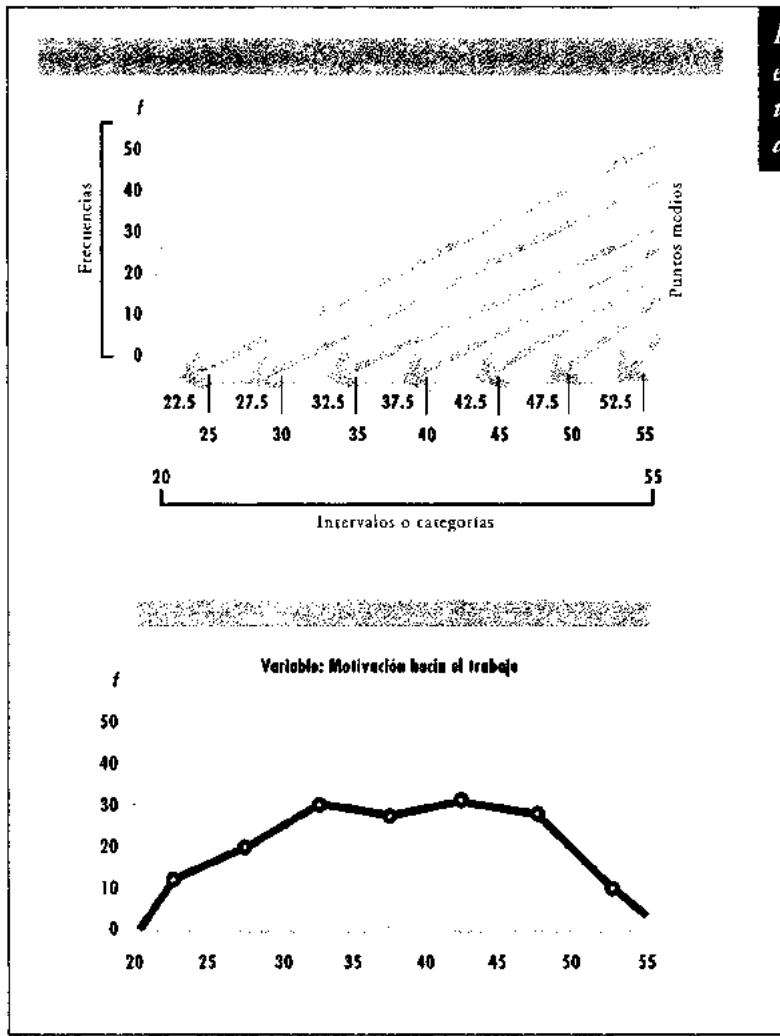


Figura 10.3

35-39.9	33
40-44.9	36
45-49.9	27
50-54.9	8
TOTAL	169

Los polígonos de frecuencias representan curvas útiles para describir los datos; más adelante se hablará de ello.

En resumen, para cada una de las variables de la investigación se obtiene su distribución de frecuencias y, de ser posible, se grafica y se traza su polígono de frecuencias correspondiente.

Pero además del polígono de frecuencias, deben calcularse las medidas de tendencia central y de variabilidad o dispersión.

¿Cuáles son las medidas de tendencia central?

Las medidas de tendencia central son puntos en una distribución, los valores medios o centrales de ésta, y nos ayudan a ubicarla dentro de la escala de medición. Las principales medidas de tendencia central son tres: *moda*, *mediana* y *media*. El nivel de medición de la variable determina cuál es la medida de tendencia central apropiada.

La *moda* es la categoría o puntuación que ocurre con mayor frecuencia. En el cuadro 10.5, la moda es "1" (sí se ha obtenido la cooperación). Se utiliza con cualquier nivel de medición.

La *mediana* es el valor que divide la distribución por la mitad. Esto es, la mitad de los casos caen por debajo de la mediana y la otra mitad se ubica por encima de la mediana. La mediana refleja la posición intermedia de la distribución. Por ejemplo, si los datos obtenidos fueran:

24 31 35 35 38 43 45 50 57

Medidas de tendencia central: valores medios o centrales de una distribución que sirven para ubicarla dentro de la escala de medición.

la mediana es 38, porque deja cuatro casos por encima (43, 45, 50 y 57) y cuatro casos por debajo (35, 35, 31 y 24). Parte a la distribución en dos mitades. En general, para descubrir el caso o la puntuación que constituye la mediana de

una distribución, simplemente se aplica la fórmula: $\frac{N+1}{2}$.

Si tenemos 9 casos, $\frac{9+1}{2} = 5$, entonces buscamos el quinto valor y éste es la mediana. En el ejemplo anterior es 38. Obsérvese que la mediana es el valor observado que se localiza a la mitad de la distribución, no el valor 5. La fórmula no nos proporciona directamente el valor de la mediana, sino el número de caso en donde está la mediana.

La mediana es una medida de tendencia central propia de los niveles de medición ordinal, por intervalos y de razón. No tiene sentido con variables nominales, porque en este nivel no hay jerarquías, ni noción de encima o debajo. También,

la mediana es particularmente útil cuando hay valores extremos en la distribución. No es sensible a éstos. Si tuviéramos los siguientes datos:

24 31 35 35 38 43 45 50 248

la mediana seguiría siendo 38.

Para ejemplificar la interpretación de la media y la mediana, se incluye un artículo al respecto en el cuadro 10.6.²

Moda: categoría o puntuación que se presenta con mayor frecuencia.

Ejemplo de interpretación de la mediana

Cuadro 10.6

¿Qué edad tiene? Si teme contestar no se preocupe, los perfiles de edad difieren de un país a otro.

Con base en el informe anual sobre "El estado de la población mundial" que dio a conocer la Organización de las Naciones Unidas, la población mundial llegó en 1987 a cinco mil millones de habitantes.

El documento señala que la edad media mundial es de 23 años, lo que significa que la mitad de los habitantes del globo terrestre sobrepasa esta mediana y que el otro medio es más joven.

Sin embargo, la mediana de edad de la población mundial se modifica con los años y de acuerdo con las estadísticas recabadas por la ONU la edad central fue de 27 años para el año 2000; y será de 31 años en el 2025. Buena noticia para el actual ciudadano global medio, porque parece ser que se encuentra en la situación de envejecer más lentamente que los demás.

Cabe señalar que la mediana varía de un lugar a otro; en los países en desarrollo la mediana de edad es de 21 años, mientras que en los países industrializados es de 33. Sucede también que en los países pobres la mediana se mantiene más joven, pero al mismo tiempo la esperanza de vida es baja. Para ilustrarlo con un ejemplo, en Kenia la edad promedio de vida es de sólo 54 años de vida, en comparación con Estados Unidos, que es de 75 años.

El informe destaca que los jóvenes y ancianos se engloban en un grupo dependiente; esto significa que son consumidores más que productores de riqueza, y dependen para su sustento de la población económicamente activa, la cual se encuentra entre los 15 y 64 años de edad.

Este factor predomina en los países industrializados, jóvenes y ancianos requieren en gran medida de los servicios gubernamentales que se mantienen con la paga de la población trabajadora. El primer grupo lo necesita durante el trayecto de su escolaridad, en tanto que los segundos tienen derecho a pensiones estatales y a una asistencia médica las más de las veces prolongada. Así, por ejemplo, en países como Francia el gasto público de salud anual por persona es de 694 dólares, en tanto que en Filipinas es de seis dólares.

En Inglaterra las tasas de natalidad son casi nulas, pues su población envejece, y esto puede traer consecuencias económicas serias. Debido al envejecimiento de su población, como sucede en la mayoría de los países europeos, se topa con la difícil situación de atender la fuerte demanda de servicios de salud.

El cuadro de países pobres aún no queda claro, ya que ni los jóvenes ni los ancianos llegan a depender fuertemente de sus gobiernos, porque atienden una mínima parte de los servicios sociales requeridos. Así, tenemos que los niños de esta parte del mundo asisten a la escuela, además de trabajar en las calles para ayudar a su familia al pago de sus útiles

² Leguizamo (1987).

Cuadro 10.6 (continuación)

escolares; en las tribus de Indonesia las abuelas se dedican a las tareas domésticas mientras el resto de la familia trabaja en el campo.

Vemos entonces que la dependencia adopta formas distintas según el tipo de población. Hoy en día se calcula que la tasa de dependencia global es de 65 por cada 100 adultos. Y nuevamente encontramos diferencias marcadas de la relación de dependencia en los países ricos y países pobres: en los primeros es de 50 por cada 100 adultos y en los segundos es de 70 por cada 100 adultos.

De la información que arrojan las estadísticas de población mundial se deduce que los "perfiles de edad" son cruciales para cualquier gobierno, en lo que se refiere al rubro de gasto público, porque, como hemos visto, los países conformados de gente joven requieren de mayor inversión en salud y educación para población infantil y juvenil. Por el contrario, para los conglomerados de ancianos, el gobierno tendrá que destinar dinero para las pensiones y los servicios de salud de largo duración.

El informe mundial de población concluye diciendo que la calidad de los servicios de salud, educación y condiciones de vivienda mejorarían notablemente si las tasas de la población dependiente fueran menos elevadas.

La media es la medida de tendencia central más utilizada y puede definirse como el promedio aritmético de una distribución. Se simboliza como \bar{X} , y es la suma de todos los valores dividida entre el número de casos. Es una medida solamente aplicable a mediciones por intervalos o de razón. Carece de sentido para variables medidas en un nivel nominal u ordinal. Su fórmula es:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + X_4}{N}$$

Por ejemplo, si tuviéramos las siguientes puntuaciones:

8 7 6 4 3 2 6 9 8

la media sería igual a:

$$\bar{X} = \frac{8+7+6+4+3+2+6+9+8}{9} = 5.88$$

La fórmula simplificada de la media es:

$$\bar{X} = \frac{\sum X}{N}$$

El símbolo " \sum " indica que debe efectuarse una sumatoria, X es el símbolo de una puntuación y N es el número total de casos o puntuaciones. En nuestro ejemplo:

$$\bar{X} = \frac{53}{9} = 5.88$$

La media sí es sensible a valores extremos. Si tuviéramos las siguientes puntuaciones:

8 7 6 4 3 2 6 9 20

la media sería:

$$\bar{X} = \frac{65}{9} = 7.22$$

Cálculo de la media o promedio

Cuando en una distribución de frecuencias los datos están agrupados en intervalos, la media se calcula así:

1. Se obtiene el punto medio de cada intervalo:

<i>Intervalos</i>	<i>Puntos medios</i>	<i>Frecuencias</i>
13-15	14	3
10-12	11	4
7-9	8	9
4-6	5	2
1-3	2	1

2. Se multiplica cada punto medio por las frecuencias que le corresponden:

<i>Intervalos</i>	<i>X = Puntos medios</i>	<i>Frecuencias (f)</i>	<i>fx</i>
13-15	14	3	42
10-12	11	4	44
7-9	8	9	72
4-6	5	2	10
1-3	2	1	2
		<u>1</u>	<u>2</u>
		<i>N = 19</i>	<i>∑fx = 170</i>

$\sum fx$ es la sumatoria de la última columna, que corresponde a los puntos medios multiplicados por sus respectivas frecuencias ($14 \times 3 = 42$ y así sucesivamente).

3. Se aplica la siguiente fórmula, para el cálculo de la media con datos agrupados de una distribución de frecuencias:

$$\bar{X} = \frac{\sum fx}{N}$$

En nuestro ejemplo, tenemos:

$$\bar{X} = \frac{170}{19} = 8.95$$

¿Cuáles son las medidas de la variabilidad?

Las medidas de la variabilidad indican la dispersión de los datos en la escala de medición y responden a la pregunta: ¿Dónde están diseminadas las puntuaciones o los valores obtenidos? Las medidas de tendencia central son valores en una distribución y las medidas de la variabilidad son intervalos, designan distancias o un número de unidades en la escala de medición. Las medidas de la variabilidad más utilizadas son *rango*, *desviación estándar* y *varianza*.

El *rango*, también llamado *recorrido*, es la diferencia entre la puntuación mayor y la puntuación menor, e indica el número de unidades en la escala de medición necesario para incluir los valores máximo y mínimo. Se calcula así: $X_M - X_m$ (puntuación mayor, menos puntuación menor). Si tenemos los siguientes valores:

17 18 20 20 24 28 28 30 33

el rango será: $33 - 17 = 16$.

Cuanto más grande sea el *rango*, mayor será la *dispersión de los datos* de una distribución.

La *desviación estándar* es el promedio de desviación de las puntuaciones con respecto a la media. Esta medida se expresa en las unidades originales de medición de la distribución. Se interpreta en relación con la media. Cuanto mayor sea la dispersión de los datos alrededor de la media, mayor será la desviación estándar. Se simboliza con: s o la sigma minúscula (σ) y su fórmula esencial es:

$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{N}}$$

Esto es, la desviación de cada puntuación respecto a la media se eleva al cuadrado, se suman todas las desviaciones cuadradas, se divide entre el número total de puntuaciones, y a esta división se le saca raíz cuadrada.

Procedimientos para calcular la desviación estándar

El procedimiento para calcularla es el siguiente:

1. Se ordenan las puntuaciones. Por ejemplo:

variable: calificación en estadística social

X (Puntuaciones)
9
7
6
6
5
4
3

Rango: diferencia entre las puntuaciones mayor y menor que indica el número necesario de unidades en la escala de medición para incluir los valores máximo y mínimo.

2. Se calcula la media:

$$\bar{X} = \frac{9+7+6+6+5+4+3}{7} = 5.71$$

3. Se determina la desviación de cada puntuación con respecto a la media:

X	$X - \bar{X}$
9	3.29
7	1.29
6	0.29
6	0.29
5	-0.71
4	-1.71
3	-2.71
<hr/>	
$\Sigma X = 40$	

4. Se eleva al cuadrado cada desviación y se obtiene la sumatoria de las desviaciones elevadas al cuadrado o $\Sigma(X - \bar{X})^2$.

X	$(X - \bar{X})^2$
9	10.82
7	1.66
6	0.08
6	0.08
5	0.50
4	2.92
3	7.34
<hr/>	
$\Sigma X = 40$	$\Sigma(X - \bar{X})^2 = 23.40$

5. Se aplica la fórmula:

$$s = \frac{\sqrt{23.40}}{7} = \sqrt{3.34}$$

$$s = 1.83$$

Cuando los datos están agrupados en una distribución de frecuencias, se procede así:

1. Se obtiene el punto medio de cada intervalo y se determina la media de la distribución (con la fórmula para datos agrupados):

Intervalos	Puntos medios	Frecuencias	fx
13-15	14	3	42
10-12	11	4	44
7-9	8	9	72
4-6	5	2	10
1-3	2	1	2
		$N = 19$	$fx = 170$

$$\bar{X} = \frac{\sum fx}{N} = \frac{170}{19} = 8.95$$

2. Se eleva la media al cuadrado:

$$\bar{X}^2 = (8.95)^2 = 80.1$$

3. Se multiplica la columna fx por los puntos medios, y se obtiene una columna, que llamaremos fx^2 , así como la sumatoria de esta última columna:

Intervalos	Puntos medios	fx	fx^2
13-15	14	42	588
10-12	11	44	484
7-9	8	72	576
4-6	5	10	50
1-3	2	2	4
			$\sum fx^2 = 1702$

Obsérvese que cada valor de la última columna fx^2 se obtiene multiplicando un punto medio por su respectivo valor en la columna fx .

4. Se aplica la siguiente fórmula para la desviación estándar con datos agrupados en una distribución de frecuencias:³

³ Levin (1979, p. 70).

$$s = \sqrt{\frac{\sum fx^2}{N} - \bar{X}^2}$$

$$s = \sqrt{\frac{1702}{19} - 80.1}$$

$$s = \sqrt{89.58 - 80.1}$$

$$s = \sqrt{9.48}$$

$$s = 3.08$$

La desviación estándar se interpreta como "cuánto se desvía, en promedio, de la media un conjunto de puntuaciones".

Supongamos que un investigador obtuvo para su muestra una media de ingreso familiar de \$8 000 (ocho mil pesos) y una desviación estándar de \$1 000 (un mil pesos). La interpretación es que los ingresos familiares de la muestra se desvían, en promedio, mil pesos respecto a la media.

La desviación estándar sólo se utiliza en variables medidas por intervalos o de razón.

Desviación estándar: promedio de desviación de las puntuaciones con respecto a la media que se expresa en las unidades originales de medición de la distribución.

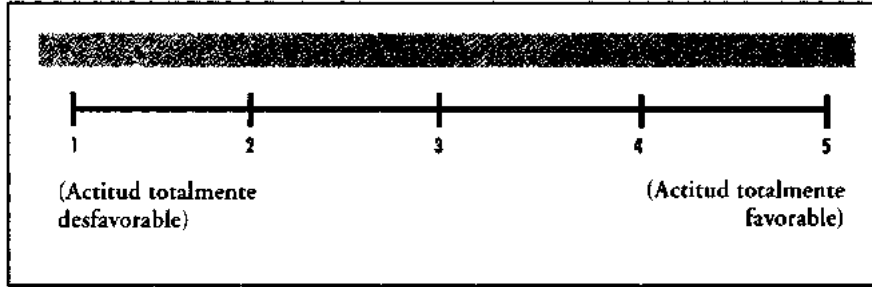
La varianza

La *varianza* es la desviación estándar elevada al cuadrado y se simboliza s^2 . Es un concepto estadístico muy importante, ya que muchas de las pruebas cuantitativas se fundamentan en él. Diversos métodos estadísticos parten de la descomposición de la varianza. Sin embargo, con fines descriptivos se utiliza preferentemente la desviación estándar.

¿Cómo se interpretan las medidas de tendencia central y de la variabilidad?

Cabe destacar que al describir nuestros datos, interpretamos las medidas de tendencia central y de la variabilidad en conjunto, no aisladamente. Tomamos en cuenta todas las medidas. Para interpretarlas, lo primero que hacemos es tomar en cuenta el rango potencial de la escala. Supongamos que aplicamos una escala de actitudes del tipo Likert para medir la "actitud hacia el presidente" de una nación (digamos que la escala tuviera 18 ítems y se promediaran sus resultados). El rango potencial es de 1 a 5:

Varianza: desviación estándar elevada al cuadrado.



Si obtuviéramos los siguientes resultados:

Variable: actitud hacia el presidente

Moda: 4.0

Mediana: 3.9

Media (\bar{X}): 4.2

Desviación estándar: 0.7

Puntuación más alta observada (máximo): 5.0

Puntuación más baja observada (mínimo): 2.0

Rango: 3

Podríamos hacer la siguiente interpretación descriptiva: la actitud hacia el presidente es favorable. La categoría que más se repitió fue 4 (favorable). El 50% de los sujetos está por encima del valor 3.9 y el restante 50% se sitúa por debajo de este valor. En promedio, los sujetos se ubican en 4.2 (favorable). Asimismo, se desvían de 4.2, en promedio, 0.7 unidades de la escala. Ninguna persona calificó al presidente de manera desfavorable (no hay "1"). Las puntuaciones tienden a ubicarse en valores medios o elevados.

En cambio, si los resultados fueran:

Variable: actitud hacia el presidente

Moda: 1

Mediana: 1.5

Media (\bar{X}): 1.3

Desviación estándar: 0.4

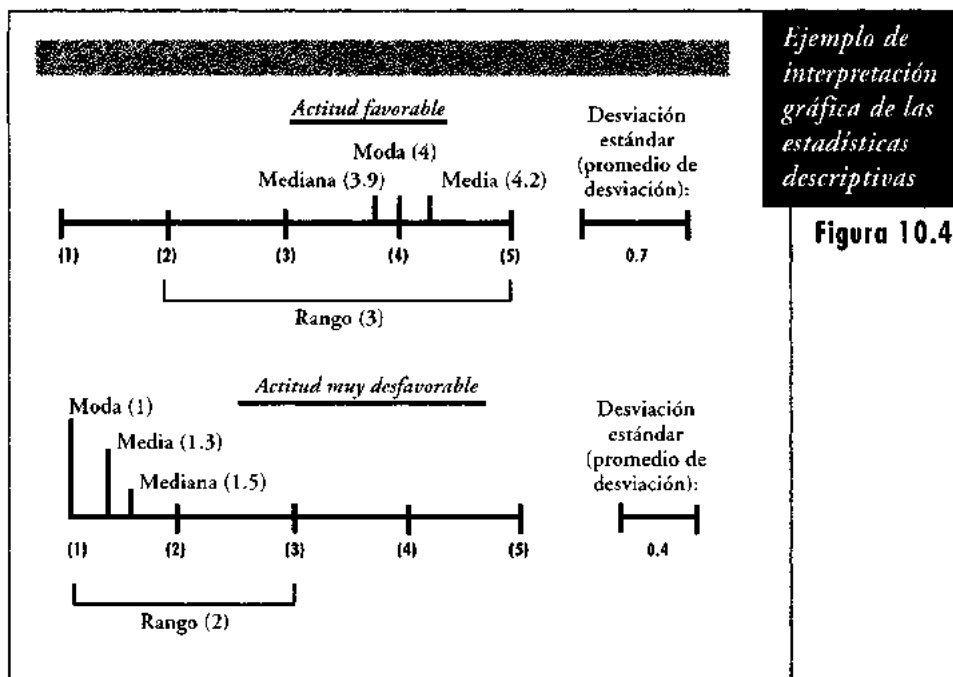
Varianza: 0.16

Máximo: 3.0

Mínimo: 1.0

Rango: 2.0

La interpretación es que la actitud hacia el presidente es muy desfavorable. En la figura 10.4 vemos gráficamente la comparación de resultados.



Ejemplo de interpretación gráfica de las estadísticas descriptivas

Figura 10.4

La variabilidad también es menor en el caso de la actitud muy desfavorable (los datos se encuentran menos dispersos).

En el cuadro 10.7 se presenta otro ejemplo de interpretación con una prueba de motivación intrínseca aplicada a 60 sujetos de un experimento (Hernández-Sampieri y Cortés, 1982). La escala tiene 17 ítems (con cinco opciones cada uno, 1 a 5) y mide la motivación intrínseca al ejecutar una tarea.

El nivel de motivación intrínseca exhibido por los sujetos tiende a ser elevado, como lo indican los resultados de la escala. El rango real de la escala iba de 17 a 85. El rango resultante para esta investigación varió de 40 a 81. Por lo tanto, es evidente que los sujetos se inclinaron hacia valores elevados en la medida de motivación intrínseca. Además, la media de los participantes es de 66.9 y la mediana de 67.8, lo cual confirma la tendencia de la muestra hacia valores altos de la escala. A pesar de que la dispersión de las puntuaciones de los sujetos es alta (la desviación estándar es igual a 9.1 y el rango es de 41), esta dispersión se manifiesta en el área más elevada de la escala. Veámoslo gráficamente:

Es decir, aunque las puntuaciones varían de 40 a 81 y la desviación estándar es de 9.1 (la media sobre la cual gravita es de 66.9), esta variación se da en la parte de los valores más altos de la escala. En resumen, la tarea resultó intrínseca-

Ejemplo de interpretación de una distribución de frecuencias

VARIABLE: MOTIVACIÓN INTRÍNSECA

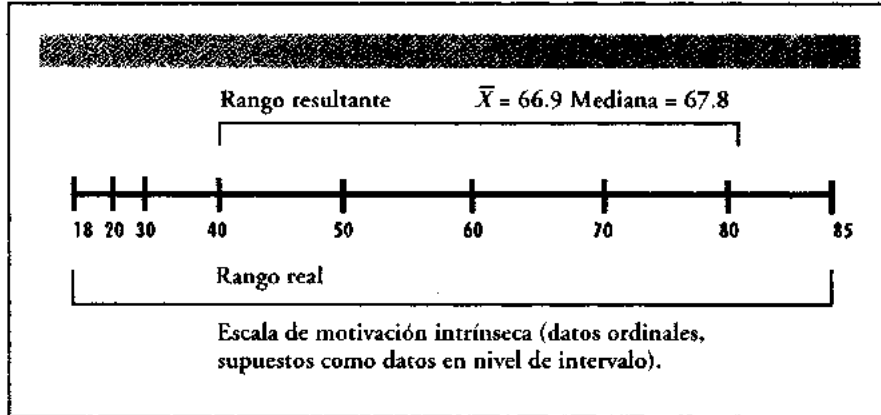
¿Qué grado de motivación intrínseca exhibieron los sujetos?

Número de ítems = 17

Cuadro 10.7

VALORES REGISTRADOS EN LA ESCALA DE MOTIVACIÓN INTRÍNSECA	FRECUENCIAS ABSOLUTAS	FRECUENCIAS RELATIVAS (%)	FRECUENCIAS AJUSTADAS (%)	FRECUENCIAS ACUMULADAS (%)
40	1	1.7	1.7	1.7
44	1	1.7	1.7	3.3
48	1	1.7	1.7	5.0
51	1	1.7	1.7	6.7
52	2	3.3	3.3	10.0
56	2	3.3	3.3	13.3
58	1	1.7	1.7	15.0
59	1	1.7	1.7	16.7
60	2	3.3	3.3	20.0
61	4	6.7	6.7	26.7
63	2	3.3	3.3	30.0
64	2	3.3	3.3	33.3
65	3	5.0	5.0	38.3
66	2	3.3	3.3	41.7
67	4	6.7	6.7	48.3
68	3	5.0	5.0	53.3
69	1	1.7	1.7	55.0
70	4	6.7	6.7	61.7
71	3	5.0	5.0	66.7
72	4	6.7	6.7	73.3
73	3	5.0	5.0	78.3
74	2	3.3	3.3	81.7
75	1	1.7	1.7	83.3
76	1	1.7	1.7	85.0
77	2	3.3	3.3	88.3
78	1	1.7	1.7	90.0
79	2	3.3	3.3	93.3
80	2	3.3	3.3	96.7
81	2	3.3	3.3	100.0
Total	60	100.0	100.0	

Media = 66.883	E.E. = 1.176	Mediana = 67.833
Moda = 61.000	s = 9.112	Varianza = 83.020
Curtosis = .587	Asimetría = -.775	Rango = 41.000
Mínimo = 40.000	Máximo = 81.000	Sumatoria = 4 013.000



mente motivante para la mayoría de los sujetos, sólo que para algunos resultó muy motivante; para otros, relativamente motivante; y para los demás, medianamente motivante, o sea que la tendencia general es hacia valores altos (observamos la columna de frecuencias acumuladas y notamos que 80% obtuvo puntuaciones superiores a 60).

En el cuadro 10.7 se introduce un nuevo elemento: la columna de frecuencias ajustadas. Ésta incluye las frecuencias relativas, pero desechando los valores perdidos (*missing values*). Cuando no hay valores perdidos, las columnas de frecuencia relativas y de frecuencias ajustadas son equivalentes.

Ahora bien, ¿qué significa un alto nivel de motivación intrínseca exhibido con respecto a una tarea? Significa que la tarea fue percibida como atractiva, interesante, divertida y categorizada como una experiencia agradable. Asimismo, implica que los sujetos, al estar ejecutándola, derivaron de ella sentimientos de satisfacción, goce y realización personal. Por lo general, quien se encuentra intrínsecamente motivado hacia una labor, la habrá de disfrutar, ya que obtendrá de la labor *per se* recompensas internas, como sentimientos de logro y autorrealización. Además de ser absorbido por el desarrollo de la tarea y, al tener un buen desempeño, la opinión de sí mismo mejorará o se verá reforzada.

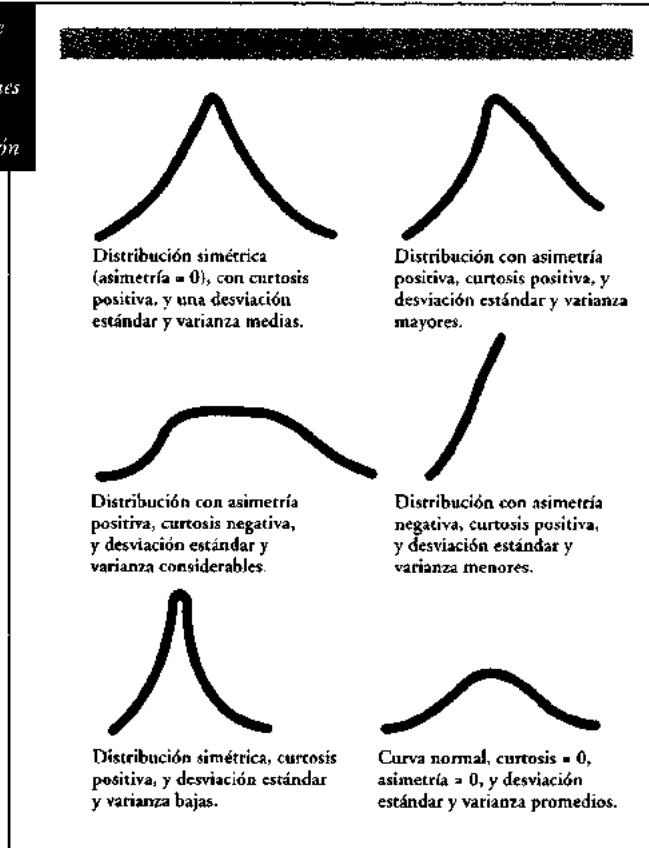
¿Hay alguna otra estadística descriptiva?

Sí, la *asimetría* y la *curtosis*. Los polígonos de frecuencia suelen representarse como *curvas* (figura 10.5) para que puedan analizarse en términos de probabilidad y visualizar su grado de dispersión. De hecho, en realidad son curvas. Dos elementos son esenciales para estas curvas o dichos polígonos de frecuencias: la asimetría y la curtosis.

La *asimetría* es una estadística necesaria para conocer cuánto se parece nuestra distribución a una distribución teórica llamada *curva normal* (la cual se repre-

Ejemplos de curvas o distribuciones y su interpretación

Figura 10.5



Asimetría: estadística que se usa para conocer cuánto se parece una distribución a la distribución teórica llamada *curva normal* y que constituye un indicador del lado de la curva en el que las frecuencias se agrupan más.

señala en la figura 10.5) y constituye un indicador del lado de la curva donde se agrupan las frecuencias. Si es cero (asimetría = 0), la curva o distribución es simétrica. Cuando es positiva, quiere decir que hay más valores agrupados hacia la izquierda de la curva (por debajo de la media). Cuando es negativa, significa que los valores tienden a agruparse hacia la derecha de la curva (por encima de la media).

La *curtosis* es un indicador de lo plana o “picuda” que es una curva. Cuando es cero ($\text{curtosis} = 0$), significa que se trata de una *curva normal*. Si es positiva, quiere decir que la curva, la distribución o el polígono es más “picuda(o)” o levantada(o). Si es negativa(o), que es más plana.

La asimetría y la curtosis requieren mínimo de un nivel de medición por intervalos. En la figura 10.5 se muestran ejemplos de curvas con su interpretación.

¿Cómo se traducen las estadísticas descriptivas al inglés?

Algunos programas y paquetes estadísticos computacionales pueden realizar el cálculo de las estadísticas descriptivas, cuyos resultados aparecen junto al nombre respectivo de éstas muchas veces en inglés. A continuación se indican las diferentes estadísticas y su equivalente en inglés.

Curtosis: indicador de lo plana o *picuda* que es una curva.

<i>Estadística</i>	<i>Equivalente en inglés</i>
— Moda	— Mode
— Mediana	— Median
— Media	— Mean
— Desviación estándar	— Standard deviation
— Varianza	— Variance
— Máximo	— Maximum
— Mínimo	— Minimum
— Rango	— Range
— Asimetría	— Skewness
— Curtosis	— Kurtosis

Nota final

Debe recordarse que en una investigación se obtiene una distribución de frecuencias para cada variable y se calculan las estadísticas descriptivas para cada variable: se calculan las que se necesiten de acuerdo con los propósitos de la investigación y los niveles de medición.

Puntuaciones z

Las *puntuaciones z* son transformaciones que se pueden hacer a los valores o las puntuaciones obtenidas, con el propósito de analizar su distancia respecto a la media, en unidades de desviación estándar. *Una puntuación z nos indica la direc-*

ción y el grado en que un valor individual obtenido se aleja de la media, en una escala de unidades de desviación estándar. Como mencionan Nie et al. (1975), las puntuaciones z son el método más comúnmente utilizado para estandarizar la escala de una variable medida en un nivel por intervalos.

Su fórmula es:

$$z = \frac{X - \bar{X}}{s}$$

donde X es la puntuación o el valor a transformar, \bar{X} es la media de la distribución y s la desviación estándar de ésta. El resultado z es la puntuación transformada en unidades de desviación estándar.

Supongamos que en una distribución de frecuencias obtuvimos una media de 60 y una desviación estándar de 10, y deseamos comparar a una puntuación de “50” con el resto de la distribución. Entonces, transformamos esta puntuación o tal valor en una puntuación z . Tenemos que:

$$X = 50$$

$$\bar{X} = 60$$

$$s = 10$$

La puntuación z correspondiente a un valor de “50” es:

$$z = \frac{50 - 60}{10} = -1.00$$

Podemos decir que el valor “50” se localiza a una desviación estándar por debajo de la media de la distribución (el valor “30” está a tres desviaciones estándar por debajo de la media).

Estandarizar los valores permite comparar puntuaciones de dos distribuciones diferentes (la forma de medición es la misma, aunque se trata de distribuciones distintas). Por ejemplo, podemos comparar una distribución obtenida en una preprueba con otra obtenida en una posprueba (en un contexto experimental). Supongamos que se trata de un estímulo que incrementa la productividad. Un trabajador obtuvo en la preprueba una productividad de 130 (la media del grupo fue de 122.5 y la desviación estándar de 10). Y en la posprueba obtuvo 135 (la media del grupo fue de 140 y la desviación estándar de 9.8). ¿Mejóro la productividad del trabajador? En apariencia la mejoría no es considerable. Sin transformar las dos calificaciones en puntuaciones z no es posible asegurarlo porque los valores no pertenecen a la misma distribución. Entonces transformamos ambos valores a puntuaciones z , los pasamos a una escala común, donde la

comparación es válida. El valor de 130 en productividad en términos de unidades de desviación estándar es igual a:

$$z = \frac{130 - 122,5}{10,0} = 0,75$$

Y el valor de 135 corresponde a una puntuación z de:

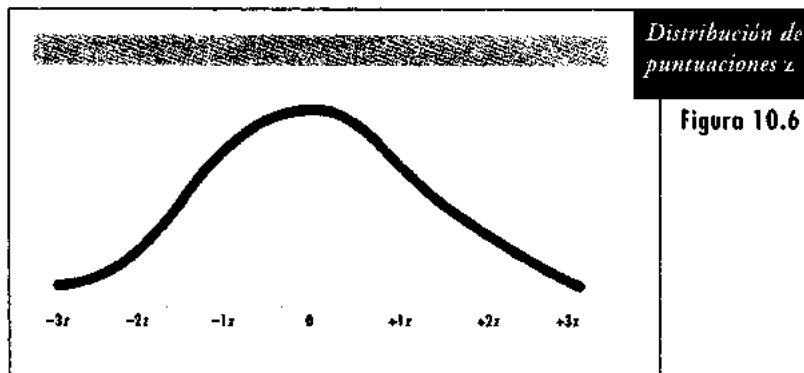
$$z = \frac{135 - 140}{9,8} = -0,51$$

Como observamos, en términos absolutos 135 es una mejor puntuación que 130, pero no en términos relativos (en relación con sus respectivas distribuciones).

La distribución de puntuaciones z no cambia la forma de la distribución original, pero sí modifica las unidades originales a "unidades de desviación estándar" (Wright, 1979). La distribución de puntuaciones z tiene una media de 0 (cero) y una desviación estándar de 1 (uno). La figura 10.6 muestra la distribución de puntuaciones z .

Las puntuaciones z también sirven para comparar mediciones de distintas pruebas o escalas aplicadas a los mismos sujetos (los valores obtenidos en cada escala se transforman a puntuaciones z y se comparan). No debe olvidarse que en la fórmula se tratan de la media y la desviación estándar que corresponden al valor a transformar (de su misma distribución). También, las puntuaciones z sirven para analizar distancias, entre puntuaciones de una misma distribución y áreas de la curva que abarcan tales distancias, o para sopesar el desempeño de un grupo de sujetos en varias pruebas.

Puntuación z : medida que indica la dirección y el grado en que un valor individual se aleja de la media, en una escala de unidades de desviación estándar.



Las puntuaciones z son un elemento descriptivo adicional que se agrega para analizar nuestros datos.

Razones y tasas

Una razón es la relación entre dos categorías. Por ejemplo:

Categorías	Frecuencias absolutas
Masculino	60
Femenino	30

La razón de hombres a mujeres es de $\frac{60}{30} = 2$. Es decir, por cada dos hombres hay una mujer.

Razón: relación entre dos categorías.

Una tasa es la relación entre el número de casos, frecuencias o eventos de una categoría y el número total de observaciones, multiplicada por un múltiplo de 10, generalmente 100 o 1 000. La fórmula es:

$$\text{Tasa} = \frac{\text{Número de eventos durante un periodo}}{\text{Número total de eventos posibles}} \times 100 \text{ o } 1\,000$$

$$\text{Ejemplo: } \frac{\text{Número de nacidos vivos en la ciudad}}{\text{Número de habitantes en la ciudad}} \times 1\,000$$

$$\text{Tasa de nacidos vivos en Tingüindín: } \frac{10\,000}{300\,000} \times 1\,000 = 33.33$$

Es decir, hay 33.33 nacidos vivos por cada 1 000 habitantes en Tingüindín.

ESTADÍSTICA INFERENCIAL: DE LA MUESTRA A LA POBLACIÓN

¿Para qué es útil la estadística inferencial?

Con frecuencia, el propósito de la investigación va más allá de describir las distribuciones de las variables; se pretende generalizar los resultados obtenidos en la muestra a la población o el universo.⁴ Los datos casi siempre se recolectan de una muestra y sus resultados estadísticos se denominan "estadígrafos"; la media o la des-

⁴ Los conceptos de muestra y población se explicaron en el capítulo 8.

viación estándar de la distribución de una muestra son estadígrafos. *A las estadísticas de la población o al universo se les conoce como "parámetros"*. Los parámetros no son calculados, porque no se recolectan datos de toda la población, pero pueden ser inferidos de los estadígrafos, de ahí el nombre de *estadística inferencial*. El procedimiento de esta naturaleza de la estadística se esquematiza en la figura 10.7.

La inferencia de los parámetros se lleva a cabo mediante técnicas estadísticas apropiadas. Estas técnicas se explicarán más adelante.

La estadística inferencial se utiliza para dos procedimientos (Wiersma, 1999, p. 335):

- a) *Probar hipótesis*
- b) *Estimar parámetros*

¿En qué consiste la prueba de hipótesis?

Una *hipótesis* en el contexto de la estadística inferencial es una proposición respecto a uno o varios parámetros, y lo que el investigador hace a través de la prueba de hipótesis es determinar si la hipótesis es congruente con los datos obtenidos en la muestra (Wiersma, 1999). Si es congruente con los datos, ésta se retiene como un valor aceptable del parámetro. Si la hipótesis no lo es, se rechaza (pero los datos no se descartan) (Wiersma, 1999). Para comprender lo que es la prueba de hipótesis en la estadística inferencial es necesario revisar el concepto de distribución muestral⁵ y nivel de significancia.

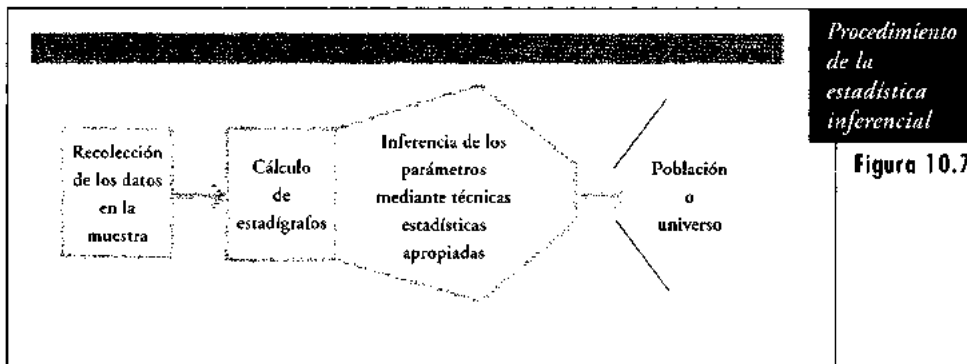


Figura 10.7

⁵ Distribución muestral y distribución de una muestra son conceptos diferentes; esta última es resultado de los datos de nuestra investigación.

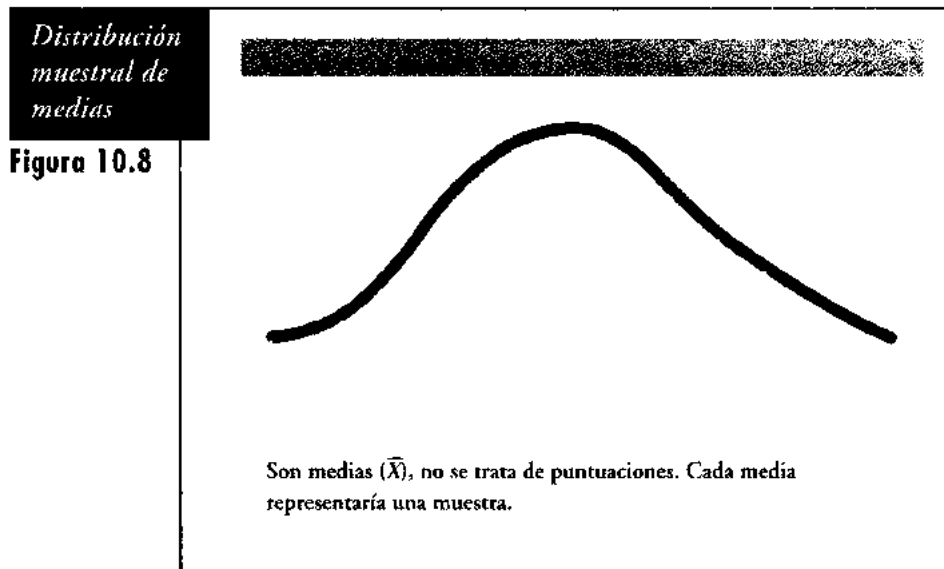
¿Qué es una distribución muestral?

Una *distribución muestral* es un conjunto de valores sobre una estadística calculada de todas las muestras posibles de determinado tamaño (Wiersma, 1999). Las distribuciones muestrales de medias son probablemente las más conocidas. Expliquemos este concepto con un ejemplo. Supongamos que nuestro universo o la población son los automovilistas de una ciudad y deseamos averiguar cuánto tiempo pasan diariamente “al volante”. De este universo podría extraerse una muestra representativa. Vamos a suponer que el tamaño adecuado de muestra es de quinientos doce automovilistas ($n = 512$). Del mismo universo se podrían extraer diferentes muestras, cada una con 512 personas.

Teóricamente, incluso podría hacerse al azar una, dos, tres, cuatro y las veces que fuera necesario, hasta agotar todas las muestras posibles de 512 automovilistas de esa ciudad (todos los sujetos serían seleccionados en varias muestras). En cada muestra se obtendría una media del tiempo que pasan los automovilistas manejando. Tendríamos pues, una gran cantidad de medias, tantas como las muestras extraídas ($\bar{X}_1, \bar{X}_2, \bar{X}_3, \bar{X}_4, \bar{X}_5, \dots, \bar{X}_k$). Y con estas medias elaboraríamos una distribución de medias. Habría muestras que, en promedio, pasaran más tiempo “al volante” que otras. Este concepto se representa en la figura 10.8.

Si calculáramos la media de todas las medias de las muestras, obtendríamos el valor de la media poblacional.

Muy rara vez se obtiene la *distribución muestral* (la distribución de las medias de todas las muestras posibles). Es más bien un concepto teórico definido



por la estadística para los investigadores. Lo que comúnmente hacemos es extraer una sola muestra.

En el ejemplo de los automovilistas, sólo una de las líneas verticales de la distribución muestral presentada en la figura 10.8 es la media obtenida para la única muestra seleccionada de 512 personas. Y la pregunta es: ¿nuestra media está cerca de la media de la distribución muestral? (o lo que es igual: ¿la media de la muestra está cercana a la media de la distribución muestral?), debido a que si está cerca podremos tener una estimación precisa de la media poblacional (el parámetro poblacional es prácticamente el mismo que el de la distribución muestral). Esto se expresa en el *teorema central del límite*, el cual se explicó en el capítulo sobre muestreo. Recordemos que dice: “Si una población (no necesariamente normal) tiene de media m y de desviación estándar σ , la distribución de las medias en el muestreo aleatorio realizado en esta población tiende, al aumentar n , a una distribución normal de media m y desviación estándar σ/\sqrt{n} , donde n es el tamaño de muestra.”

El teorema especifica que la distribución muestral tiene una media igual a la de la población, una varianza igual a la varianza de la población dividida entre el tamaño de muestra (su desviación estándar es σ/\sqrt{n} , y se distribuye normalmente) (Wiersma, 1999). σ es un parámetro normalmente desconocido, aunque es posible estimarlo por la desviación estándar de la muestra.

El concepto de *distribución normal* es importante otra vez y se ofrece una breve explicación en la figura 10.9.

Tasa: relación entre el número de frecuencias de una categoría y el número total de observaciones, multiplicada por un múltiplo de 10, casi siempre 100 o 1 000.

¿Qué es el nivel de significancia?

Wiersma (1999) ofrece una explicación sencilla del concepto, en la cual nos basaremos para analizar su significado.

La probabilidad de que un evento ocurra oscila entre 0 y 1, donde 0 significa la imposibilidad de ocurrencia y 1 la certeza de que el fenómeno ocurra. Al lanzar al aire una moneda no cargada, la probabilidad de que salga “cruz” es 0.50 y la probabilidad de que la moneda caiga en “cara” también es de 0.50. Con un dado, la probabilidad de obtener cualquiera de sus caras al lanzarlo es de $1/6 = 0.1667$. La suma de posibilidades siempre es de 1.

Aplicando el concepto de probabilidad a la distribución muestral, tomaremos el área de ésta como 1.00; en consecuencia, cualquier área comprendida entre dos puntos de la distribución corresponderá a la probabilidad de la distribución. Para probar hipótesis inferenciales respecto a la media, el investigador debe evaluar si es alta o baja la probabilidad de que la media de la muestra esté cerca de la media de la distribución muestral. Si es baja, el investigador dudará

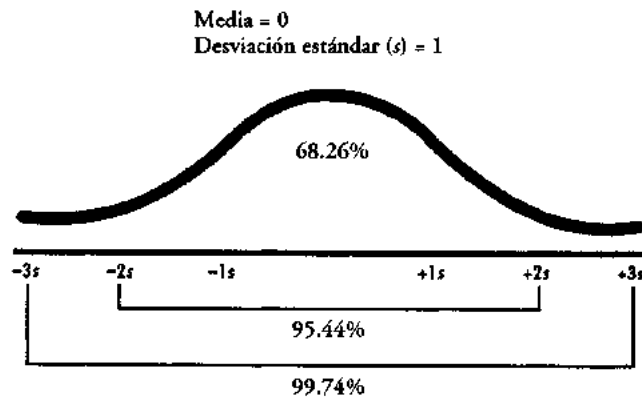
Concepto
de curva o
distribución
normal

Figura 10.9

Una gran cantidad de los fenómenos del comportamiento humano se manifiestan de la siguiente forma: la mayoría de las puntuaciones se concentran en el centro de la distribución, en tanto que en los extremos encontramos sólo algunas puntuaciones. Por ejemplo, la inteligencia: hay pocas personas muy inteligentes (genios), pero también hay pocas personas con muy baja inteligencia (por ejemplo, retardados mentales). La mayoría de los seres humanos somos medianamente inteligentes. Esto podría representarse así:



Debido a ello, se creó un modelo de probabilidad llamado curva normal o distribución normal. Como todo modelo es una distribución teórica que difícilmente se presenta en la realidad tal cual, pero sí se presentan aproximaciones a éste. La curva normal tiene la siguiente configuración:



68.26% del área de la curva normal es cubierta entre $-1s$ y $+1s$,
95.44% del área de esta curva es cubierta entre $-2s$ y $+2s$ y
99.74% se cubre con $-3s$ y $+3s$.

Las principales características de la distribución normal son:

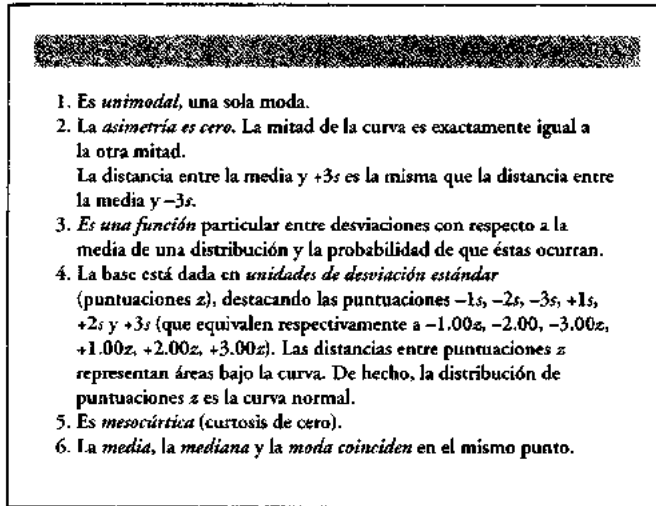


Figura 10.9
(continuación)

de generalizar a la población. Si es alta, el investigador podrá hacer generalizaciones. Es aquí donde entra *el nivel de significancia o nivel alfa* (nivel α : es un nivel de la probabilidad de equivocarse y se fija antes de probar hipótesis inferenciales.

Se explicará esto con un ejemplo coloquial.

Si fuera a apostar en las carreras de caballos y tuviera 95% de probabilidades de atinarle al ganador, contra sólo 5% de perder, ¿apostaría? Obviamente sí, siempre y cuando le aseguraran ese 95% en favor. O bien, si le dieran 95 boletos de 100 para la rifa de un automóvil, ¿sentiría confianza en que va a estrenar vehículo? Por supuesto que sí. No tendría la certeza total; ésta no existe en el universo, al menos para los seres humanos.

Pues bien, algo similar hace el investigador social. Obtiene una estadística en una muestra (por ejemplo, la media) y analiza qué porcentaje tiene de confianza en que dicha estadística se acerque al valor de la distribución muestral (que es el valor de la población o el parámetro). Busca un alto porcentaje de confianza, una probabilidad elevada para estar tranquilo, porque sabe que tal vez haya error de muestreo y, aunque la evidencia parece mostrar una aparente “cercanía” entre el valor calculado en la muestra y el parámetro, tal “cercanía” puede no ser real o deberse a errores en la selección de la muestra.

¿Con qué porcentaje de confianza el investigador generaliza, para suponer que tal cercanía es real y no por un error de muestreo? *Existen dos niveles convenidos en ciencias sociales:*

Nivel de significancia: valor de certeza, respecto de no equivocarse, que el investigador fija *a priori*.

- a) *El nivel de significancia de .05*, el cual implica que el investigador tiene 95% de seguridad para generalizar sin equivocarse y sólo 5% en contra. En términos de probabilidad, .95 y .05, respectivamente; ambos suman la unidad.
- b) *El nivel de significancia de .01*, que implica que el investigador tiene 99% en su favor y 1% en contra (.99 y .01 = 1.00) para generalizar sin temor.

A veces el nivel de significancia puede ser todavía más riguroso y confiable (por ejemplo, .001, .00001, .00000001), pero lo mínimo es .05. No se acepta un nivel de .06 (94% a favor de la generalización confiable), porque se busca hacer ciencia, no intuición.

El nivel de significancia es un valor de certeza que el investigador fija a priori, respecto a no equivocarse. Volveremos más adelante sobre este punto.

¿Cómo se relacionan la distribución muestral y el nivel de significancia?

El nivel de significancia se expresa en términos de probabilidad (.05 y .01) y la distribución muestral también como probabilidad (el área total de ésta como 1.00). Pues bien, para ver si existe o no confianza al generalizar acudimos a la distribución muestral, una probabilidad apropiada para la investigación social. El nivel de significancia lo tomamos como un área bajo la distribución muestral, como se observa en la figura 10.10, dependiendo de si elegimos un nivel de .05 o de .01.

Así, el nivel de significancia representa áreas de riesgo o confianza en la distribución muestral.

Una vez definido el nivel de significancia, ¿qué hacemos para ver si nuestra hipótesis sobre la media poblacional es aceptada o rechazada?

Antes de estudiar el *procedimiento*, es necesario hacer las siguientes consideraciones:

- a) La distribución muestral es una distribución normal de puntuaciones z , la base de la curva son puntuaciones z o unidades de desviación estándar.
- b) Las puntuaciones z son distancias que indican áreas bajo la distribución normal. En este caso, áreas de probabilidad.
- c) El área de riesgo es tomada como el área de rechazo de la hipótesis; y el área de confianza, como el área de aceptación de la hipótesis.
- d) Se habla de una hipótesis acerca del parámetro (en este caso, media poblacional).

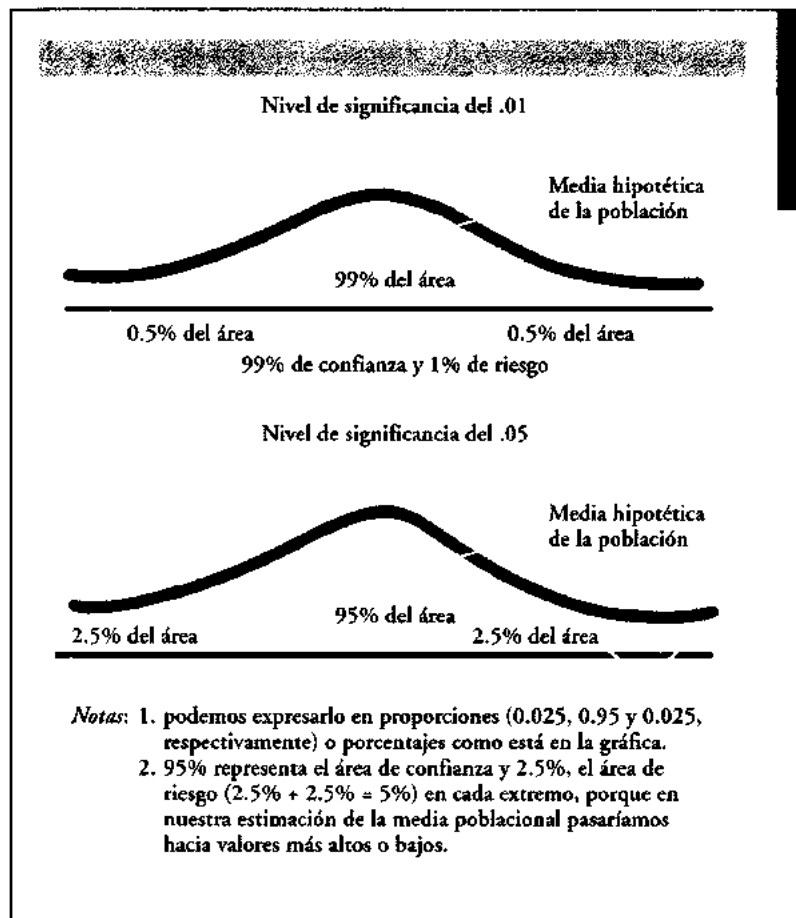


Figura 10.10

Partiendo de estas consideraciones, el procedimiento es:

1. Sobre bases firmes (revisión de la literatura e información disponible), establecer una hipótesis acerca del parámetro poblacional. Por ejemplo: "El promedio de horas diarias que se exponen los niños de la ciudad de Celaya a la televisión en fin de semana es de 3.0."
2. Definir el nivel de significancia. Por ejemplo, $\alpha = .05$.
3. Recolectar los datos en una muestra representativa. Vamos a suponer que obtuvimos una media de 2,9 horas y una desviación estándar de 1.2 horas; la muestra incluyó 312 niños.

4. Estimar la desviación estándar de la distribución muestral de la media utilizando la siguiente fórmula:

$$S_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Donde $S_{\bar{x}}$ es la desviación estándar de la distribución muestral de la media, s representa la desviación estándar de la muestra y n es el tamaño de la muestra.

En el ejemplo:

$$S_{\bar{x}} = \frac{1.2}{\sqrt{312}}$$

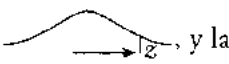
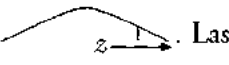
$$S_{\bar{x}} = 0.0679$$

5. Transformar la media de la muestra en una puntuación z , en el contexto de la distribución muestral, con una variación de la fórmula ya conocida para obtener puntuaciones z :

$$z = \frac{X - \bar{X}}{S_{\bar{x}}}$$

donde X es la media de la muestra (recordemos que la distribución muestral es de medias y no de puntuaciones), \bar{X} es la media hipotetizada de la distribución muestral (parámetro poblacional), $S_{\bar{x}}$ es la desviación estándar de la distribución muestral de medias. Así, tenemos:

$$z = \frac{2.9 - 3.0}{0.0679} = -1.47$$

6. En la *tabla de áreas bajo la curva normal* (apéndice 5, tabla 1)⁶, buscar aquella puntuación z que deje a 2.5% por encima de ella, que es 1.96. En la tabla 1 se presenta la distribución de puntuaciones z , sólo la mitad, pues debemos recordar que es una distribución simétrica y se aplica igual para ambos lados de la media. Así se incluye en los textos de estadística social. Se busca 2.5%, porque la tabla sólo abarca la mitad de la distribución y el riesgo que estamos afrontando es de 5% (2.5% del extremo de cada lado). La tabla contiene cuatro columnas: la primera indica puntuaciones z ; la segunda expresa la distancia de la puntuación z a la media; la tercera, el área que está por debajo de esa puntuación desde el comienzo de la distribución , y la cuarta, el área que está por encima de esa puntuación . Las áreas se expresan en proporciones. Lo que buscamos es una puntuación z que deje por encima un área de .0250 o 2.5% (la encontramos en la cuarta

⁶ Material incluido en el CD que acompaña este libro.

columna de la tabla); esta puntuación z es 1.96. Siempre que nuestro nivel de significancia sea .05, tomamos la puntuación z de 1.96.

7. Comparar la media de la muestra transformada a puntuación z con el valor 1.96; si es menor, aceptar la hipótesis; si es mayor, rechazarla. Veamos el ejemplo:

Media de la muestra transformada a z	Nivel de significancia del .05
1.47	±1.96

Decisión: Aceptar la hipótesis a un nivel de significancia de .05 (95% a favor y 5% de riesgo de cometer un error).

Si la media obtenida,
al transformarse en z ,
hubiera sido: 3.25,

7.46 o un valor mayor → Rechazar la hipótesis

Por ejemplo:

Media de la muestra = 2.0
Desviación estándar de la muestra = 0.65

$$\begin{aligned} n &= 700 \\ S\bar{x} &= 0.0246 \\ z &= 40.65 \end{aligned}$$

La media está situada a más de 40 desviaciones estándar de la media; se localiza en la zona crítica (más allá de 1.96 desviaciones estándar). Rechazar la hipótesis.

¿Por qué es importante otro concepto: el intervalo de confianza?

Se ha hablado de la distribución muestral por lo que respecta a la prueba de hipótesis, pero otro procedimiento de la estadística inferencial es construir un *intervalo* donde se localiza un parámetro (Wiersma, 1999). Por ejemplo, en lugar de pretender probar una hipótesis acerca de la media poblacional, puede buscarse un intervalo donde se ubique dicha media. Lo anterior requiere un nivel de confianza, al igual que en la prueba de hipótesis inferenciales. El nivel de confianza es al intervalo de confianza, lo que el nivel de significancia es a la prueba de hipótesis. Es decir, se trata de una probabilidad definida de que un parámetro se va a ubicar en un determinado intervalo. Los niveles de confianza más comunes en la investigación social son 0.95 y 0.99. Su sentido es el de 0.95;

quiere decir que tenemos 95% en favor de que el parámetro se localice en el intervalo estimado, contra 5% de elegir un intervalo equivocado. El nivel de 0.99 señala 99% de probabilidad de seleccionar el intervalo adecuado. Tales niveles de confianza se expresan en unidades de desviación estándar. Una vez más se acude a la distribución muestral, concretamente a la tabla de áreas bajo la curva normal (apéndice 5, tabla 1), y se selecciona la puntuación z correspondiente al nivel de confianza seleccionado. Una vez hecho esto, se aplica la siguiente fórmula:

$$\text{Intervalo de confianza} = \text{estadígrafo} + \left(\begin{array}{l} \text{Puntuación } z \\ \text{que expresa el} \\ \text{nivel de confian-} \\ \text{za elegido} \end{array} \right) \left(\begin{array}{l} \text{Desviación estándar} \\ \text{de la distribu-} \\ \text{ción muestral co-} \\ \text{rrespondiente} \end{array} \right)$$

donde el estadígrafo es la estadística calculada en la muestra, la puntuación z es 1.96 con un nivel de .95 y 2.58 con un nivel de .99, en tanto que el error estándar depende del estadígrafo en cuestión. Veámoslo con el ejemplo de la media en el caso de la exposición diaria a la televisión (en fin de semana) por parte de los niños de Celaya:

Media = 2.9 horas

$s = 1.2$ horas

$S\bar{x} = 0.0679$

(desviación estándar de la distribución muestral de la media).

Nivel de confianza = .95 ($z = 1.96$)

Intervalo de confianza = $2.9 \pm (1.96)(0.0679)$
 $= 2.9 \pm (0.133)$

Intervalo de confianza: la media poblacional está entre 2.767 y 3.033 horas, con 95% de probabilidades de no cometer error.

¿Se pueden cometer errores al realizar estadística inferencial?

Nunca estaremos completamente seguros de nuestra estimación. Trabajamos con altos niveles de confianza o seguridad, pero, aunque el riesgo es mínimo, podría cometerse un error. *Los resultados posibles al probar hipótesis serían:*

1. Aceptar una hipótesis verdadera (decisión *correcta*).
2. Rechazar una hipótesis falsa (decisión *correcta*).

3. Aceptar una hipótesis falsa (conocido como *error del Tipo II o error beta*).
4. Rechazar una hipótesis verdadera (conocido como *error del Tipo I o error alfa*).

Ambos tipos de error son indeseables; puede *reducirse la posibilidad* de que se presenten mediante:

- a) *Muestras representativas probabilísticas.*
- b) *Inspección cuidadosa de los datos.*
- c) *Selección de las pruebas estadísticas apropiadas.*
- d) *Mayor conocimiento de la población.*

ANÁLISIS PARAMÉTRICOS

Hay dos tipos de análisis que pueden realizarse: los *análisis paramétricos* y los *no paramétricos*. Cada tipo posee sus características y presuposiciones que lo sustentan; la elección de qué clase de análisis efectuar depende de estas presuposiciones. De igual forma, cabe destacar que en una misma investigación es posible llevar a cabo análisis paramétricos para algunas hipótesis y variables y análisis no paramétricos para otras.

¿Cuáles son los presupuestos o las presuposiciones de la estadística paramétrica?

Para realizar *análisis paramétricos* debe partirse de los siguientes supuestos:

1. *La distribución poblacional de la variable dependiente es normal:* el universo tiene una distribución normal.
2. *El nivel de medición* de la variable dependiente es *por intervalos o razón*.
3. *Cuando dos o más poblaciones son estudiadas, tienen una varianza homogénea:* las poblaciones en cuestión poseen una dispersión similar en sus distribuciones (Wiersma, 1999).

¿Cuáles son los métodos o las pruebas estadísticas paramétricas más utilizados?

Las *pruebas estadísticas paramétricas más utilizadas* son:

- Coeficiente de correlación de Pearson y regresión lineal
- Prueba *t*
- Prueba de contraste de la diferencia de proporciones

- Análisis de varianza unidireccional (ANOVA *Oneway*)
- Análisis de varianza factorial (ANOVA)
- Análisis de covarianza (ANCOVA)

¿Qué es el coeficiente de correlación de Pearson?

Definición: Es una prueba estadística para analizar la relación entre dos variables medidas en un nivel por intervalos o de razón.

Se simboliza: r

Hipótesis a probar: Correlacional, del tipo de “A mayor X, mayor Y”, “A mayor X, menor Y”, “Altos valores en X están asociados con altos valores en Y”, “Altos valores en X se asocian con bajos valores de Y”.

Variables: Dos. La prueba en sí no considera a una como independiente y a otra como dependiente, ya que no evalúa la causalidad. La noción de causa-efecto (independiente-dependiente) es posible establecerla teóricamente, pero la prueba no considera dicha causalidad.

El coeficiente de correlación de Pearson se calcula a partir de las puntuaciones obtenidas en una muestra en dos variables. Se relacionan las puntuaciones obtenidas de una variable con las puntuaciones obtenidas de otra variable, en los mismos sujetos.

Nivel de medición de las variables: Intervalos o razón.

Interpretación: El coeficiente r de Pearson puede variar de -1.00 a $+1.00$ donde:

-1.00 = correlación negativa perfecta. (“A mayor X, menor Y”, de manera proporcional. Es decir, cada vez que X aumenta una unidad, Y disminuye siempre una cantidad constante.) Esto también se aplica “a menor X, mayor Y”.

-0.90 = Correlación negativa muy fuerte.

-0.75 = Correlación negativa considerable.

-0.50 = Correlación negativa media.

-0.10 = Correlación negativa débil.

0.00 = No existe correlación alguna entre las variables.

$+0.10$ = Correlación positiva débil.

$+0.50$ = Correlación positiva media.

$+0.75$ = Correlación positiva considerable.

+0.90 = Correlación positiva muy fuerte.

+1.00 = *Correlación positiva perfecta.*

(“A mayor X, mayor Y” o “A menor X, menor Y”, de manera proporcional. Cada vez que X aumenta, Y aumenta siempre una cantidad constante.)

El *signo* indica la dirección de la correlación (positiva o negativa); y el *valor numérico*, la magnitud de la correlación.

Los principales programas de análisis estadístico computacionales reportan si el coeficiente es o no significativo de la siguiente manera:

$s = 0.001$ significancia
0.7831 valor del coeficiente

Si s es menor del valor .05, se dice que el coeficiente es *significativo* al nivel de .05 (95% de confianza en que la correlación sea verdadera y 5% de probabilidad de error). Si s es menor a .01, el coeficiente es *significativo* al nivel de .01 (99% de confianza de que la correlación sea verdadera y 1% de probabilidad de error).

Consideraciones:

Cuando el coeficiente r de Pearson se eleva al cuadrado (r^2), el resultado indica la *varianza de factores comunes*. Esto es, el porcentaje de la variación de una variable debido a la variación de la otra variable y viceversa.

Por ejemplo: La correlación entre “productividad” y “asistencia al trabajo” es de 0.80.

$$r = 0.80$$
$$r^2 = 0.64$$

“La productividad” constituye a, o explica, 64% de la variación de “la asistencia al trabajo”. “La asistencia al trabajo” explica 64% de “la productividad”.

En los artículos de revistas científicas o textos, la significancia se suele indicar así:

$$0.48^*$$
$$p < .05$$

Quiere decir que el coeficiente es significativo al nivel de .05. La probabilidad de error es menor de 5%. Si $p < .01$,

Coefficiente de correlación de Pearson:
prueba estadística para analizar la relación entre dos variables medidas en un nivel por intervalos o de razón.

el coeficiente es significativo al nivel de .01. También suele señalarse, con asteriscos, de la siguiente manera:

	X
XY	.11
AZ	.62**
BX	.47*
CW	.009
* $p < .05$ ** $p < .01$	

X, Y, A, Z, B, C y W son variables. Cada coeficiente es una correlación.

Hi:	"A mayor motivación intrínseca, mayor puntualidad."
Resultado:	$r = .721$ $s = 0.0001$
Interpretación:	Se acepta la hipótesis de investigación al nivel de .01. La correlación entre la motivación intrínseca y la productividad es considerable.
Hi:	"A mayor ingreso, mayor motivación intrínseca."
Resultado:	$r = .214$ $s = 0.081$
Interpretación:	Se acepta la hipótesis nula. El coeficiente no es significativo: 0.081 es mayor que 0.05; recordemos que 05 es el nivel mínimo para aceptar la hipótesis.
Nota precautoria:	Recuérdese lo referente a correlaciones espúreas que se comentó en el capítulo de tipos de estudio.

¿Qué es la regresión lineal?

- Definición:** Es un modelo matemático para estimar el efecto de una variable sobre otra. Está asociado con el coeficiente r de Pearson.
- Hipótesis:** Correlacionales y causales.
- Variables:** Dos. Una se considera como independiente y otra como dependiente. Pero, para poder hacerlo, debe tenerse un sólido sustento teórico.

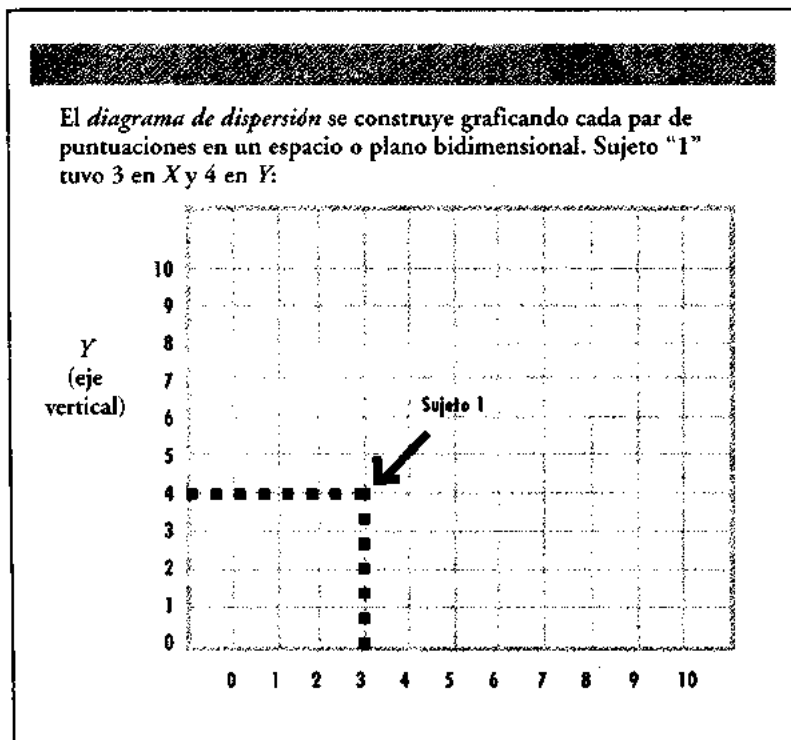
Nivel de medición de las variables:

Intervalos o razón.

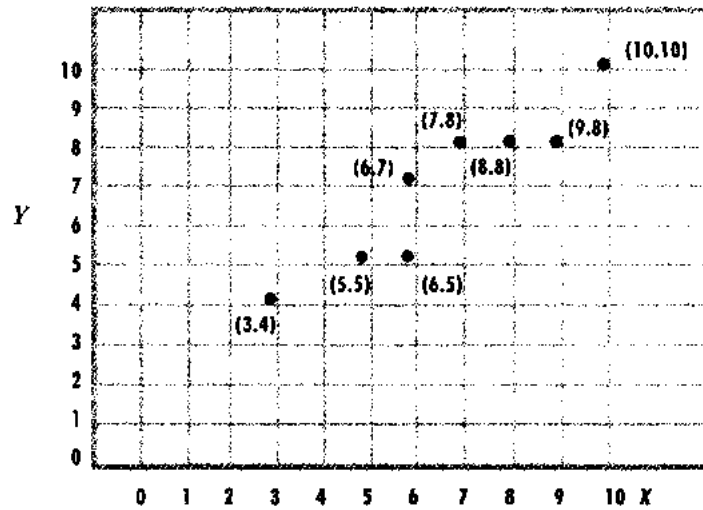
Procedimiento e interpretación:

La regresión lineal se determina con base en el *diagrama de dispersión*. Éste consiste en una gráfica donde se relacionan las puntuaciones de una muestra en dos variables. Veámoslo con un ejemplo sencillo de 8 casos. Una variable es la calificación en Filosofía y la otra variable es la calificación en Estadística; ambas medidas, hipotéticamente, de 0 a 10.

<i>Sujetos</i>	<i>Puntuaciones filosofía (X)</i>	<i>Estadística (Y)</i>
1	3	4
2	8	8
3	9	8
4	6	5
5	10	10
6	7	8
7	6	7
8	5	5



Así se grafican todos los pares:
 X (eje horizontal)



Los *diagramas de dispersión* son una manera de visualizar gráficamente una correlación. Por ejemplo:



Correlación positiva muy fuerte: la tendencia es ascendente, altas puntuaciones en X y altas puntuaciones en Y .

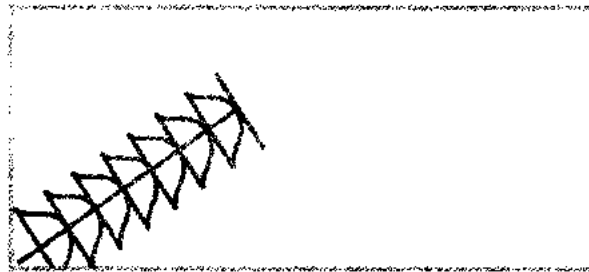


Correlación negativa considerable



Ausencia de correlación

Así, cada punto representa un caso y un resultado de la intersección de las puntuaciones en ambas variables. El diagrama de dispersión puede ser resumido a una línea (producto de las medias de las puntuaciones).



Conociendo la línea y la tendencia, podemos predecir los valores de una variable conociendo los de la otra variable.

Ésta línea se expresa mediante la *ecuación de regresión lineal*:

$$Y = a + bX$$

donde *Y* es un valor de la variable dependiente que se desea predecir, *a* es la ordenada en el origen y *b* la pendiente o inclinación.

Los programas y paquetes de análisis estadístico computacionales, que incluyen la *regresión lineal*, proporcionan los datos de *a* y *b*.

a o *intercept* y *b* o *slope*.

Para predecir un valor de *Y*, se sustituyen los valores correspondientes en la ecuación.

$$a \text{ (intercept)} = 1.2$$

$$b \text{ (slope)} = 0.8$$

Entonces podemos hacer la predicción: ¿a un valor de 7 en Filosofía qué valor le corresponde en Estadística?

$$Y = \frac{1.2}{\boxed{a}} + \frac{(0.8)}{\boxed{b}} \cdot \frac{(7)}{\boxed{X}}$$

$$Y = 6.8$$

Predecimos que a un valor de 7 en X le corresponderá un valor de 6.8 en Y .

Consideraciones: La *regresión lineal* es útil con relaciones lineales, no con *relaciones curvilineales* de los tipos que se muestran en la figura 10.11.

Hi: “La autonomía laboral es una variable para predecir la motivación intrínseca en el trabajo. Ambas variables están relacionadas.”
Las dos variables fueron medidas en una escala por intervalos de 1 a 5.

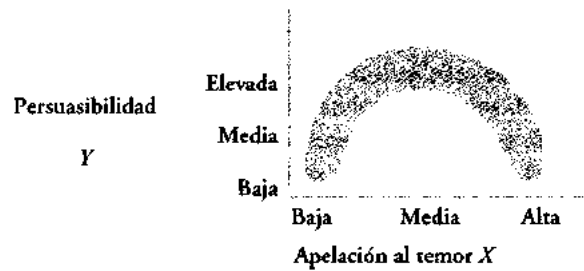
Resultado: $a \text{ (intercept)} = 0.42$
 $b \text{ (slope)} = 0.65$

Interpretación: Cuando X (autonomía) es 1, la predicción estimada de Y es 1.07; cuando X es 2, la predicción estimada de Y es 1.72; cuando X es 3, Y será 2.37; cuando X es 4, Y será 3.02; y cuando X es 5, Y será 3.67.

$$Y = a + bX$$
$$1.07 = 0.42 + 0.65 (1)$$
$$1.72 = 0.42 + 0.65 (2)$$
$$2.37 = 0.42 + 0.65 (3)$$
$$3.02 = 0.42 + 0.65 (4)$$
$$3.67 = 0.42 + 0.65 (5)$$

De la
regresión
lineal

Las relaciones curvilineales son aquellas en las cuales la tendencia varía: primero es ascendente y luego descendente, o viceversa. Se ha demostrado que una estrategia persuasiva con niveles altos de apelación al temor, por ejemplo, un comercial televisivo muy dramático, provoca una baja persuasibilidad, lo mismo que una estrategia persuasiva con niveles muy bajos de apelación al temor. La estrategia persuasiva más adecuada es la que utiliza niveles medios de apelación al temor. Esta relación es curvilineal; se representaría así:



Otras gráficas de relaciones curvilineales serían:

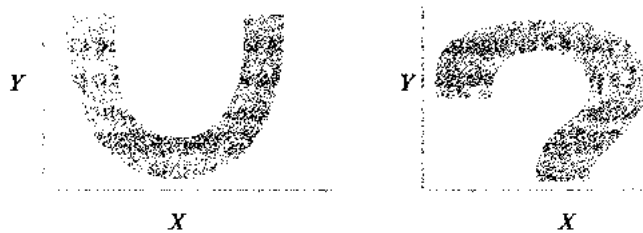


Figura 10.11

¿Qué es la prueba t?

Definición: Es una prueba estadística para evaluar si dos grupos difieren entre sí de manera significativa respecto a sus medias.

Se simboliza: *t*

Hipótesis: De diferencia entre dos grupos. La hipótesis de investigación propone que los grupos difieren significativamente entre sí y la hipótesis nula propone que los grupos no difieren significativamente.

Variable: La comparación se realiza sobre una variable. Si hay diferentes variables, se efectuarán varias pruebas t (una por cada variable), aunque la razón que motiva la creación de los grupos puede ser una variable independiente. Por ejemplo, un experimento con dos grupos, donde a uno se le aplica el estímulo experimental y el otro grupo es de control.

Nivel de medición de la variable de comparación:

Intervalos o razón.

Interpretación:

El valor t se obtiene en muestras grandes mediante la fórmula:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

donde \bar{X}_1 es la media de un grupo, \bar{X}_2 es la media del otro grupo, s_1^2 es la desviación estándar del primer grupo elevada al cuadrado, n_1 es el tamaño del primer grupo, s_2^2 es la desviación estándar del segundo grupo elevada al cuadrado y n_2 es el tamaño del segundo grupo. En realidad, el denominador es el *error estándar de la distribución muestral de la diferencia entre medias*.

Para saber si el valor t es significativo, se aplica la fórmula y se calculan los *grados de libertad*. La prueba t se basa en una distribución muestral o poblacional de diferencia de medias conocida como la distribución t de Student. Esta distribución se identifica por los *grados de libertad*, los cuales *constituyen el número de maneras en que los datos pueden variar libremente*. Son determinantes, ya que nos indican qué valor debemos esperar de t , dependiendo del tamaño de los grupos que se comparan. *Cuanto mayor número de grados de libertad se tengan, la distribución t de Student se acercará más a ser una distribución normal* y usualmente, si los grados de libertad exceden los 120, la distribución normal se utiliza como una aproximación adecuada de la distribución t de Student (Wiersma, 1999).

Los *grados de libertad* se calculan así:

$$gl = (n_1 + n_2) - 2$$

n_1 y n_2 son el tamaño de los grupos que se comparan.

Una vez calculados el valor t y los grados de libertad, se elige *el nivel de significancia* y se compara el valor obtenido contra el valor que le correspondería en la tabla 2 del apéndice 5 (*tabla de la distribución t de Student*). Si el valor calculado es igual o mayor al que aparece en la tabla, se acepta la hipótesis de investigación. Pero si es menor, se acepta la hipótesis nula.

En la tabla se busca el valor con el cual vamos a comparar el que hemos calculado, basándonos en el nivel de confianza elegido (0.05 o 0.01) y los grados de libertad. La tabla contiene los niveles de confianza como columnas y los grados de libertad como renglones. Los niveles de confianza adquieren el significado del que se ha hablado (el .05 significa 95% de que los grupos en realidad difieran significativamente entre sí y 5% de posibilidad de error).

Cuanto mayor sea el valor t calculado respecto al valor de la tabla y menor sea la posibilidad de error, mayor será la certeza en los resultados.

Cuando el valor t se calcula utilizando un paquete estadístico computacional, la significancia se proporciona como parte de los resultados y ésta debe ser menor a .05 o .01, dependiendo del nivel de confianza seleccionado.

Consideraciones: La prueba t se utiliza para comparar los resultados de una preprueba con los resultados de una posprueba en un contexto experimental. Se comparan las medias y las varianzas del grupo en dos momentos diferentes:

$(\bar{X}_1) \times (\bar{X}_2)$ o bien, para comparar las prepruebas o pospruebas de dos grupos que participan en un experimento:

X (\bar{X}_1)

(\bar{X}_2) t O son las pospruebas

Hi: "Los varones le atribuyen mayor importancia al atractivo físico en sus relaciones heterosexuales que las mujeres."

Ho: "Los varones no le atribuyen más importancia al atractivo físico en sus relaciones heterosexuales que las mujeres."

La variable atractivo físico fue medida a través de una prueba estandarizada y el nivel de medición es por intervalos. La escala varía de 0 a 18.

La hipótesis se somete a prueba con los estudiantes de clase media de dos universidades de la ciudad de Monterrey.

n_1 (hombres) = 128

n_2 (mujeres) = 119

Resultados:

\bar{X}_1 (hombres) = 15

\bar{X}_2 (mujeres) = 12

$$\begin{aligned}
 s_1 \text{ (hombres)} &= 4 \\
 s_2 \text{ (mujeres)} &= 3 \\
 t &= \frac{15 - 12}{\sqrt{\frac{(4)^2}{128} + \frac{(3)^2}{119}}} \\
 t &= 6.698 \\
 Gl &= (128 + 119) - 2 \\
 Gl &= 245
 \end{aligned}$$

Al acudir a la tabla de la distribución t de Student (apéndice 5, tabla 2, incluido en el CD), buscamos los grados de libertad correspondientes y elegimos en la columna de Gl , el renglón α , que se selecciona siempre que se tienen más de 200 grados de libertad. La tabla contiene los siguientes valores:

Gl	.05	.01
α (mayor de 200)	1.645	2.326

Nuestro valor calculado de t es 6.698 y resulta superior al valor de la tabla en un nivel de confianza de .05 ($6.698 > 1.645$). Entonces, la conclusión es que aceptamos la hipótesis de investigación y rechazamos la nula. Incluso, el valor t calculado es superior en un nivel de confianza del .01 ($6.698 > 2.326$).

Comentario: Efectivamente, en el contexto de la investigación, los varones le atribuyen más importancia al atractivo físico en sus relaciones heterosexuales que las mujeres.

Si tuviéramos 60 grados de libertad y un valor t igual a 1.27, al comparar este valor con los de la tabla obtendríamos:

Gl	.05	.01
60	1.6707	2.390

El valor t calculado es menor a los valores de la tabla. *Se rechaza la hipótesis de investigación y se acepta la hipótesis nula.*

¿Qué es la prueba de diferencia de proporciones?

Definición: Es una prueba estadística para analizar si dos proporciones difieren significativamente entre sí.

Hipótesis: De diferencia de proporciones en dos grupos.

Variable: La comparación se realiza sobre una variable. Si hay varias, se efectuará una prueba de diferencia de proporciones por variable.

Nivel de medición de la variable de comparación:

Intervalos o razón, expresados en proporciones o porcentajes.

Procedimiento e interpretación:

Se obtienen las proporciones de los grupos. Se aplica la siguiente fórmula:

$$z = \frac{P_1 - P_2}{\sqrt{\frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2}}}$$

$\left\{ \begin{array}{l} q_1 = 1 - P_1 \\ q_2 = 1 - P_2 \end{array} \right.$

La puntuación z resultante se compara con la puntuación z de la distribución de puntuaciones z (normal) que corresponda al nivel de confianza elegido. El valor calculado de z (resultante de aplicar la fórmula) debe ser igual o mayor que el valor de la tabla de áreas bajo la curva normal correspondiente (tabla 1, apéndice 5). Si es igual o mayor, se acepta la hipótesis de investigación. Si es menor, se rechaza.

Hi: "El porcentaje de liberales en la ciudad de Arualm es mayor que en Linderbuck."

% de liberales en Arualm

% de liberales en Linderbuck

55%

48%

$n_1 = 410$

$n_2 = 301$

Los porcentajes se transforman en proporciones y se calculan q_1 y q_2 :

<i>Aruahn</i>	<i>Linderbuck</i>
$P_1 = 0.55$	$P_2 = 0.48$
$N_1 = 410$	$N_2 = 301$
$q_1 = 1 - .55 = 0.45$	$q_2 = 1 - .48 = 0.52$
$\alpha = .05 = 1.96 z$ (puntuación z que, como se ha explicado anteriormente, corresponde al nivel alfa del .05)	
$z = \frac{0.55 - 0.48}{\sqrt{\frac{(.55)(.45)}{410} + \frac{(.48)(.52)}{301}}} = 1.849$	
<p>Como la z calculada es menor a 1.96 (nivel alfa expresado en una puntuación z), aceptamos la hipótesis nula y rechazamos la de investigación.</p>	

¿Qué es el análisis de varianza unidireccional? (one way)

Definición:

Es una prueba estadística para analizar si más de dos grupos difieren significativamente entre sí en cuanto a sus medias y varianzas. *La prueba t* se utiliza para *dos grupos* y el *análisis de varianza unidireccional* se usa para *tres, cuatro o más grupos*. Aunque con dos grupos se puede utilizar el análisis de varianza unidireccional.

Hipótesis:

De diferencia entre más de dos grupos. La hipótesis de investigación propone que los grupos difieren significativamente entre sí y la hipótesis nula propone que los grupos no difieren significativamente.

Variables:

Una variable independiente y una variable dependiente.

Nivel de medición de las variables:

La variable independiente es categórica y la dependiente es por intervalos o razón.

El hecho de que la variable independiente sea categórica significa que es posible formar grupos diferentes. Puede ser una variable nominal, ordinal, por intervalos o de razón (pero en estos últimos dos casos la variable debe reducirse a categorías).

Por ejemplo:

— Religión.

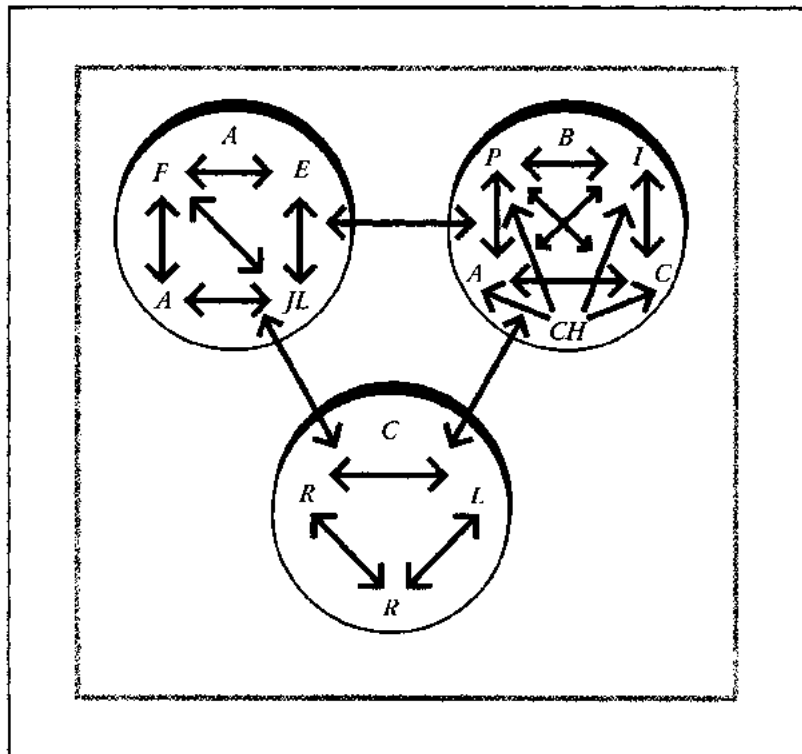
— Nivel socioeconómico (muy alto, alto, medio, bajo y muy bajo).

— Antigüedad en la empresa (de 0 a 1 año, más de un año a cinco años, más de cinco años a diez, más de diez años a 20 y más de 20 años).

Interpretación:

El *análisis de varianza unidireccional* produce un valor conocido como *F* o *razón F*, que se basa en una distribución muestral, conocida como *distribución F*, que es otro miembro de la familia de distribuciones muestrales. La *razón F* compara las variaciones en las puntuaciones debidas a dos diferentes fuentes: variaciones entre los grupos que se comparan y variaciones dentro de los grupos.

Si los grupos difieren realmente entre sí, sus puntuaciones variarán más de lo que puedan variar las puntuaciones entre los integrantes de un mismo grupo. Veámoslo con un ejemplo cotidiano. Si tenemos tres familias *A*, *B* y *C*. La familia *A* está integrada por Felipe, Angélica, Elena y José Luis. La familia *B* está compuesta por Chester, Pilar, Íñigo, Alonso y Carlos. Y la familia *C* está integra-



da por Rodrigo, Laura y Roberto. ¿Qué esperamos? Pues que los integrantes de una familia se parezcan más entre sí que a los miembros de otra familia. Esto se graficaría como en la figura que viene en la página 545.

Es decir, esperamos *homogeneidad* intrafamilias y *heterogeneidad* interfamilias.

¿Qué sucedería si los miembros de las familias se parecieran más a los integrantes de las otras familias que a los de la suya propia? Quiere decir que no hay diferencia entre los grupos (en el ejemplo, familias).

Esta misma lógica se aplica a la *razón F*, la cual nos indica si las diferencias entre los grupos son mayores que las diferencias intragrupos (dentro de éstos). Estas diferencias se miden en términos de *varianza*. La *varianza* es una medida de dispersión o variabilidad alrededor de la media y se calcula en términos de desviaciones elevadas al cuadrado. Recuérdese que la *desviación estándar* es un promedio de desviaciones respecto a la media $(X - \bar{X})$ y la *varianza* es un promedio de

desviaciones respecto a la media elevadas al cuadrado $(X - \bar{X})^2$. La *varianza* por eso se simboliza como S^2 y su fórmula es $\sum(X - \bar{X})^2/n$. En consecuencia, la *razón F*, que es una razón de varianzas; se expresa así:

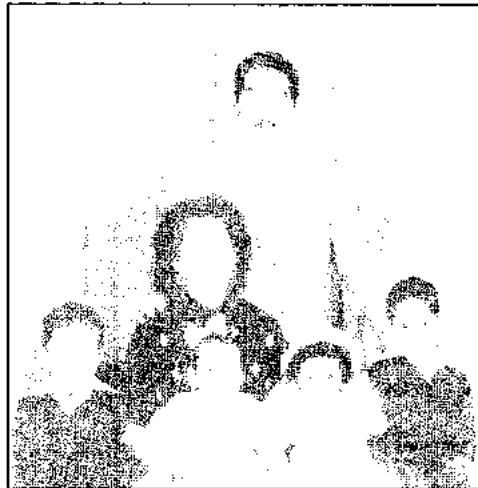
$$F = \frac{\text{Media cuadrática entre los grupos}}{\text{Media cuadrática dentro de los grupos}}$$

donde *media cuadrática* implica un promedio de varianzas elevadas al cuadrado.

La *media cuadrática entre los grupos* se obtiene calculando la media de las puntuaciones de todos los grupos (media total), después se obtiene la desviación de la media de cada grupo respecto a la media total y se eleva al cuadrado cada una de estas desviaciones, y luego se suman. Finalmente se sopesa el número de individuos en cada grupo y la *media cuadrática* se obtiene con base en los *grados de libertad intergrupales* (no se calcula

Análisis de varianzas: prueba estadística para analizar si más de dos grupos difieren entre sí de manera significativa en sus medias y varianzas.

En un análisis de integrantes de un mismo grupo, se espera homogeneidad: los miembros de una misma familia se parecen más entre sí.



con base en el número de puntuaciones). La *media cuadrática dentro de los grupos* se calcula obteniendo primero la desviación de cada puntuación respecto a la media de su grupo; posteriormente esta fuente de variación se suma y combina para obtener una media de la *varianza intragrupal* para todas las observaciones, tomando en cuenta los grados de libertad totales (Wright, 1979).

Las fórmulas de la media cuadrática son:

$$\text{Media cuadrática entre grupos} = \frac{\text{Suma de cuadrados entre grupos}}{\text{Grados de libertad entre grupos}}$$

Los grados de libertad entre grupos = $K - 1$ (donde K es el número de grupos).

$$\text{Media cuadrática dentro de los grupos} = \frac{\text{Suma de cuadrados intragrupos}}{\text{Grados de libertad intragrupos}}$$

Los *grados de libertad intragrupos* = $n - K$ (n es el tamaño de la muestra, la suma de los individuos de todos los grupos, y K recordemos que es el número de grupos).

Para el procedimiento de cálculo manual de la *razón F* se recomiendan Levin (1979) o cualquier texto de estadística social.

Pues bien, cuando *F* resulta significativa, quiere decir que los grupos difieren significativamente entre sí. Es decir, se acepta la hipótesis de investigación y se rechaza la hipótesis nula.

Cuando se efectúa el *análisis de varianza* por medio de un programa computacional o se utiliza un paquete estadístico, se genera una tabla de resultados con los elementos de los cuadros 10.8 y 10.9.

El valor α (*alfa*) o probabilidad a elegir es una vez más .05 o .01. Si es menor que .05 es significativo a este nivel, y si es menor que .01 es significativo también a este nivel. Cuando el programa o paquete estadístico no incluye la significancia se acude a la *tabla 3 del apéndice 5 (tabla de la distribución F)*. Esta tabla contiene una lista de razones significativas (razones *F*) que debemos obtener para poder aceptar la hipótesis de investigación en los niveles de confianza de .05 y .01. Al igual que en caso de la razón *t*, el valor exacto de *F* que debemos obtener depende de sus grados de libertad asociados. Por lo tanto, la utilización de la tabla se inicia buscando los dos valores *G*_l, los *grados de libertad entre los grupos* y los *grados de libertad intragrupos*. Los grados de libertad entre grupos se indican en la parte superior de la página, mientras que los grados de libertad intragrupos se han colocado al lado izquierdo de la tabla. El cuerpo de la tabla de la distribución *F* presenta razones *F* significativas a los niveles de confianza de .05 y .01.

$$\begin{aligned} \text{Si } F &= 1.12 \\ G_l \text{ entre} &= 2 \\ G_l \text{ intra} &= 60 \end{aligned}$$

Elementos para interpretar un análisis de varianza unidireccional realizado con un paquete estadístico

Cuadro 10.8

FUENTE DE VARIACIÓN (SOURCE)	SUMAS DE CUADRADOS (SUMS OF SQUARES)	GRADOS DE LIBERTAD (DEGREES OF FREEDOM)	MEDIAS CUADRÁTICAS (MEAN SQUARES)	RAZÓN F (F RATIO)	SIGNIFICANCIA DE F (F PROB.)
Entre grupos (between groups)	SS entre	Gf entre	SS entre/Gf entre	$\frac{M C \text{ entre}}{M C \text{ intra}}$	α
Intragrupos (within groups)	SS intra	Gf intra	SS intra/Gf intra		
Total	SS entre + SS intra	Gf entre + Gf intra			

Este *valor F* se compara con el valor que aparece en la tabla de la distribución *F* que es 3.15 y como *el valor F* calculado es menor al de dicha tabla, rechazaríamos la hipótesis de investigación y aceptaríamos la hipótesis nula. Para que *el valor F* calculado sea significativo debe ser igual o mayor al de la tabla.

Hi: "Los niños que se expongan a contenidos de elevada violencia televisiva exhibirán una conducta más agresiva en sus juegos, respecto a los niños que se expongan a contenidos de mediana o baja violencia televisada."

Ho: "Los niños que se expongan a contenidos de elevada violencia televisiva no exhibirán una conducta más agresiva en sus juegos, respecto a los niños que se expongan a contenidos de mediana o baja violencia televisada."

La variable independiente es el grado de exposición a la violencia televisada y la variable dependiente es la agresividad exhibida en los juegos, medida por el número de conductas agresivas observadas (intervalos).

Para probar la hipótesis se diseña un experimento con tres grupos:

- | | | |
|--------------------------------|---|--------------------------------|
| G_1, X_1 (elevada violencia) | o | } Número de actos
agresivos |
| G_2, X_2 (mediana violencia) | o | |
| G_3, X_3 (baja violencia) | o | |
| G_4 — (conducta prosocial) | o | |

En cada grupo hay 25 niños.

La razón *F* resultó significativa: se acepta la hipótesis de investigación. La diferencia entre las medias de los grupos es significativa, el contenido altamente violento tiene un efecto sobre la conducta agresiva de los niños en sus juegos. El estímulo experimental tuvo un efecto. Esto se corrobora comparando las medias de las pospruebas de los cuatro grupos,

porque el análisis de varianza unidireccional sólo nos señala si la diferencia entre las medias y las distribuciones de los grupos es o no significativa, pero no nos indica en favor de qué grupos lo es; esto es posible hacerlo comparando las medias y las distribuciones de los grupos. Y si adicionalmente queremos comparar cada par de medias (X_1 con X_2 , X_1 con X_3 , X_2 con X_3 , etcétera) y determinar con exactitud dónde están las diferencias significativas, podemos aplicar un contraste *a posteriori*, calculando una prueba *t* para cada par de medias; o bien, a través de algunas estadísticas que suelen ser parte de los análisis efectuados mediante paquetes estadísticos computacionales. Tales estadísticas se incluyen en el cuadro 10.10.

FUENTE DE VARIACIÓN	SUMAS DE CUADRADOS	GRADOS DE LIBERTAD	MEDIAS CUADRÁTICAS	RAZÓN F	SIGNIFICANCIA DE F	Ejemplo hipotético de resultados del cuadro 10.8
Entre grupos	150.18	3	50.06			
Intragrupos	857.64	96	8.93	5.6	0.001	
Total	1 007.82	99				

Cuadro 10.9

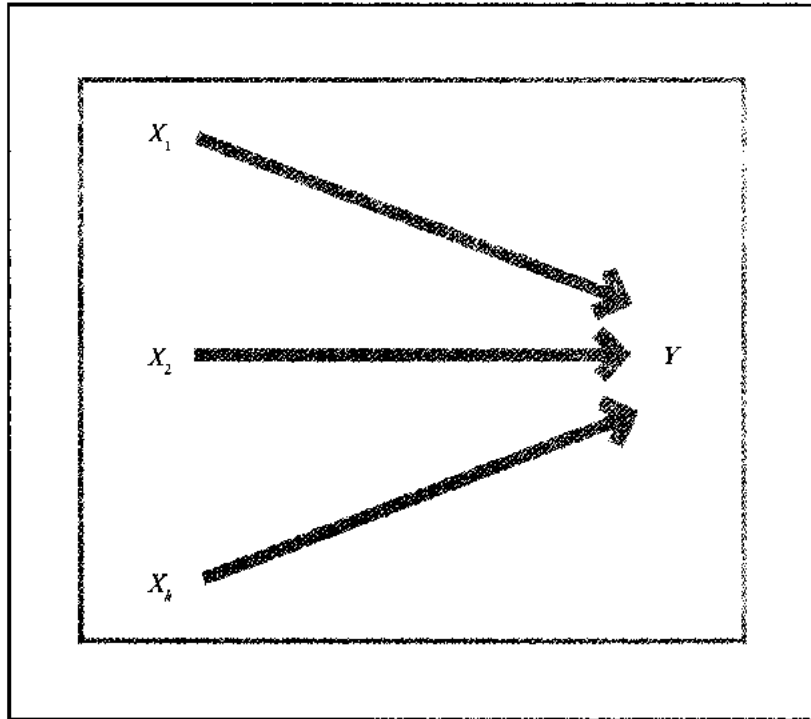
NOMBRE	ABREVIATURA	NIVEL DE CONFIANZA EN QUE SE UTILIZA	Principales estadísticas para contrastes <i>a posteriori</i> en el análisis de varianza unidireccional, realizado con el paquete estadístico más conocido en América Latina: SPSS. ⁷
– Diferencia significativa mínima	LSD	Cualquier nivel	
– La prueba de Duncan de múltiples rangos	DUNCAN	.10, .05, .01	
– Student-Newman-Kenik	SNK	.05	
– Procedimiento alternativo de Tukey	TUKEYB	.05	
– Diferencia significativa honesta	TUKEY (o DSH en algunos paquetes en español)	.05	
– LSD modificado	LSDMOD	Cualquier nivel	
– Procedimiento de Scheffe	SCHEFFE	Cualquier nivel	

Cuadro 10.10

⁷ Paquete estadístico para ciencias sociales, el cual se comentará cuando se hable de paquetes estadísticos.

¿Qué es el análisis factorial de varianza? (ANOVA) (análisis de varianza de k direcciones)

Definición: Es una prueba estadística para evaluar el efecto de dos o más variables independientes sobre una variable dependiente. Responde a esquemas del tipo:



Constituye una extensión del análisis de varianza unidireccional, solamente que incluye más de una variable independiente. Evalúa los efectos por separado de cada variable independiente y los efectos conjuntos de dos o más variables independientes.

Variables: Dos o más variables independientes y una dependiente.

Nivel de medición de las variables: La variable dependiente (criterio) debe estar medida en un nivel por intervalos o razón, y las variables independientes (factores) pueden estar en cualquier nivel de medición, pero expresadas de manera categórica.

H₁: “La similitud en valores, la atracción física y el grado de retroalimentación positiva son factores que inciden en la satisfacción sobre la relación en parejas de novios cuyas edades oscilan entre los 24 y los 32 años.”

El ANOVA efectuado mediante un paquete estadístico computacional produce los siguientes elementos básicos:

- *Fuente de la variación* (source of variation). Es el factor que origina variación en la variable dependiente. Si una fuente no origina variación en la dependiente, no tiene efectos.
- *Efectos principales* (main effects). Es el efecto de cada variable independiente por separado; no está contaminado del efecto de otras variables independientes ni de error. Suele proporcionarse la suma de todos los efectos principales.
- *Interacciones de dos direcciones* (2-way interactions). Representa el efecto conjunto de dos variables independientes, aislado de los demás posibles efectos de las variables independientes (individuales o en conjuntos). Suele proporcionarse la suma de los efectos de todas estas interacciones.
- *Interacciones de tres direcciones* (3-way interactions). Constituye el efecto conjunto de tres variables independientes, aislado de otros efectos. Suele proporcionarse la suma de los efectos de todas estas interacciones.
- Puede haber efecto de *K*-direcciones, dependiendo del número de variables independientes.

En nuestro ejemplo, tenemos los resultados que se muestran en el cuadro 10.11.

Como podemos ver en el cuadro 10.11, la similitud, la atracción y la retroalimentación tienen un efecto significativo sobre la satisfacción en la relación. Respecto a los efectos de dos variables independientes conjuntas, sólo la similitud y la atracción tienen un efecto, y hay un efecto conjunto de las tres variables independientes. La hipótesis de investigación se acepta y la nula se rechaza. Asimismo, se recuerda al lector que en el capítulo sobre diseños experimentales (en el apartado sobre diseños factoriales) se explica la noción de interacción entre variables independientes. Y cabe agregar que el ANOVA es un método estadístico propio para los diseños experimentales factoriales.

¿Qué es el análisis de covarianza?

Definición: Es una prueba estadística que analiza la relación entre una variable dependiente y dos o más independientes, eliminando y controlando el efecto de al menos una de estas independientes.

Ejemplo de resultados en el ANOVA

Cuadro 10.11

VARIABLE DEPENDIENTE: SATISFACCIÓN EN LA RELACIÓN					
FUENTE DE VARIACIÓN (SOURCE OF VARIATION)	SUMA DE CUADRADOS (SUM OF SQUARES)	GRADOS DE LIBERTAD (DEGREES OF FREEDOM)	MEDIAS CUADRÁTICAS (MEAN SQUARES)	RAZÓN	SIGNIFICANCIA DE F
- Efectos principales (main effects)					
SIMILITUD	-	-	-	22.51	0.001**
ATRACCIÓN	-	-	-	31.18	0.001**
RETROALIMENTACIÓN	-	-	-	21.02	0.001**
	-	-	-	11.84	0.004**
- Interacción de dos direcciones (2-way interactions)					
SIMILITUD	-	-	-	7.65	0.010*
ATRACCIÓN	-	-	-	4.32	0.040*
SIMILITUD	-	-	-	2.18	0.110
RETROALIMENTACIÓN	-	-	-	1.56	0.190
ATRACCIÓN	-	-	-		
RETROALIMENTACIÓN	-	-	-		
- Interacción de tres direcciones (3-way interaction)					
SIMILITUD	-	-	-	8.01	0.020*
ATRACCIÓN	-	-	-	8.01	0.020*
RETROALIMENTACIÓN	-	-	-		
- Residual	-	-	-		
- Total	-	-	-		

NOTA: A los estudiantes que se inician en el ANOVA normalmente les interesa saber si las razones *F* resultaron o no significativas; por tanto, sólo se incluyen estos valores. Y para quien se inicia en este análisis los autores recomiendan concentrarse en ello y evitar confusiones. Desde luego, el investigador experimentado acostumbra estudiar todos los valores.

** - Razón *F* significativa al nivel del .01 ($p < .01$)

* - Razón *F* significativa al nivel del .05 ($p < .05$)

Perspectivas

o usos:

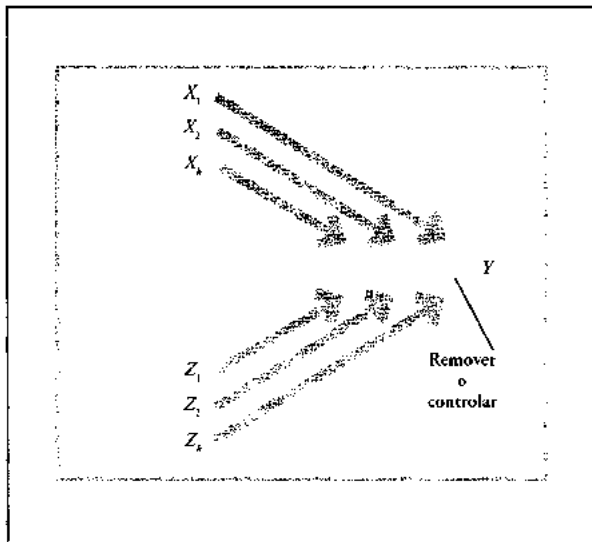
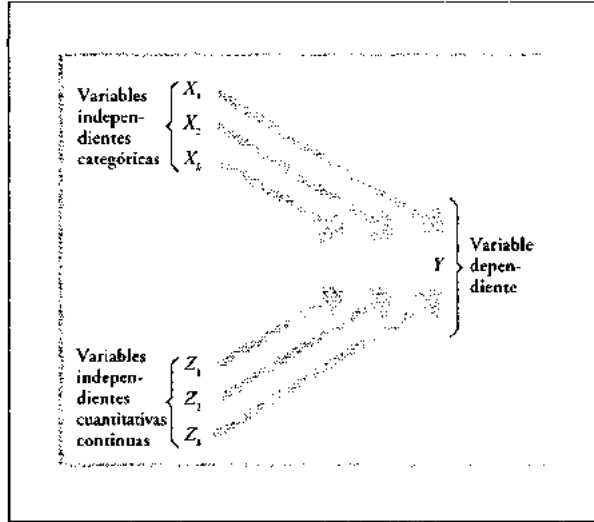
Wildt y Ahtola (1978, pp. 8-9) destacan tres perspectivas para el análisis de covarianza:

A. Perspectiva experimental. Se aplica a aquellas situaciones en que el interés del investigador se centra en las di-

ferencias observadas en la variable dependiente, a través de las categorías de la variable independiente (o variables independientes). Pero el experimentador asume que hay otras variables independientes cuantitativas que contaminan la relación y cuya influencia debe ser controlada. Es decir, se tiene el esquema de la derecha.

Y el investigador únicamente se interesa por conocer la relación entre las variables independientes categóricas y la variable dependiente. Deseando remover y controlar el efecto de las variables independientes cuantitativas no categóricas. Es decir, desea tener un esquema como el de abajo:

El objetivo es "purificar la relación entre las independientes categóricas y



la dependiente, controlando el efecto de las independientes no categóricas o continuas”.

Ejemplos de variables independientes categóricas serían: sexo (masculino, femenino), inteligencia (alta, media, baja), ingreso (menos de 1 salario mínimo, 2 a 4 salarios mínimos, 5 a 10 salarios mínimos, 11 o más salarios mínimos). Los niveles de medición nominal y ordinal son categóricos en sí mismos, y los niveles de intervalos y razón deben transformarse en categorías más discretas. Estos últimos son en sí: cuantitativos, continuos y de categorías múltiples continuas. Por ejemplo, el ingreso en su estado natural varía de la categoría 0 hasta la categoría (K)^k, ya que puede haber millones de categorías.

Variable categórica — unas cuantas categorías o un rango medio.

Variable continua — muchas categorías (a veces una infinidad).

A dichas variables independientes cuantitativas continuas, cuya influencia se remueve y controla, se les denomina “covariables”. Una covariable es incluida en el análisis para remover su efecto sobre la variable dependiente, e incrementar el conocimiento de la relación entre las variables independientes categóricas y la dependiente, aumentando la precisión del análisis.

En esta perspectiva, el *análisis de covarianza* puede ser concebido primero como un ajuste en la variable dependiente respecto a diferencias en la covariable o las covariables y, posteriormente, como una evaluación de la relación entre las variables independientes categóricas y los valores ajustados de la variable dependiente (Wildt y Ahtola, 1978).

B. Perspectiva de interés por la covariable. Esta perspectiva se ejemplifica con aquellas instancias, en las cuales el interés principal se centra en analizar la relación entre la variable dependiente y la covariable (variable cuantitativa continua) o las covariables. Aquí el enfoque es distinto; la influencia que se remueve es la de las variables independientes categóricas. Primero se controla el efecto (en este caso contaminante) de estas variables y después se analiza el efecto “purificado” de las covariables.

C. Perspectiva de regresión. En esta tercera perspectiva, tanto las variables independientes categóricas como las covariables resultan de interés para el investigador, quien puede desear exa-

minar el efecto de cada variable independiente (covariables y no covariables, todas) y después ajustar o corregir los efectos de las demás variables independientes.

En cualquier caso, el *análisis de covarianza elimina influencias no deseadas sobre la variable dependiente*. Se puede utilizar en contextos experimentales y no experimentales.

Wildt y Ahtola (1978, p. 13) definen algunos usos del análisis de covarianza:

1. *Incrementar la precisión en experimentos con asignación al azar.*
2. *Eliminar influencias extrañas o contaminantes* que pueden resultar cuando las pruebas o los individuos no son asignados al azar a las diferentes condiciones experimentales (grupos de un experimento).
3. *Eliminar efectos de variables que confundan* o distorsionen la interpretación de resultados en estudios no experimentales.

Nivel de medición de las variables: La *variable dependiente* siempre está medida por intervalos o razón y las *variables independientes* pueden estar medidas en cualquier nivel. Aunque las *covariables* deben medirse en un nivel de intervalos o razón.

Interpretación: Dependiendo de cada caso específico, el análisis de covarianza efectuado, mediante un paquete estadístico computacional, produce un cuadro de resultados muy parecido al del análisis de varianza. Los elementos más comunes del cuadro son:

<i>Fuente de variación (source of variation)</i>	<i>Sumas de cuadrados y productos cruzados (sum of squares and cross products)</i>	<i>Sumas de cuadrados ajustados (adjusted sum of squares)</i>	<i>Grados de libertad (degrees of freedom)</i>	<i>Medias cuadráticas ajustadas (adjusted mean squares)</i>	<i>Razón "F" (F-ratio)</i>	<i>Significancia (F prob.)</i>
--	--	---	--	---	----------------------------	--------------------------------

La *razón F* es, igual que en análisis de varianza, una razón de varianzas. El razonamiento estadístico es el mismo y *F* se interpreta igual, incluso se utiliza el mismo cuadro de la distribución *F* (tabla 3, apéndice 5).⁸ Solamente que las inferencias y conclusiones se hacen tomando en cuenta que las medias de la variable dependiente, a través de las categorías de las variables independientes, han sido ajustadas, eliminando el efecto de la covariable.

⁸ Material incluido en el CD que acompaña esta obra.

Ejemplo

Hi: "Los trabajadores que reciban retroalimentación verbal sobre el desempeño de parte de su supervisor mantendrán un nivel mayor de productividad que los trabajadores que reciban retroalimentación sobre el desempeño por escrito, y que los trabajadores que no reciban ningún tipo de retroalimentación."

$$H_i: \bar{X}_1 > \bar{X}_2 > \bar{X}_3$$

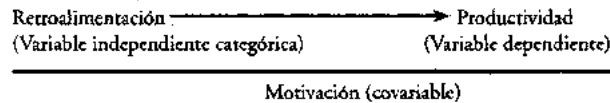
(verbal) (por escrito) (ausencia)

El investigador plantea un diseño experimental para intentar probar su hipótesis. Sin embargo, no puede asignar aleatoriamente a los trabajadores a los tres grupos de experimento. El diseño sería con intactos (cuasiexperimental) y se esquematizaría así:

G_1	X_1	\bar{X}_1
G_2	X_2	\bar{X}_2
G_3	—	\bar{X}_3

Asimismo, el investigador sabe que hay un factor que llega a contaminar los resultados (actuar como fuente de invalidación interna): la motivación. Diferencias iniciales en motivación pueden invalidar el estudio. Como no hay asignación al azar no se sabe si los resultados se ven influidos por dicho factor. Entonces, el experimentador decide eliminar y controlar el efecto de la motivación sobre la productividad para conocer los efectos de la variable independiente: tipo de retroalimentación. La motivación se convierte en covariable.

El esquema es:



Cabe destacar que, para introducir una *covariable* en el análisis, debe medirse de preferencia antes del inicio del experimento.

El *análisis de covarianza* "quita" a la variabilidad de la dependiente lo que se debe a la *covariable*. *Ajusta la varianza de la variable dependiente en las categorías de la independiente*, basándose en la covariable. En el ejemplo, ajusta la varianza de la productividad debida a la motivación, en las categorías experimentales (tratamientos o grupos). El ajuste se realiza sobre la base de la correlación entre la covariable y la dependiente. Esto se muestra esquemáticamente en el cuadro 10.12.

Una vez realizado el *análisis de covarianza*, se evalúa si *F* es o no significativa. *Cuando F resulta significativa se acepta la hipótesis de investigación.*

Si el resultado fuera:

$$G_1 = 35$$

$$G_2 = 36$$

La correlación entre la calificación en motivación y las puntuaciones en productividad es la base para el ajuste.

$$G_3 = 38$$

$$Gl_{entre} = K - 1 = 3 - 1 = \boxed{2}$$

$$Gl_{intra} = N - K = \boxed{107}$$

$$F = 1.70$$

Comparamos con el valor de la tabla respectiva que .05 es igual a 3.07; nuestra razón F 1.70 es menor a este valor.

Por lo tanto, rechazamos la hipótesis de investigación y aceptamos la hipótesis nula. Esto se contrasta con las medias ajustadas de los grupos que proporcione el análisis de covarianza (no las medias obtenidas en el experimento por cada grupo, sino las ajustadas con base en la covariable).

ANÁLISIS NO PARAMÉTRICOS

¿Cuáles son las presuposiciones de la estadística no paramétrica?

Para realizar análisis no paramétricos debe partirse de las siguientes consideraciones:

1. La mayoría de *estos análisis no requieren de presupuestos acerca de la forma de la distribución poblacional*. Aceptan distribuciones no normales.
2. Las *variables no necesariamente tienen que estar medidas en un nivel por intervalos o de razón*; pueden analizar datos nominales u ordinales. De hecho, si se quieren aplicar análisis no paramétricos a datos por intervalos o razón, éstos se resumen a categorías discretas (a unas cuantas). Las variables deben ser categóricas.

¿Cuáles son los métodos o las pruebas estadísticas no paramétricas más utilizados?

Las *pruebas no paramétricas más utilizadas* son:

1. La ji cuadrada o χ^2 .
2. Los coeficientes de correlación e independencia para tabulaciones cruzadas.
3. Los coeficientes de correlación por rangos ordenados de Spearman y Kendall.

Ejemplo de un diseño de investigación que utiliza el análisis de covarianza como instrumento para ajustar diferencias en motivación entre los grupos

	COVARIABLE	VARIABLE INDEPENDIENTE	VARIABLE DEPENDIENTE
	Calificación en motivación	Tipo de retroalimentación	Puntuaciones en productividad ajustadas por la covariable
G ₁	0	X ₁	0
G ₂	0	X ₂	0
G ₃	0	—	0

Cuadro 10.12

¿Qué es la ji cuadrada o χ^2 ?

Definición: Es una prueba estadística para evaluar hipótesis acerca de la relación entre dos variables categóricas.

Se simboliza: χ^2 .

Hipótesis a probar: Correlacionales.

Variables involucradas: Dos. La prueba ji cuadrada no considera relaciones causales.

Nivel de medición de las variables: Nominal u ordinal (o intervalos o razón reducidos a ordinales).

Procedimiento: La *ji cuadrada* se calcula por medio de una *tabla de contingencia o tabulación cruzada*, que es un cuadro de dos dimensiones, y cada dimensión contiene una variable. A su vez, cada variable se subdivide en dos o más categorías.

Un ejemplo de una tabla de contingencia se presenta en el cuadro 10.13.

El cuadro 10.13 demuestra el concepto de *tabla de contingencia o tabulación cruzada*. Las variables aparecen señaladas a los lados del cuadro, cada una con sus dos categorías. Se dice que se trata de una tabla 2×2 , donde cada dígito significa una variable y el valor de éste indica el número de categorías de la variable:⁹

⁹ Un concepto similar fue expuesto al hablar de diseños factoriales en el capítulo sobre experimentos, sólo que en aquellos casos se habla de dos o más variables y las celdas o recuadros incluían promedios de la variable dependiente. Aquí se está especificando que se trata únicamente de dos variables y las celdas contienen frecuencias.

Ejemplo de una tabla de contingencia 2 x 3 con las frecuencias observadas incluidos

Cuadro 10.14

IDENTIFICACIÓN POLÍTICA	ZONA DEL DISTRITO ELECTORAL		Total
	Norte	Sur	
Partido Derechista	180	100	280
Partido del Centro	190	280	470
Partido Izquierdista	170	120	290
Total	540	500	1 040

donde N es el número total de frecuencias observadas.

Para la primera celda (zona norte y partido derechista) la frecuencia esperada sería:

$$f_e = \frac{(280)(540)}{1\,040} = 145.4$$

Veamos de dónde salieron los números:

		280 total de renglón
540 Total de columnas		1 040 N

Para el ejemplo del cuadro 10.14, la tabla de frecuencias esperadas sería el cuadro 10.15.

Cuadro de
frecuencias
esperadas para
el cuadro 10.14

145.4	134.6	280
244.0	226.0	470
150.6	139.4	290
540	500	1 040

Cuadro 10.15

Una vez obtenidas las frecuencias esperadas, se aplica la siguiente fórmula de *ji cuadrada*:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

donde \sum significa sumatoria.

O es la frecuencia observada en cada celda.

E es la frecuencia esperada en cada celda.

Es decir, se calcula para cada celda la diferencia entre la frecuencia observada y la esperada; esta diferencia se eleva al cuadrado y se divide entre la frecuencia esperada. Finalmente, se suman tales resultados y la sumatoria es el valor de χ^2 obtenida.

Otra manera de calcular χ^2 es mediante el cuadro 10.16.

El valor χ^2 para los valores observados es de 47.33.

Interpretación: Al igual que t y F , la *ji cuadrada* proviene de una distribución muestral denominada distribución χ^2 , y los resultados obtenidos en la muestra están identificados por los grados de libertad. Esto es, para saber si un valor de χ^2 es o no significativo, debemos calcular los grados de libertad. Éstos se obtienen mediante la siguiente fórmula:

$$Gl = (r - 1)(c - 1)$$

donde r es el número de renglones del cuadro de contingencia y c el número de columnas. En nuestro caso:

$$Gl = (3 - 1)(2 - 1) = 2$$

Procedimiento para calcular el ji cuadrada (χ^2):

Cuadro 10.16

CELDA	O	E	O - E	(O - E) ²	$\frac{(O - E)^2}{E}$
Zona norte/partido derechista	180	145.4	34.6	1 197.16	8.23
Zona norte/partido centro	190	244.4	-54.4	2 959.36	12.11
Zona norte/partido izquierdista	170	150.6	19.4	376.36	2.50
Zona sur/partido derechista	100	134.6	-34.6	1 197.16	8.89
Zona sur/partido centro	280	226.0	54.0	2 916.00	12.90
Zona sur/partido izquierdista	120	139.4	-19.4	376.36	2.70

$\chi^2 = 47.33$

Y acudimos con los grados de libertad que corresponden a la tabla 4 del apéndice 5 (*Distribución de ji cuadrada*), eligiendo nuestro nivel de confianza (.05 o .01). Si nuestro valor calculado de χ^2 es igual o superior al de la tabla, decimos que las variables están relacionadas (χ^2 fue significativa). En el ejemplo, el valor que requerimos empatar o superar al nivel del .05 es 5.991. El valor de χ^2 calculado por nosotros es de 47.33, que es muy superior al de la tabla: χ^2 resulta significativa.

Hi: "Los tres canales de televisión a nivel nacional difieren en la cantidad de programas prosociales, neutrales y antisociales que difunden." "Hay relación entre la variable 'canal de televisión nacional' y la variable 'emisión de programas prosociales, neutrales y antisociales'."

Resultados: $\chi^2 = 7.95$
 $Gf = 4$

Para que χ^2 sea significativa a .01, con cuatro grados de libertad, se necesita un valor mínimo de 13.277 y para que sea significativa a .05, se necesita un valor mínimo de 9.488. Por lo tanto, se rechaza la hipótesis de investigación y se acepta la nula. No hay relación entre las variables.

Esquema:	Canal 24	Canal 32	Canal 56
Programas prosociales			
Programas neutrales			
Programas antisociales			

Cuadro de contingencia correspondiente al ejemplo

Cuadro 10.17

Comentario: Cuando al calcular χ^2 se utiliza un paquete estadístico computacional, el resultado de χ^2 se proporciona junto con su significancia; si ésta es menor a .05 o a .01, se acepta la hipótesis de investigación.

¿Qué son los coeficientes de correlación e independencia para tabulaciones cruzadas?

Además de la χ^2 , existen *otros coeficientes* para evaluar si las variables incluidas en la tabla de contingencia o tabulación cruzada están correlacionadas. A continuación, se mencionan algunos de estos coeficientes. No en todos se utilizan frecuencias.

Coficiente	Para cuadros de contingencia	Nivel de medición de las variables (ambas)	Interpretación
- $F_i(\phi)$	2 x 2	Nominal	Varía de 0 a +1, donde "cero" implica ausencia de correlación entre las variables y "más uno" significa que las variables están correlacionadas de manera perfecta.
- Coeficiente de contingencia ϕ o C de Pearson [C]	Cualquier tamaño	Nominal	Su valor mínimo es 0 (ausencia de correlación), pero su valor máximo depende del tamaño de la tabla de contingencia. Con

Coeficiente	Para cuadros de contingencia	Nivel de medición de las variables (ambas)	Interpretación
– V de Cramer (V)	Mayores de 2 x 2	Nominal	cuadros 2 x 2 varía de 0 a .707. Si se trata de tablas 3 x 3, varía de 0 a 0.816. Es un ajuste de F_i en cuadros mayores a 2 x 2. Varía de 0 a +1 con variables nominales ["cero" es nula correlación y "más uno" representa una correlación perfecta].
– Lambda (λ_b)	Cualquier tamaño	Nominal	Se utiliza con variables nominales y varía de 0 a +1 (+1 significa que puede predecirse sin error a la variable dependiente definida en la tabla, sobre la base de la independiente).
– Gamma (r)	Cualquier tamaño	Ordinal	Varía de -1 a +1 [-1 es una relación negativa perfecta, y +1 una relación positiva perfecta].
– Tau-b de Kendall (Tau-b)	Cualquier tamaño, pero más apropiado para tablas con igual número de renglones y columnas	Ordinal	Varía de -1 a +1.
– D de Somers	Cualquier tamaño	Ordinal	Varía de -1 a +1.
– Eta	Cualquier tamaño	Variable independiente nominal y dependiente por intervalos o razón; aquí no se calculan frecuencias en el cuadro, sino medias.	Es un indicador de cuán disimilares son las medias en la variable dependiente dentro de las categorías de la independiente. Si son idénticas, es igual a 0. Cuando son muy diferentes y las varianzas dentro de las categorías de la independiente son pequeñas, ésta puede incrementarse hasta 1 (Nie <i>et al.</i> , 1975).

¿Qué otra aplicación tienen las tablas de contingencia?

Las *tablas de contingencia*, además de servir para el cálculo de χ^2 y otros coeficientes, son útiles para describir conjuntamente dos o más variables. Esto se efectúa convirtiendo las frecuencias observadas en frecuencias relativas o porcentajes. En una tabulación cruzada puede haber tres tipos de porcentajes respecto de cada celda.

- Porcentaje en relación con el total de frecuencias observadas (N).
- Porcentaje en relación con el total marginal de la columna.
- Porcentaje en relación con el total marginal del renglón.

Veamos con un ejemplo hipotético de una tabla 2 x 2 con las variables sexo y preferencia por un conductor. Las frecuencias observadas serían:

		Sexo		
		masculino	femenino	
Preferencia por el conductor	A	25	25	50
	B	40	10	50
		65	35	100

Las celdas podrían representarse así:

<i>a</i>	<i>c</i>
<i>b</i>	<i>d</i>

Tomemos el caso de *a* (celda superior izquierda). La celda *a* (25 frecuencias observadas) con respecto al total ($N = 100$) representa 25%. En relación con el total marginal de columna (cuyo total es 65) representa 38.46% y respecto del total marginal de renglón (cuyo total es 50) significa 50%. Esto puede expresarse así:

Frecuencias observadas			
	25		
En relación con N	25.00%		
En relación con " $a + b$ "	38.46%	<i>c</i>	$a + c = 50$
En relación con " $a + c$ "	50.00%		
	<i>b</i>	<i>d</i>	$b + d$
	$a + b = 65$	$c + d$	$100 = N$

Así procedemos con cada categoría, como ocurre en el cuadro 10.18.

Comentario: Una cuarta parte de la muestra está constituida por hombres que prefieren al conductor A, 10.0% son mujeres que prefieren al conductor B. Más de 60% (61.5%) de los hombres prefieren a B, etcétera.

Ejemplo de una tabla de contingencia para describir conjuntamente dos variables

Cuadro 10.18

	Preferencia por el conductor		SEXO
	Masculino	Femenino	
A	25 25.0% 38.5% 50.0%	25 25.0% 71.4% 50.0%	50
B	40 40.0% 61.5% 80.0%	10 10.0% 28.6% 20.0%	50
	65	35	100

Debe observarse que tales *frecuencias relativas* se basan en las *frecuencias observadas*, pero no tienen nada que ver con frecuencias esperadas (estas últimas son frecuencias absolutas). La tabulación cruzada para describir conjuntamente variables y la tabulación cruzada para calcular estadísticas de correlación se basan en los mismos datos iniciales; pero representan funciones muy distintas.

¿Qué son los coeficientes y la correlación por rangos ordenados de Spearman y Kendall?

Los coeficientes *rho* de Spearman, simbolizado como r_s , y tau de Kendall, simbolizado como τ , son *medidas de correlación* para variables en un *nivel de medición ordinal*, de tal modo que los individuos u objetos de la muestra pueden ordenarse por rangos (jerarquías). Por ejemplo, supongamos que tenemos las variables "preferencia en el sabor" y "atractivo del envase", y pedimos a personas representativas del mercado que evalúen conjuntamente 10 refrescos embotellados y los ordenen del 1 al 10 (doble "1" es la categoría o el rango máximo en ambas variables), con los siguientes resultados:

Refresco ¹⁰	Variable 1 Preferencia en el sabor	Variable 2 Atractivo del envase
— Loy	1	2
— Caravana	2	5
— Mauna-Loa	3	1

¹⁰ Nombres ficticios.

— Recreo	4	3
— Carma	5	4
— Manzanol	6	6
— Cereza	7	8
— Pezcara	8	7
— Casa	9	10
— Manzanita	10	9

Para analizar los resultados, utilizaríamos los coeficientes r_s y t . Ahora bien, debe observarse que todos los sujetos u objetos tienen que jerarquizarse por rangos que contienen las propiedades de una escala ordinal (se ordena de mayor a menor). Ambos coeficientes varían de -1.0 (correlación negativa perfecta) a $+1.0$ (correlación positiva perfecta). Se trata de estadísticas sumamente eficientes para datos ordinales. La diferencia entre ellos es explicada por Nie *et al.* (1975, p. 289) de la manera siguiente: El coeficiente de Kendall resulta un poco más significativo cuando los datos contienen un número considerable de rangos empataados. El coeficiente de Spearman parece ser una aproximación cercana al coeficiente r de Pearson, cuando los datos son continuos (por ejemplo, no caracterizados por un número considerable de empates en cada rango).

También se interpreta su significancia.

CÁLCULO DEL COEFICIENTE DE CONFIABILIDAD ALFA-CRONBACH

De acuerdo con Carmines y Zeller (1988, pp. 44 y 45) existen dos procedimientos para calcular el coeficiente α :

1. Sobre la base de la varianza de los ítems, aplicando la siguiente fórmula:

$$\alpha = \frac{N}{(N-1) \left[\frac{1 - \sum s^2(Y_i)}{s^2x} \right]}$$

donde N es igual al número de ítems de la escala. " $\sum s^2(Y_i)$ " es igual a la sumatoria de las varianzas de los ítems y s^2x es igual a la varianza de toda la escala.

2. Sobre la base de la matriz de correlación de los ítems, el procedimiento sería:
 - a) Se aplica la escala.
 - b) Se obtienen los resultados.
 - c) Se calculan los coeficientes de correlación r de Pearson entre todos los ítems (todos contra todos de par en par).

d) Se elabora la matriz de correlación con los coeficientes obtenidos. Por ejemplo:

	Ítems			
	1	2	3	4
1	—	.451	.399	.585
2	ya fue calculado	—	.489	.501
3	ya fue calculado	ya fue calculado	—	.541
4	ya fue calculado	ya fue calculado	ya fue calculado	—

Los coeficientes que se mencionan como “ya fue calculado”, vienen en la parte superior de las líneas horizontales (guiones). Es decir, cada coeficiente se incluye una sola vez y los coeficientes se excluyen en las mismas puntuaciones (1 con 1, 2 con 2, 3 con 3 y 4 con 4).

e) Se calcula \bar{p} (promedio de las correlaciones y NP el número de correlaciones no repetidas o no excluidas).

$$\bar{p} = \frac{\sum P}{NP} \quad (\sum P \text{ es la sumatoria de las correlaciones y } NP \text{ el número de correlaciones no repetidas o no excluidas}).$$

$$\bar{p} = \frac{.451 + .399 + .585 + .489 + .501 + .541}{6}$$

$$\bar{p} = 0.494$$

f) Se aplica la fórmula siguiente:

$$\alpha = \frac{N\bar{p}}{1 + \bar{p}(N - 1)}$$

donde N es el número de ítems y \bar{p} el promedio de las correlaciones entre ítems.

En el ejemplo:

$$\alpha = \frac{4(0.494)}{1 + 0.494(4 - 1)}$$

$$\alpha = \frac{1.98}{2.48}$$

$$\alpha = 0.798$$

$$\alpha = 0.80 \text{ (cerrando)}$$

Es un coeficiente muy aceptable.¹¹

¹¹ Nota: Los procedimientos señalados incluyen varianza o correlación r de Pearson. Es decir, el nivel de medición de la variable es por intervalos o razón.

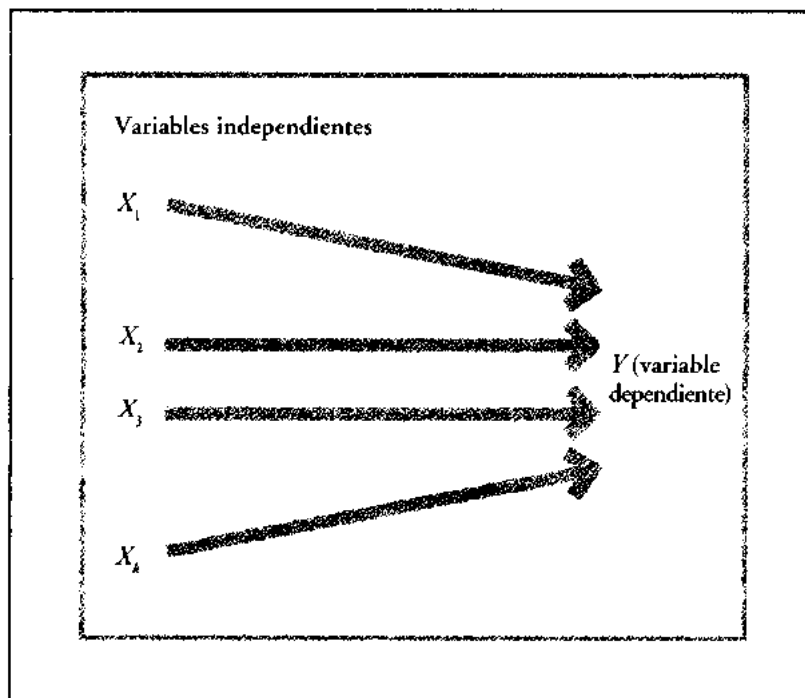
ANÁLISIS MULTIVARIADO

¿Qué son los métodos de análisis multivariado?

Los *métodos de análisis multivariado* son aquellos en que se analiza la relación entre *diversas variables independientes y al menos una dependiente*. Son métodos más complejos que requieren del uso de computadoras para efectuar los cálculos necesarios; normalmente se enseñan a nivel posgrado. A continuación se mencionan algunos de los principales métodos de análisis multivariado sin profundizar en ellos, pues esto rebasa los propósitos del libro.

¿Qué es la regresión múltiple?

Es un método para analizar el efecto de dos o más variables independientes sobre una dependiente. Asimismo, es una extensión de la regresión lineal sólo que con mayor número de variables independientes. O sea, sirve para predecir el valor de una variable dependiente conociendo el valor y la influencia de las variables independientes incluidas en el análisis. Si queremos conocer la influencia que ejercen las variables “satisfacción sobre los ingresos percibidos”, “antigüedad en la empresa”, “motivación intrínseca en el trabajo” y “percepción del crecimiento y desarrollo personal en el trabajo” sobre la variable “duración en la empresa”, el modelo de regresión múltiple es el adecuado para aplicarlo a los datos obtenidos. Este método es útil para analizar esquemas del tipo siguiente:



La información básica que proporciona la regresión múltiple es el coeficiente de correlación múltiple (R), que señala la correlación entre la variable dependiente y todas las demás variables independientes tomadas en conjunto.

El coeficiente puede variar de 0 a 1.00; cuanto más alto sea su valor, las variables independientes explicarán en mayor medida la diversidad de la variable dependiente o que son factores más efectivos para predecir el comportamiento de esta última. R^2 (el coeficiente de correlación múltiple elevado al cuadrado) nos indica el porcentaje de variación en la dependiente debida a las independientes.

Otra información relevante producida por el análisis de regresión múltiple son los valores “beta” (β) que indican el peso o influencia que tiene cada variable independiente sobre la dependiente. También se proporcionan coeficientes de correlación bivariados entre la dependiente y cada independiente.

Para predecir la variable dependiente se aplica la ecuación de regresión múltiple:

$$y = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots b_kX_k$$

donde a es una constante de regresión para el conjunto de puntuaciones obtenidas, $b_1, b_2, b_3, \dots b_k$ son los pesos “beta” de las variables independientes y X_1, X_2, X_3 y... X_k son valores de las variables independientes que fija el investigador para hacer la predicción.

La variable dependiente debe estar medida en un nivel por intervalos o de razón. Las independientes, en cualquier nivel de medición.

¿Qué es el análisis lineal de patrones o path analysis?

Es una técnica estadística multivariada para representar interrelaciones entre variables a partir de regresiones. Así como analizar la magnitud de la influencia de unas variables sobre otras, influencia directa e indirecta. Se trata de un modelo causal. Supongamos que tenemos el esquema causal que se observa en la figura 10.12 y deseamos probarlo.

El *análisis path* es un método para someterlo a prueba. La información principal que proporciona son los *coeficientes path*, los cuales representan la fuerza de las relaciones entre las variables (son coeficientes de regresión estandarizados).

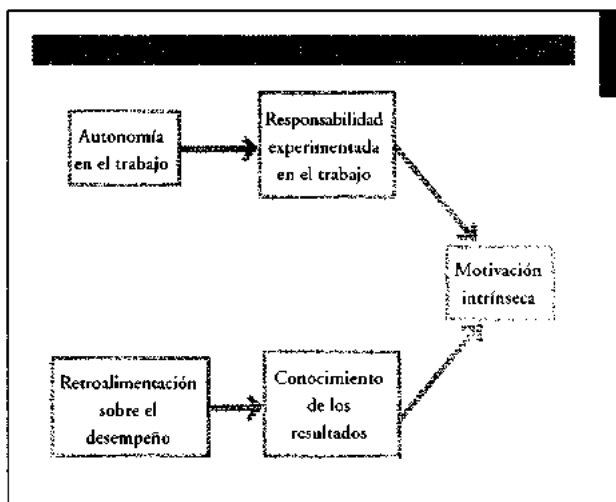
También proporciona información acerca de otras variables analizadas. Cuantifica efectos. En la figura 10.13 se muestra un ejemplo hipotético para ilustrar este tipo de análisis.

Cuanto más se acerque un *coeficiente “path”* a cero menor efecto tendrá.

¿Qué es el análisis de factores?

Es un método estadístico multivariado que se usa para determinar el número y la naturaleza de un grupo de *constructos* subyacentes en un conjunto de mediciones.

Un *constructo* es un atributo para explicar un fenómeno (Wiersma, 1999). En este análisis se generan "variables artificiales" (denominadas factores) que representan constructos. Los *factores* se obtienen de las variables originales y deben ser interpretados de acuerdo con éstas. Como menciona Naghi (1984), es una técnica para explicar un fenómeno complejo en función de unas cuantas variables.



Un análisis lineal path

Figura 10.12

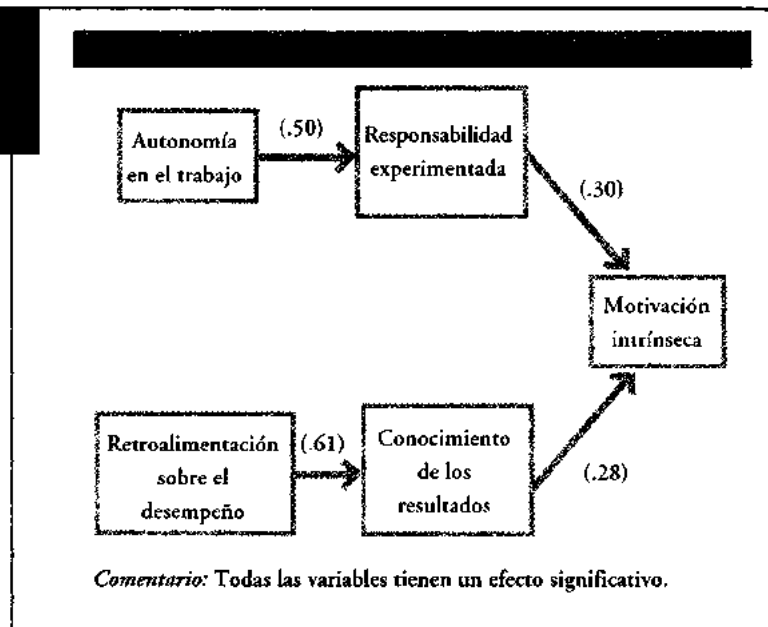
Un ejemplo del uso de esta técnica lo constituye una investigación realizada por Paniagua (1988) con la colaboración de los autores. El estudio pretendía analizar los factores que determinan la relación entre los vendedores y los compradores industriales de la ciudad de México. Se midieron variables entre las que destacan: coordinación (coord.), conflicto (confl.), frecuencia de la relación comprador-vendedor (frec.), reciprocidad económica en la relación (RF2), reciprocidad en el manejo de consideraciones administrativas (RF1) e importancia de la relación (monto de las operaciones) (impor.). Los resultados se muestran en el cuadro 10.19.

Obsérvese que debajo de las columnas *FI* a *FVI* aparecen coeficientes que corresponden a los ítems de una escala. Si estos coeficientes son medios o elevados, se dice que los ítems *cargan* o forman parte del factor correspondiente. Por ejemplo, los ítems 23, 24, 25 y 26 cargan en el primer factor (obtienen valores de .84392, .71642, .67853 y .74737, respectivamente) y no en otros factores (tienen valores bajos). Así, descubrimos una estructura de seis factores en 19 ítems. Los *factores* reciben un nombre para saber qué constructos se encuentran subyacentes. El análisis de factores también proporciona la varianza explicada y puede diagramarse en forma gráfica en las coordenadas *X* y *Y*. La técnica es compleja y habrá que conocerla muy bien. Es muy útil para la validez de constructo.

Análisis de factores: método estadístico multivariado que sirve para determinar el número y la naturaleza de un grupo de constructos subyacentes en un conjunto de mediciones.

Ejemplo de un análisis lineal path

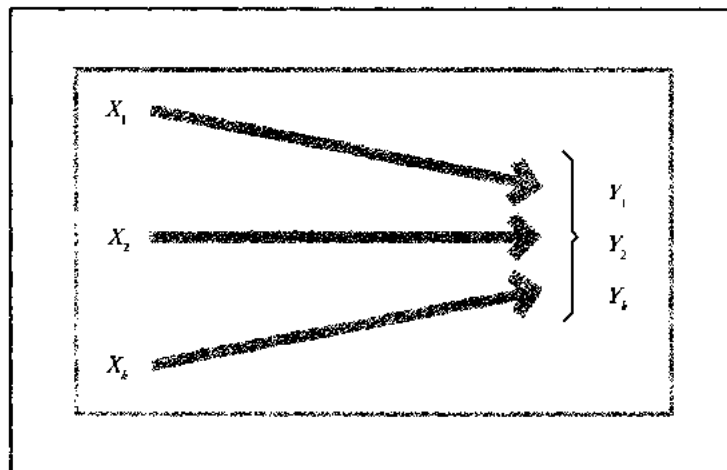
Figura 10.13



Las variables deben estar medidas en un nivel por intervalos o razón. El ejemplo de Núñez (2001) es otro caso de este análisis.

¿Qué es el análisis multivariado de varianza (MANOVA)?

Es un modelo para analizar la relación entre dos o más *variables independientes* y dos o más *variables dependientes*. Es decir, es útil para *estructuras causales* del tipo:



MATRIZ DE PATRÓN FACTORIAL ELEGIDA SUBMUESTRA "COMPRAS"										Ejemplo de algunos resultados en un análisis de factores
VARIABLE O COMPONENTE ESCALADO	ÍTEM	COMUNALIDAD	F1	FII	FIII	FIV	FVI	FVII	FACTOR EN QUE CARGA	
Coord	23	.66997	.84392	-.00895	-.11828	-.03405	.06502	-.27645		
	24	.46446	.71642	-.05609	-.01958	.07106	.00043	-.07127	F1	
	25	.59113	.67853	-.11175	.02581	-.09507	-.02857	.14978		
	26	.63988	.74737	.04695	.13472	-.04837	.07117	.02225		
Confl	47	.55724	-.05110	.62553	-.20945	-.05248	-.27963	-.06387		
	48	.59072	.06230	.65163	.17884	-.10916	.31061	.04245	FII	
	49	.35530	-.05665	.55503	-.11163	-.12946	-.07981	-.17370		
	50	.54716	-.07908	.61007	.08675	.20431	-.24058	.14142		
Frec	15	.46081	-.14748	-.06276	.63660	-.12476	-.06299	.13488		
	16	.44245	.03606	.08501	.62797	.00401	.11692	-.03968	FIII	
	18	.50307	-.05359	-.02787	.71169	.03786	-.10795	-.06775		
	19	.71311	.06664	.02881	.84213	.07161	-.16806	-.11058		
RF2	42	.46781	-.01380	-.18030	.09416	.63133	.17716	.06309		
	43	.50097	.10175	-.07970	-.16207	.64202	-.02808	-.18530	FIV	
RF1	40	.72202	.01579	-.03548	.04181	-.18914	.77312	.14292		
	41	.48405	.15684	-.18489	-.06425	.01958	.58187	.19379	FVI	
impor.	53	.31524	.02822	.02945	-.10069	.08605	.01579	.55431		
	54	.44550	-.04376	.08383	.01731	-.18396	.13956	.58137	FVII	
	58	.53236	.26836	-.05219	.10026	-.11741	-.02893	.55080		
EIGENVALUE			5.36433	2.53081	2.47521	1.55248	1.23464	1.06932		
% VARIANZA EXPLICADA			37.7%	17.8%	17.4%	10.9%	8.7%	7.5%	T = 100.0%	
DELTA = .00										
F1 = Coesificación		[explica 37.7% de la varianza]								
FII = Conflicto		[explica 17.8% de la varianza]								
FIII = Frecuencia		[explica 17.4% de la varianza]								

Cuadro 10.19

He aquí una síntesis de la explicación de Wiersma (1999) y sobre este tipo de análisis:

Al incluir dos o más variables dependientes simultáneamente no se consideran las diferencias entre las medias en cada variable, sino las diferencias en variables canónicas. El interés no sólo es saber si los grupos definidos por las variables independientes difieren en las *variables canónicas*, sino conocer la naturaleza de

éstas. Una *variable canónica* es una variable artificial generada a partir de los datos. Representa *constructos* y se compone de variables reales (las variables dependientes), las cuales deben ser descritas en términos de variables dependientes. Lo anterior se efectúa a través de las cargas de los coeficientes de correlación entre una variable dependiente y una variable canónica. Si una carga entre la *variable canónica* y la dependiente es positiva y elevada, significa que altos valores en la dependiente están asociados con altos valores en la canónica. Por ejemplo, si una variable dependiente consiste en puntuaciones a una prueba sobre innovación y creatividad, y dichas puntuaciones se correlacionan en forma considerable con una variable canónica, inferimos que la variable canónica representa un constructo que involucra la creatividad.

En los cálculos que se hacen en el *MANOVA*, se generan variables canónicas hasta que se encuentra que no hay una diferencia estadística significativa entre las categorías o los grupos; o bien, hasta que se agotan los grados de libertad de la variable independiente (lo que ocurra primero). El número de variables canónicas no puede exceder el número de variables dependientes, pero es común que el número sea mayor que el de *variables canónicas* estadísticamente significativas o los grados de libertad.

La hipótesis de investigación en el *MANOVA* postula que las medias en la variable dependiente de los grupos o las categorías de la variable independiente difieren entre sí. La hipótesis nula postula que dichas medias serán iguales. Se calculan diversas estadísticas para evaluar ambas hipótesis, entre las que destacan la *prueba Wilks-Lambda* y la *T² de Hotelling*; si resultan significativas a un nivel de confianza, se acepta la hipótesis de investigación de diferencia de medias. Esto indica que hay, por lo menos, una variable canónica significativa y que se presentan diferencias entre los grupos de la variable independiente en la variable canónica.

Los paquetes estadísticos que contiene el *MANOVA* suelen posicionar los grupos de las variables independientes por puntuaciones discriminantes; éstas son calculadas con una función discriminante, que es una ecuación de regresión para un compuesto de variables dependientes. A cada grupo se le asigna una puntuación discriminante en cada variable canónica. Las puntuaciones discriminantes de una variable independiente serían cero (redondeando). Una puntuación discriminante positiva, y elevada para un grupo, indica que éste se coloca por encima de los demás en la respectiva variable canónica. Y deben considerarse las cargas, las cuales son positivas o negativas. Cuando una variable dependiente tiene una carga fuerte (elevada) y negativa, aquellos grupos con puntuaciones discriminantes negativas “cargan” más fuerte en la contribución de la variable dependiente a la *variable canónica*. Las puntuaciones discriminantes son utilizadas para interpretar las separaciones de los grupos en las variables canónicas, en

Análisis multivariado de varianza: modelo para analizar la relación entre dos o más variables independientes y dos o más variables dependientes.

tanto que las cargas se usan para evaluar y ligar los resultados de las variables dependientes (Wiersma, 1999). Un ejemplo hipotético de las cargas de los coeficientes de correlación entre las variables dependientes y las variables canónicas se muestran en el cuadro 10.20, mientras que un ejemplo hipotético de las *puntuaciones discriminantes* viene en el cuadro 10.21.

Como observamos en el cuadro 10.21, se obtuvieron tres *constructos subyacentes* en las puntuaciones recolectadas de la muestra: motivación intrínseca, atribución de causalidad externa y desempeño laboral. Vemos en el cuadro 10.21 que los grupos están separados en las tres variables canónicas (los grupos difieren), particularmente en la primera variable canónica (motivación intrínseca) y los obreros ocupan la posición más baja. Las variables dependientes enmarcadas en un recuadro en la primera variable canónica se cargan en ella (cuadro 10.20); en consecuencia, los ejecutivos tienen las puntuaciones más altas en motivación intrínseca medida por la escala mencionada, en atribuciones internas y en sentimientos de éxito en el trabajo. Así se interpretan todas las variables canónicas y dependientes.

En el *MANOVA* se incluyen *razones F* y *análisis univariados de varianza*. Algunos paquetes estadísticos incluyen una prueba denominada *correlación canónica*, que es muy similar al *MANOVA*. Ésta es la máxima correlación que llega a obtenerse entre los conjuntos de puntuaciones y las relaciones entre las variables independientes, entre las variables dependientes y entre los conjuntos de ambas (dependientes e independientes) (Kerlinger, 1979). Las variables en el *MANOVA* y la correlación canónica asumen que las variables están medidas en un nivel de intervalos o razón. Tal correlación se interpreta como otras; pero el contexto de interpretación varía de acuerdo con el número de variables involucradas.

VARIABLE DEPENDIENTE	VARIABLES CANÓNICAS			Cargas de los coeficientes de correlación entre las variables dependientes y las variables canónicas
	I (MOTIVACIÓN INTRINSECA)	II (ATRIBUCIÓN DE CAUSALIDAD EXTERNA)	III (DESEMPEÑO LABORAL)	
– Motivación intrínseca [escala intrínseca del inventario de características del trabajo]	.90	.05	.07	Cuadro 10.20
– Atribuciones internas	.86	-.07	-.09	
– Sentimientos de éxito en el trabajo	.70	.08	.00	
– Atribuciones externas	.03	.61	.12	
– Productividad	-.12	.07	.74	
– Eficiencia	.18	-.04	.48	
– Calidad	.19	.13	.57	

Puntuaciones discriminantes con cuatro grupos en tres variables canónicas	GRUPO	VARIABLES CANÓNICAS		
		I	II	III
	Ejecutivos	1.97	0.95	-1.69
	Secretarios	-0.19	1.18	1.25
	Empleados	1.40	-1.01	-0.49
	Obreros	-3.18	-1.12	0.93

Cuadro 10.21

¿Hay otros métodos multivariados?

En la actualidad, hay muchos métodos multivariados de análisis, los cuales se han desarrollado con la evolución de la computadora. Los investigadores disponen del *análisis discriminante*, cuando las variables independientes son medidas por intervalos o razón, y la dependiente es categórica. Tal análisis sirve para predecir la pertenencia de un caso a una de las categorías de la variable dependiente, sobre la base de varias independientes (dos o más). Se utiliza una ecuación de regresión llamada *función discriminante*. Si queremos predecir el voto obtenido por dos partidos contendientes (variable dependiente nominal con dos categorías) sobre la base de cuatro variables independientes, aplicaremos el análisis discriminante, para resolver una ecuación de regresión; así se obtienen las predicciones individuales. En el ejemplo, hay dos categorías (votar por *A* o votar por *B*); por tanto, los valores a predecir son 0 y 1 (*A* y *B*, respectivamente). Si el sujeto obtiene una puntuación más cercana a cero, se predice que pertenece al grupo que votará por *A*; si logra una puntuación más cercana a 1, se predice que pertenece al grupo que votará por *B*. Además, se consigue una medida del grado de discriminación del modelo.

Se cuenta también con el *análisis de agrupamientos o conglomerados* (técnica para agrupar los casos o elementos de una muestra en grupos sobre la base de una o más variables), el *escalamiento multidimensional* (para diseñar escalas que midan a los sujetos en diversas variables y los ubiquen simultáneamente en los ejes de las distintas variables, así como para conocer la estructura de las variables entre sí), el *análisis de series cronológicas o de tiempo* (para analizar la evolución de los casos en una o más variables a través del tiempo y predecir el comportamiento de las variables o sucesos) y la elaboración de *mapas multidimensionales* (donde establecemos distancias entre casos, basándonos en mediciones múltiples de varias dimensiones o variables), para los cuales se requieren bases sólidas en materia de estadística y en matemáticas avanzadas. Sugerimos Ferrán (2001), para una revisión de tales pruebas y modelos.

¿CÓMO SE LLEVAN A CABO LOS ANÁLISIS ESTADÍSTICOS?

Actualmente, los análisis estadísticos se llevan a cabo a través de *programas computacionales* utilizando *paquetes estadísticos*. Estos paquetes son sistemas integrales de programas diseñados para el análisis de datos. Cada paquete tiene su propio formato, sus instrucciones, sus procedimientos y sus características. Para conocer un paquete es necesario consultar el manual respectivo. Los manuales de los paquetes más importantes han sido publicados y difundidos con amplitud.

El procedimiento para analizar los datos, una vez que tenemos la matriz, es crear o aplicar un programa basándonos en el manual. Tal programa incluye el llamado de la matriz de datos y las pruebas estadísticas seleccionadas. Después se corre el programa o se ejecutan los comandos (instrucciones) y se obtienen los resultados, los cuales se interpretan.

Los principales *paquetes estadísticos* conocidos hoy son SPSS y Minitab. Ambos se encuentran disponibles en diferentes versiones para sistemas Windows y Macintosh.

SPSS® (Statistical Package for the Social Sciences)

El Paquete Estadístico para Ciencias Sociales, desarrollado en la Universidad de Chicago, es uno de los más difundidos. Contiene todos los análisis estadísticos descritos en este capítulo.

En América Latina, algunas instituciones educativas tienen versiones antiguas del SPSS; otras, una versión titulada SPSS-X (que es más flexible que las versiones anteriores), en tanto que ciertas instituciones cuentan con versiones más recientes (que son completamente flexibles), como son:

- SPSS para Windows
- SPSS para Macintosh
- SPSS para UNIX

Desde luego, estas versiones sólo pueden utilizarse en computadoras personales y portátiles con la capacidad necesaria para el paquete.

Como ocurre con todos los programas o software, SPSS constantemente se actualiza con versiones en inglés y español. Cada año surgen textos o manuales acordes con las nuevas versiones. Por ello, no vale la pena mencionar alguna referencia, pues será obsoleta cuando el lector revise estas líneas. Lo mejor para mantenerse al día en materia de SPSS, como en cualquier otro programa, es “navegar” en Internet, ya sea entrando mediante un directorio o disparador de búsqueda con palabras clave tales como: “SPSS Bolivia”, “SPSS Venezuela” o “SPSS Mexico” (si en nuestro país no hay página, acudimos a otro). O bien,

consultando páginas específicas de SPSS. Para la actualización, las palabras clave serían: “SPSS manuals” (recordemos que para cruzar palabras, éstas tienen que ir entre comillas “ ”). Además, habrá que consultar la revista periódica *Keywords*, cuyo propósito es mantener al corriente a los usuarios.

MINITAB®

Es un paquete que goza de popularidad por su relativamente bajo costo. Incluye un considerable número de pruebas estadísticas, y cuenta con un tutorial para aprender a utilizarlo y practicar; además, es muy sencillo de manejar.

Minitab tiene una página Web a la que es posible entrar mediante la palabra clave “minitab”. Por otro lado, podemos bajar a nuestro disco duro un *demo* del programa. En Internet existen decenas de programas de análisis estadístico para cualquier ciencia o disciplina; entre ellos:

1. *ESP* (Paquete Econométrico de Software). Especialmente útil para análisis estadístico de series cronológicas. Se puede tener en máquinas IBM, aunque hay adaptaciones para otras máquinas.
2. *BMDP* (Programa Biomédico Computarizado). Desarrollado por la Universidad de California en Los Ángeles, es utilizable en máquinas IBM y en otros sistemas (CYBER, Honeywell, Univac, Xerox, etcétera). Aunque fue diseñado para el área biomédica, contiene gran cantidad de análisis aplicables a ciencias sociales.
3. *OSIRIS* (Organized Set of Integrated Routines for Investigation with Statistics, Conjunto Organizado de Rutinas Integradas para la Investigación con Estadísticas). Fue desarrollado por el Instituto de Investigación Social de la Universidad de Michigan. Disponible en máquinas IBM y en otras. El manual puede pedirse a dicho instituto.
4. *SAS* (Sistema de Análisis Estadístico). Fue diseñado en la Universidad de Carolina del Norte y es distribuido por SAS Institute, Inc., de Raleigh, de ese estado. Es muy poderoso y su utilización se ha incrementado notablemente. Resulta un paquete muy completo para computadoras personales que contiene variedad de pruebas estadísticas.

En el texto se incluye un software sencillo que hemos titulado *STATS®*.

En el caso de SPSS y Minitab, ambos paquetes son fáciles de usar, pues lo único que hay que hacer es definir las variables, seleccionar las opciones y solicitar los análisis requeridos.

Por ejemplo, en SPSS para Windows:

1. Seleccionamos la ventana de SPSS.

2. SPSS nos muestra en pantalla una matriz de datos, como se ve en la figura 10.14. Recordemos que SPSS tiene dos iconos que nos enseñan: la matriz de datos (vista de datos, *data view*) y el libro de códigos, así como la definición de variables (vista de variables, *variable view*).
3. Definimos cada variable. En el caso de versiones anteriores a la 10, lo haremos seleccionando “Data” (Datos), luego “Define” (Definir) y “Variable”; en tanto que en versiones 10 o superiores, mediante la vista de variables.
4. Llenamos la matriz con los datos correspondientes.
5. Seleccionamos la opción de “Analyze” (Analizar) y elegimos las pruebas apropiadas.

Los análisis, que son efectuados por la computadora, los vemos en pantalla o los mandamos a imprimir.

Si deseamos gráficos, usamos la ventana u opción de “Graphs” (Gráficos). Cuando queremos crear escalas o transformar variables, tenemos la ventana de Transform (Transformar).

Para conocer mejor el programa SPSS y los análisis disponibles, recomendamos Ferrán (2001) y Camacho (2001).

En la figura 10.14 se observa la ventana principal de SPSS para Windows, correspondiente a la versión 10 en español.

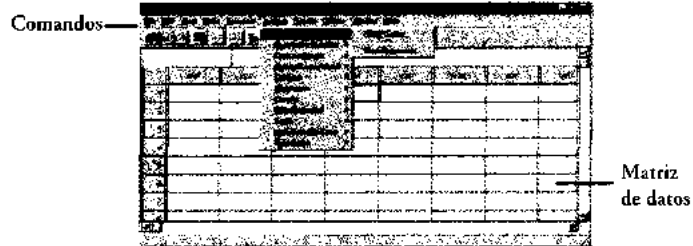
ANÁLISIS CUALITATIVO DE LOS DATOS

El primer reto del análisis cualitativo es que generalmente nos encontramos con gran volumen de datos (numerosas páginas de transcripciones de entrevistas y sesiones, cintas de audio y video, notas, bitácoras, diagramas, fotografías, etcétera). Por ello, el material debe encontrarse muy bien organizado (clasificado por temas, subtemas, orden de recolección, etcétera).

La primera tentación del analista es reducir datos; de hecho, parte del análisis cualitativo consiste en hacerlo, pero teniendo sumo cuidado de no perder información ni descartar datos valiosos. Las transcripciones tienen que estar completas, aunque a veces sean repetitivas en cuanto al contenido.

En los estudios cualitativos, el análisis de los datos no está determinado completamente, sino que es “prefigurado, coreografiado o esbozado”. Es decir, se comienza a efectuar bajo un plan general, pero su desarrollo va sufriendo modificaciones de acuerdo con los resultados (Dey, 1993). Por ejemplo, puede ocurrir que clasifiquemos contenidos de entrevistas en temas y categorías, y al hacerlo elaboremos una gráfica o un esquema, después que incorporemos comentarios de nuestras notas y nos demos cuenta de que es importante agregar el análisis de fotografías, así como un video, etcétera.

Figura 10.14



El SPSS, versión para ambiente Windows, trabaja de una manera muy sencilla: El investigador usuario selecciona las opciones más apropiadas para su análisis, tal y como se hace en otros programas que se encuentran en dicho ambiente.

File (Archivo): Para construir un nuevo archivo, localizar un archivo ya construido, guardar archivos, especificar impresora, imprimir, cerrar, enviar archivos por correo electrónico...

Edit (Editar): Para modificar archivos, buscar datos, copiar, cortar, eliminar...

Data (Datos): Definir tiempo, insertar variables y/o casos, dividir el archivo o matriz de datos en grupos –sexo, ocupación, edad–, seleccionar casos para análisis parciales y definir las variables (en el caso de versiones anteriores a la versión 10: nombre, tipo, categorías o *labels*), especificar valores perdidos y formato –ancho y puntos decimales–, sopesar casos, y otras funciones.

Transform (Transformar): Recodificar, conjuntar o unir y modificar variables y datos.

Analyze (Analizar): Solicitar análisis estadísticos básicamente:

1. Sumarize –Sumarios–: Estadísticas descriptivas (tablas de frecuencias y estadísticas de tendencia central y dispersión).
2. Comparar medias (prueba *t* y ANOVA unidireccional).
3. ANOVA (análisis de varianza).
4. Correlación.
5. Regresión.
6. Pruebas no paramétricas.
7. Respuestas múltiples (escalas y confiabilidad).

Graphs (Gráficas): Solicitar gráficos (histogramas, barras de pastel, diagramas de dispersión, etcétera).

Utilities (Herramientas): Definir ambientes, variables, etcétera.

Window (Ventana): Moverse a través de archivos y hacia otros programas.

Help (Ayuda): Contenidos, cómo utilizar SPSS, comandos, guías y demás elementos de Windows aplicados al paquete.

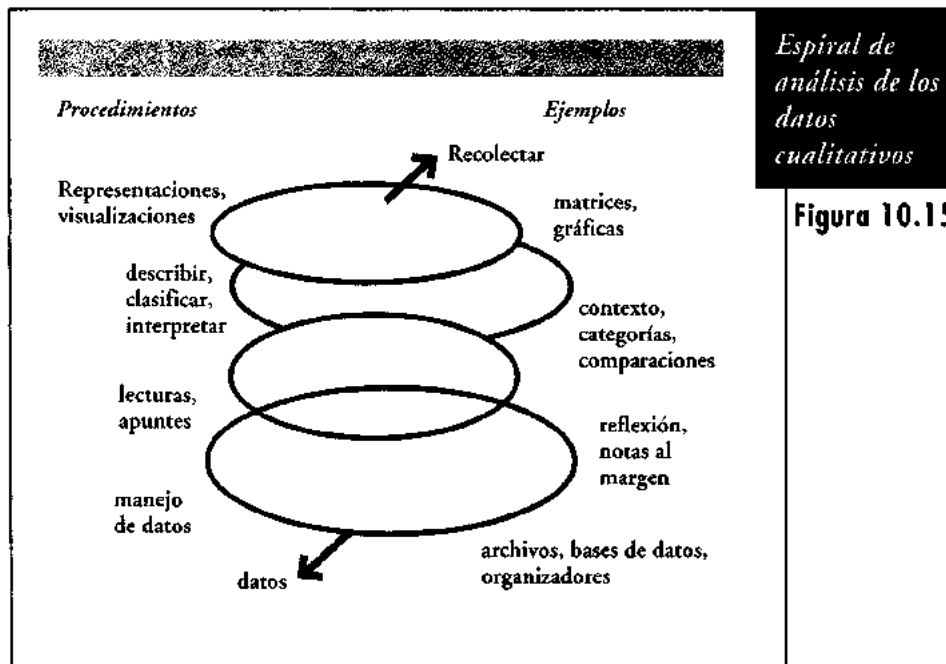
Nota. SPSS es 100% compatible con Excel, por lo que con la base de datos en SPSS se pueden obtener las gráficas y cálculos de Excel.

Creswell (1998) simboliza el desarrollo del análisis cualitativo como una espiral, en la cual se cubren varias facetas o diversos ángulos del mismo objeto de estudio. Esto se muestra en la figura 10.15.

Sin embargo, el análisis cualitativo tiene sus propósitos centrales:

- Darle orden a los datos (Patton, 1990).
- Organizar las unidades, las categorías, los temas y los patrones (Patton, 1990; Grinnell, 1997).
- Comprender en profundidad el contexto que rodea los datos.
- Describir las experiencias de las personas estudiadas bajo su óptica, en su lenguaje y con sus expresiones (Grinnell, 1997).
- Interpretar y evaluar unidades, categorías, temas y patrones (Patton, 1990).
- Explicar contextos, situaciones, hechos, fenómenos (Baptiste, 2001).
- Generar preguntas de investigación e hipótesis.
- Reconstruir historias (Baptiste, 2001).
- Relacionar los resultados del análisis con la teoría fundamentada o construir teorías (Charmaz, 2000; Baptiste, 2001).

Algunas personas piensan que el análisis cualitativo se limita a descripciones narrativas, como si se tratara de una novela costumbrista. Pareciera que la consig-



na fuera: "entra a una comunidad, observa, pregunta lo que sucede y narra una historia" (por ejemplo, cómo viven un grupo étnico —qué come, cuáles son sus costumbres, cómo está constituido, quién es el líder— o un grupo social). Esto puede ser parte de la actividad del analista cualitativo, pero no se limita a ello.

Así, análisis cualitativo *no* es sinónimo de pura descripción ni de caos o desorden. Como señala Grinnell (1997), sin plan de análisis se corre el riesgo de obtener conclusiones débiles y hasta tendenciosas. Contar con éste no significa que debamos ser rígidos a la hora de analizar los datos. El enfoque cualitativo requiere flexibilidad.

Dentro del análisis cualitativo se encuentran diversidad de estrategias, que pueden ser tan variadas como posiciones hay sobre el enfoque cualitativo. Sin embargo, al revisar a diversos autores que tratan el análisis cualitativo (Baptiste, 2001; Berg, 1998; Creswell, 1998; Denzin y Lincoln, 2000, y Grinnell, 1997), así como con base tanto en nuestras experiencias como en reflexiones, proponemos la siguiente "coreografía" del proceso de análisis cualitativo, que se resume gráficamente en la figura 10.16 y que explicaremos a detalle.

1. Revisar el material

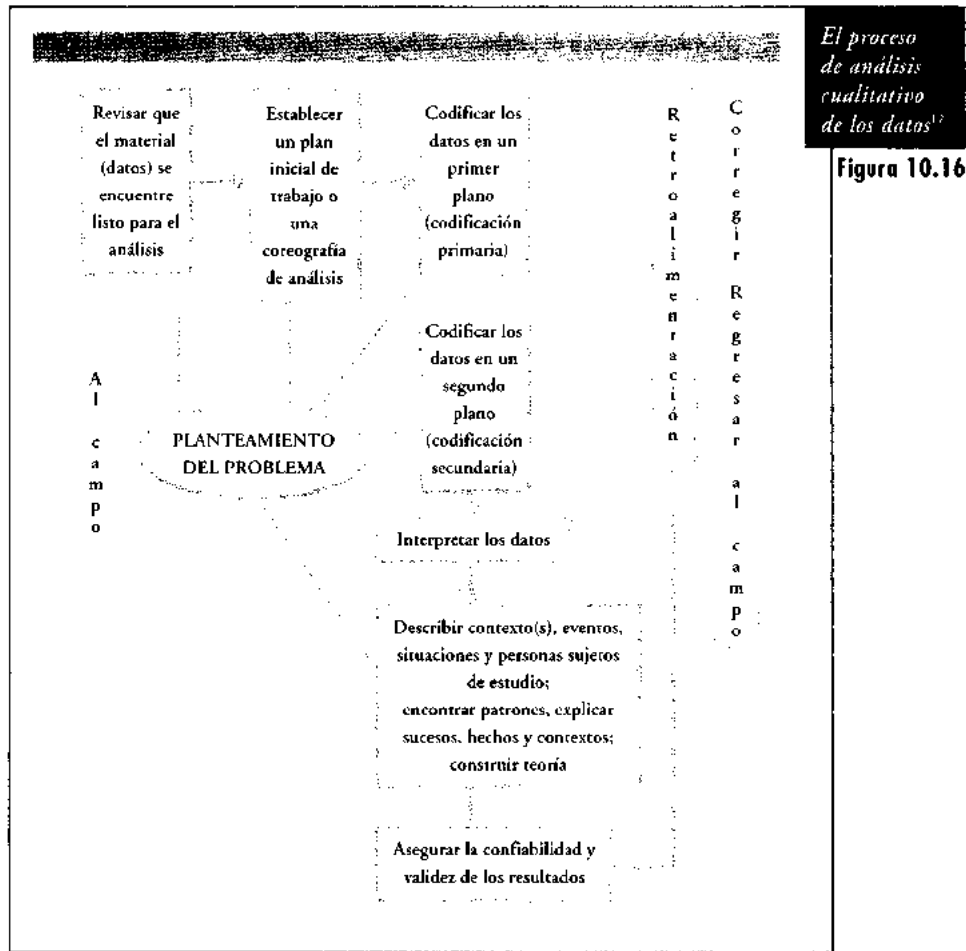
El primer paso del análisis cualitativo, independientemente del método de recolección de datos que hayamos seleccionado, consiste en revisar que los datos hayan sido preparados en forma adecuada para el análisis; esto es, que se encuentren organizados y clasificados por un criterio lógico (temas, actores relevantes,

etcétera), además de ser legibles, audibles o visibles (dependiendo de si se trata de notas o textos, audiocasetes, CD's, fotografías, mapas, videos), y que estén completos.

En este momento es importante asegurarnos de que contemos con el equipo apropiado para manejar los datos (computadora con el software adecuado, equipo de audio y video, escáner, etcétera). Asimismo, debemos de-

En una investigación cualitativa, se pueden escuchar y analizar grabaciones contenidas directamente en cassetes.





cidir cómo vamos a analizar los datos: si directamente del medio en que fueron producidos o de transcripciones escritas. A veces se hace de las dos formas.

¹² Esta visión de análisis como proceso no contradice la naturaleza cambiante y evolutiva del análisis cualitativo, puesto que los procesos son dinámicos, cambiantes e itinerantes.

Por ejemplo, grabaciones de audio de una sesión cualitativa pueden transcribirse a un procesador de palabras y analizarse (obvio, en las transcripciones se incluyen las expresiones paralingüísticas de los sujetos: pausas, muecas, etcétera; ¡oh!, ¡mmm!, ¡eh!...); pero también es posible analizar la grabación contenida en el casete.

Los videos pueden analizarse con el texto transcrito en el procesador de palabras, al reproducirse en una videocasetera u otro aparato; además, es posible introducir segmentos de video en la computadora (con el software apropiado).

Si se van a analizar transcripciones escritas en un procesador de texto, deben formatearse los contenidos y cada línea se numera (los programas computacionales para análisis cualitativo hacen esto automáticamente como se observará más adelante).

También habrá que recordar que la información cualitativa en sí no es anónima (se trata de escritos de personas, testimonios de voz, conductas filmadas, etcétera), por lo cual debe mantenerse siempre la confidencialidad de los sujetos (es cuestión de ética).

2. Establecer un plan de trabajo inicial (coreografía)

En el análisis cualitativo, el investigador goza de mayor libertad para trabajar los datos; pero esta forma también representa un esfuerzo sistemático y riguroso que requiere de reglas que guían las tareas analíticas y de un plan de trabajo (Grinnell, 1997), el cual llega a sufrir alteraciones o cambios, dependiendo de la evolución que siga el mismo análisis.

El plan incluye: *a*) revisar todos los datos (entrevistas, sesiones, documentos, notas de observación, biografías, etcétera); *b*) revisar la bitácora elaborada durante la etapa de recolección de los datos, las notas y los memos de campo; *c*) establecer un sistema de codificación de datos; *d*) definir el método para analizar los datos; *e*) definir los programas computacionales de análisis que se van a utilizar (software requerido); *f*) organizar el equipo de investigadores que habrán de participar, así como el papel que jugará cada uno de los miembros, y *e*) establecer fechas tentativas para la realización del análisis. Tales actividades se mencionan a continuación, pues son las acciones subsecuentes del análisis de los datos.

Grinnell (1997) sugiere que durante el análisis se documente el proceso mediante una bitácora, la cual consiste en un diario donde se anotan y describen cada actividad realizada, las ideas que van surgiendo producto del análisis, comentarios respecto de la credibilidad del estudio, información contradictoria, la manera como se ajustan la codificación, los problemas y logros en el análisis, los significados que paulatinamente se encuentren al analizar los datos, las razones por las cuales se procede de una u otra formas y las notas en torno al método utilizado.

Esta bitácora de análisis será un instrumento invaluable para la validez y confiabilidad del análisis (al igual que la bitácora de campo nos servía para establecer la validez y confiabilidad del método de recolección de datos).

La bitácora se escribe diariamente (anotando la fecha) y cada investigador sigue su propio sistema para llenarla. Sin embargo, podemos categorizar las anotaciones en: 1. comentarios acerca del método de análisis, 2. comentarios sobre los problemas durante el proceso, 3. comentarios sobre la codificación, 4. ideas y comentarios de los investigadores (incluyendo diagramas, mapas conceptuales, dibujos, esquemas, matrices), 5. material de apoyo localizado (fotografías, videos, etcétera) y 6. significados, descripciones y conclusiones preliminares.

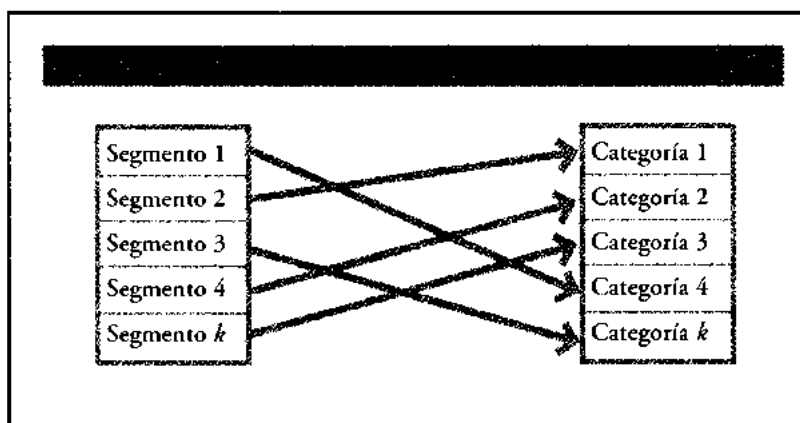
3. Codificar los datos en un primer nivel o plano

En la mayoría de los estudios cualitativos se codifican los datos para tener una descripción más completa de éstos, resumirlos, eliminar información irrelevante, realizar análisis cuantitativo y generar mayor sentido de entendimiento del material analizado.

La codificación tiene dos planos o niveles: en el primero, se codifican las unidades en categorías; en el segundo, se comparan las categorías entre sí para agruparlas en temas y buscar posibles vinculaciones. Veamos el primer nivel.

La codificación implica clasificar y, en esencia, requiere asignar unidades de análisis a categorías de análisis. Algunos autores la denominan categorización. La codificación comienza cuando observamos diferencias y similitudes entre segmentos de los datos. Por ejemplo, entre líneas o párrafos.

El proceso de codificar puede representarse gráficamente de la siguiente manera:



Los segmentos son las unidades de análisis. El investigador debe decidir cuáles van a ser, pues varían de acuerdo con el tipo de material a analizar. Ejemplos de segmentos o unidades son:

En textos:

1. *Palabras.* “Alcoholismo”, “Ricardo”, “Divorcio”.
2. *Líneas.* “Mi esposo me abandonó después de que me embaracé por tercera vez.”
3. *Párrafos.* “No puedo dejar de pensar (¡mmm!) en que mis hijas vean a su padre completamente ebrio. Es algo en lo que pienso todas las noches antes de acostarme. Ojalá y dejara la bebida (¡uhh!), pero lo veo como algo imposible. No ha podido dejar de hacerlo desde que lo conozco... (¡mmm!) pero antes bebía mucho menos...”
4. *Intervenciones de sujetos* (desde que comienza hasta que concluye su intervención cada uno).

Jesús: *No puedo dejar de beber, no puedo (jugg!).*

Alejandra: *Ni lo quieres intentar. Piensa en lo mal que te sientes, en tus hijas. ¿Qué va a pasar cuando sean grandes?*

En este caso tenemos dos unidades de análisis.

5. *Páginas.*
6. *Cambios de tema* (cada vez que aparece un nuevo tema).
7. *Todo el texto.*

En grabaciones de audio o video (aunque hayan sido transcritas a texto):

1. *Palabras o expresiones.*
2. *Intervenciones de sujetos.*
3. *Cambios de tema.*
4. *Periodos* (segundo, minuto, cada *k* minutos, hora...).
5. *Sesión completa* (entrevista, grupo de enfoque...).

Biografías:

1. *Día, mes, año, periodo, pasaje de vida.*
2. *Cambios de tema.*
3. *Actos conocidos.*

Música:

1. *Línea de canción.*
2. *Estrofa.*
3. *Canción completa.*
4. *Obra.*

Otros:

1. *Diagramas.*
2. *Mapas.*
3. *Fotografías.*
4. *Arrefactos* (objetos personales, cuadros pictóricos, esculturas, piezas de ropa, productos domésticos, etcétera).

Las unidades de análisis son –como puede verse– unidades de significado. Como menciona Grinnell (1997), se trata de los bloques de construcción en un esquema de categorías. En el cuadro 10.22 se muestra un escrito con 13 unidades de análisis, definiendo a la unidad como línea:

En nuestra bitácora de análisis, por ejemplo, anotaríamos que las unidades de análisis elegidas fueron las líneas. Las unidades pueden ser más cortas o largas, el caso es que sean “manejables”; ante la duda, más vale reducir su extensión, al menos al principio (Grinnell, 1997). También suele ocurrir que durante el proceso cambiemos las unidades de análisis.

Posteriormente, se definen las categorías de análisis (de manera similar al análisis de contenido que fue revisado en el capítulo anterior) y se etiquetan (nombre y código). Las categorías son “cajones” conceptuales.

En la codificación de primer plano se generan categorías de manera más concreta y en el segundo plano, de forma más abstracta. En este segundo nivel, lo que obtenemos son categorías generales o temas.

Las categorías pueden surgir de los datos o ser impuestas por el investigador (Baptiste, 2001), pero siempre tienen que estar estrechamente relacionadas con

1. Carolina: Aquel día fue muy triste. Nunca había llorado
2. tanto (**pausa**). Mi padre había fallecido. Era terrible, nunca
3. había sentido eso que sentí: vacío, soledad, miedo, mucho
4. miedo. No sabía qué hacer, no quería hablar con nadie, nada, no
5. quería nada. Sabía que ya no iba a ser lo mismo: ¿con quién voy
6. a platicar? (**¡oyyy!**), ¿a quién le voy a contar que tuve una
7. caída?, ¿quién me va a aconsejar a partir de ahora? ¡Ojalá,
8. aunque no hable con él, me señale qué debo hacer en el
9. futuro (**¡mmm!**). No sé, no sé, no sé que voy a hacer. La vida ya
10. no será igual sin él, sin “papito” (**melancólica**). He perdido
11. hasta el gusto por vivir. ¿Qué me espera? (**pausa**). Temo que
12. hasta pueda beber.
13. Mariane: A mí me ocurrió algo parecido pero diferente...

Ejemplo de
unidades de
análisis

Cuadro 10.22

los datos. De hecho, las categorías van apareciendo conforme se revisan las unidades de análisis. Así, tomamos la primera unidad de análisis y definimos la categoría a la que pertenece; luego le damos un nombre y un código. Después, tomamos la segunda unidad de análisis y definimos la categoría a la que pertenece (tal vez la misma que para la primera unidad o una nueva); de inmediato le damos un nombre y un código. Posteriormente, tomamos la tercera, la cuarta y cada una de las unidades, y repetimos el proceso.

Veámoslo con un ejemplo sencillo.

Investigación sobre la experiencia negativa de una mujer golpeada por su esposo y los tipos de violencia que ejercen los maridos que abusan de sus parejas.

Recolección de los datos: Entrevistas en profundidad.

Unidad de análisis: Línea.

Contexto: Entrevista con una joven esposa de 20 años, dos años de casada, de origen humilde, que vive en los suburbios de Valledupar, Colombia.

1. Carolina: Mi esposa me ha golpeado varias veces (jehhl). [pausa]
2. No sé cómo decirlo. Me pega con la mano abierta y con el puño.
3. La última vez me dijo: "Eres una ramera". También me ha
4. dicho que soy malnacida, perra. Siempre me insulta. Y la
5. verdad es que nunca he dado motivo. Nunca (pausa). Me dice que
6. los hombres se meten en mí como culebras. Que me gusta hacerlo
7. quedar mal. Me mira con odio del malo. Me amenaza con los ojos.
8. Y a veces le contesto y le pego también. El otro día le rompí
9. una lámpara en la cabeza...

La primera línea (unidad de análisis) caería en la categoría "violencia física"; la segunda también en "violencia física"; y la tercera en "violencia verbal". La cuarta es "violencia verbal". La séptima, "violencia psicológica".

Categorías: "cajones" conceptuales creados a partir de la revisión de datos, por comparación de una unidad respecto de las unidades que le precedieron.

En esta transcripción observamos tres tipos de violencia. Asimismo, notamos que el proceso de generar categoría se realiza sobre la base de la comparación constante entre unidades de análisis. Es decir, cada unidad se clasifica como similar o diferente de otras, en una categoría, por sus cualidades, y a la cual identificamos con nombre y etiqueta. También pensamos en la regla que establece que la unidad sea colocada en esa categoría.

Las categorías surgirán más rápidamente si primero leemos todo el material (unidades) y nos familiarizamos con éste.

No debemos olvidarnos de tomar notas en el diario de análisis, sobre las características de las unidades que las hacen similares o diferentes. El número de categorías crece conforme revisamos más unidades de análisis. Desde luego, al principio de la comparación entre unidades se crean varias categorías; pero conforme avanzamos hacia el final, el ritmo de generación de nuevas categorías disminuye.

La comparación entre unidades como base para la creación de categorías puede ilustrarse de la siguiente manera:

El investigador analiza la primera unidad:

1. Carolina: Mi esposo me ha golpeado varias veces (jehh!) (pausa).

Analiza la unidad y la asigna a una categoría que denomina: “violencia física”. La regla es: cualquier acto del marido que implique golpes con cualquier parte del cuerpo u objetos materiales (podría agregarse o no –según el juicio del investigador–: que genere hematomas o heridas en el cuerpo o incluso disparos de arma de fuego).

Analiza la segunda unidad:

2. No sé cómo decirlo. Me pega con la mano abierta y con el puño.

Deduce si es similar o no a la anterior (comparación constante). En este caso, “sí” lo es. Cae en la misma categoría, porque aplica la regla.

Analiza la tercera unidad.

3. La última vez me dijo: “Eres una ramera”. También me ha

Una vez más la compara con las otras dos unidades anteriores y deduce si es o no similar. No lo es, no sigue la misma regla. Entonces la asigna a otra categoría distinta, a la cual denomina “violencia verbal”, y establece la regla: Cada vez que el marido pronuncia “groserías” o “palabras que son consideradas culturalmente impropias” o, como se dice en algunas partes de América Latina, “malas palabras”, dirigidas a denigrar o insultar a su pareja.

Analiza la cuarta unidad:

4. dicho que soy malnacida, perra. Siempre me insulta. Y la

Vuelve a comparar y encuentra que es diferente a la primera y segunda unidades, pero similar a la tercera; la asigna a la categoría “violencia verbal”.

Analiza la quinta unidad de análisis:

5. verdad es que nunca he dado motivo. Nunca (pausa). Me dice que

No es similar a ninguna unidad; debe crearse otra categoría, pero si el análisis está dirigido a describir los tipos de violencia utilizados por el marido, tal unidad no es pertinente para generar categorías. Sin embargo, si el análisis pretende evaluar, además de los tipos de violencia presentes en las interacciones, el contexto en que se dan y la atribución de la esposa respecto de las razones por las cuales los maridos abusan de ellas, habría que crear una categoría y su regla (por ejemplo: "desconocimiento de la razón o motivo", cuando la mujer no expresa una razón o manifiesta no conocerla).

Y así seguiríamos con cada unidad de análisis, comparándola con las demás.

En algunas ocasiones, las unidades de análisis o significado no pueden ser ubicadas de manera clara en las categorías desarrolladas o descubiertas durante la revisión del material. Entonces se acostumbra crear la categoría "otras" ("varios", "miscelánea"...). Estas unidades son colocadas en dicha categoría, junto con otras difíciles de clasificar. Tal como señala Grinnell (1997), debemos tomar nota de la razón por la cual no son ubicadas en ninguna categoría establecida. Es posible que más adelante, al revisar otras unidades de análisis, generemos una nueva categoría en la que tengan cabida dos o más unidades que fueron asignadas a la categoría "otras". Al terminar de revisar todas las unidades, será conveniente revisar dicha categoría miscelánea y evaluar qué unidades habrán de juntarse en nuevas categorías. Cabe señalar que si una unidad de análisis no puede clasificarse en el sistema de categorías, no debe desecharse, sino agregarse a la categoría miscelánea.

Cuando nos encontramos que la categoría "otras" incluye demasiadas unidades de significado, resulta recomendable volver a revisar el proceso, y asegurarnos de que nuestro esquema de categorías y las reglas establecidas para clasificar sean claras y nos permitan discernir entre categorías. Grinnell (1997) sugiere que la categoría "otras" no puede ser mayor de 10% respecto del conjunto total del material analizado.

Una vez que hemos terminado de colocar las unidades en las categorías y evaluado la categoría miscelánea, es muy importante obtener evidencia acerca de la confiabilidad y validez del proceso de codificación. Esto puede hacerse proporcionando el material y el esquema de categorías a varios codificadores. Así, los codificadores aplican individualmente el sistema de clasificación a las unidades de análisis del material recolectado. Si notamos serias diferencias entre la codificación (asignación de unidades a categorías) hecha por distintas personas, es momento de volver a revisar el esquema de categorías, las reglas y los resulta-

Bitácora de análisis: valioso instrumento para lograr validez y confiabilidad del análisis y que se escribe diariamente (anotando la fecha); cada investigador tiene su propio sistema para llevarla.

dos. Como estamos hablando de análisis cualitativo, no hay un índice de objetividad entre codificadores que nos permita decir “hasta este punto debemos reconsiderar el sistema clasificatorio, las reglas y los procedimientos”. Pero sobre la base de la experiencia, no debe tolerarse una diferencia mayor de 10% en la clasificación efectuada por dos codificadores, ni mayor a 15% entre tres o más codificadores. Desde luego, esto es un asunto subjetivo y cada caso es diferente, por lo que dependerá del juicio del investigador.

Si hay diferencias, es probable que las reglas de clasificación o categorización no sean lo suficientemente claras o precisas. Pero si decidimos modificarlas, debe haber una razón y una justificación. Al hacerlo, habrá que anotar en la bitácora los motivos por los cuales cambian.

El principio es muy sencillo: si otras personas con las mismas reglas identifican unidades de significado asociadas con categorías diferentes a las que nosotros asociamos, ello indica que el sistema es más bien inconsistente o incongruente. Aunque la investigación cualitativa es más libre, ello no implica que no haya trabajo sistemático ni que cada quien interprete como se le ocurra el significado de las categorías. Análisis cualitativo no es “análisis caótico” ni “mi verdad es absoluta” (ejemplo: estaría equivocado si la línea “Carolina: Mi esposo me ha golpeado varias veces (¡ehh!) (pausa)”, se clasifica como “violencia verbal”. Ciertamente grado de subjetividad es tolerable e incluso deseable en el análisis cualitativo, mientras haya riqueza interpretativa, pero no subjetividad irracional. Ahora bien, independientemente de ello, el investigador debe reportar las reglas y el esquema de clasificación que utilizó, para que otros investigadores hagan su propio juicio al respecto. Más que un asunto ontológico (de lo que es real), se trata de un asunto metodológico.

El número de categorías que encontremos o generemos depende de lo que estemos buscando clasificar, el tipo de material revisado y los tópicos a tratar.

Cabe señalar también que es posible generar un sistema de categorías de múltiples dimensiones y que las unidades puedan ser ubicadas en varias categorías (una por dimensión). Por ejemplo, en el caso tratado del abuso de los maridos sobre sus esposas, ubicar las unidades en la dimensión “tipo de violencia” y, asimismo, en la dimensión “rol del esposo”.

Entonces una unidad se ubicaría en dos categorías, una referente a cada dimensión. Por ende, la unidad:

25. Carolina: “Me dijo que era una estúpida y que él manda y sólo él habla en esta casa”,

puede clasificarse como “violencia verbal” (categoría de la dimensión *tipo de violencia*) y como “autocrático o impositivo” (categoría *rol del esposo en la relación*). Podríamos agregar una tercera dimensión: “abuso de estupefacientes o be-

bidas alcohólicas”. Y de igual manera, clasificaríamos a las unidades en categorías de esta dimensión.

Desde luego, algunas unidades **no** podrían ser colocadas en ninguna categoría de una dimensión. Por ejemplo, la unidad **“Me dijo que era una estúpida y que él manda y sólo él habla en esta casa”** no tiene que ver con la última dimensión (“abuso de estupefacientes o bebidas alcohólicas”) y suele ubicarse en una categoría miscelánea denominada “no aplica” (no se ajusta a ninguna categoría de esta dimensión, porque el contenido no tiene que ver con ella).

Pero la unidad: Carolina: **“Me golpeó con un bastón, diciéndome que él manda, sólo él, nada más. Estaba ebrio”**, podría ser clasificada como “violencia física” (*tipo de violencia*), “autocrático o impositivo” (*rol del esposo en la relación*) y “presencia de abuso de alcohol” (*abuso de estupefacientes o bebidas alcohólicas*).

Una vez más, las unidades de análisis pueden ubicarse dentro de diferentes categorías en distintas dimensiones. Un ejemplo burdo e ilustrativo de lo que estamos comentando sería clasificar a las mujeres (unidades de análisis) en varias dimensiones: “color de pelo” (albino, rubio, castaño claro, etcétera), “estatura” (muy alta, alta, media, chaparra o bajita, enana), “por su complexión” (obesa, gorda, delgada) y “edad” (anciana, avanzada, madura, adulta joven, joven, muy joven, adolescente, niña). Así, Talía sería ubicada como de pelo negro, bajita, delgada y muy joven; Danea, como de pelo castaño oscuro, muy alta, de complexión estándar y adolescente; Mónica, como de pelo negro, estatura media, de complexión estándar y joven.

Con frecuencia, las categorías de una misma dimensión son mutuamente excluyentes (Paola es morena o es rubia, pero no las dos cuestiones a la vez). En otros casos, las unidades de análisis pueden caer en dos o más categorías de una misma dimensión (ser mutuamente excluyentes las categorías; pero no la clasificación de las unidades dentro de éstas).

La frase **“Me golpeó y me dijo que era una basura y me encerró en el cuarto”** puede categorizarse como “violencia física” (...golpeó), “violencia verbal” (...era una basura) y “violencia psicológica” (...me encerró).

Un ejemplo de clasificación de categorías lo ofrece Morse (1999), obtenido al observar el proceso de confortación que ofrecían enfermeras a pacientes traumatizados (en estado de gravedad), en la sala de emergencia de los hospitales (que fue esbozado en el capítulo anterior). Para la dimensión “estados conductuales de los pacientes”, encontró las siguientes categorías:

- ◆ Inconsciente
- ◆ Tranquilo y relajado
- ◆ Asustado
- ◆ Aterrorizado
- ◆ Fuera de control

Estas categorías son mutuamente excluyentes y reflejan el estado del paciente durante la emergencia.

Otra dimensión que se pudo derivar es: “estrategias de confortación”:

1. Hablarles a los pacientes en situaciones dolorosas.
2. Permitirles soportar la agonía, empleando un estilo particular de conversación y posturas que denominamos registro de conversación para confortar.
3. Normalizar la situación al prevenir los gritos, la excitación y el pánico, lo mismo que controlar la propia expresión mientras se atendían las lesiones.
4. Bromear con los pacientes en condiciones serias, de tal forma que la situación no pareciera grave.
5. Apoyar a los médicos en sus tareas y recordarles cuánto tiempo había pasado desde que comenzaron los esfuerzos de resucitación, cuándo era tiempo de mover a los pacientes, si era necesario darles más analgésicos...
6. Llevar a los familiares de los pacientes, lo cual implicaba ocultar cualquier signo de severidad del padecimiento (limpiar la sangre), describir a la gente lo que debería hacer al entrar a la sala de traumatología.
7. Apoyar a los parientes y explicarles cómo hablarles a las personas queridas.

La creación de categorías, a partir del análisis de unidades de contenido, es una muestra clara de por qué el enfoque cualitativo es esencialmente inductivo.

Los nombres de las categorías y las reglas de clasificación deben ser lo suficientemente claras para evitar reprocesos en la codificación; por ejemplo, podría haber la dimensión de análisis “Autoestima” (con las categorías “alta”, “mediana”, “baja”, “muy baja”) y la unidad: 128. **“al morir mi madre me sentí devaluada, mi autoconcepto no es bueno”** (un asunto es la autoestima y otro la autoevaluación; ambas se relacionan, pero no son lo mismo). Entonces la dimensión y las categorías que aplican a la unidad se refieren más bien a la autoevaluación que a la autoestima; la primera refleja más el significado de la unidad.




Debemos recordar que en el análisis cualitativo hay que reflejar lo que nos dicen las personas estudiadas en sus “propias palabras”.

En ocasiones, las reglas de clasificación son complejas; pero tenemos que tratar de que sean lo más precisas posible. Por ejemplo, para clasificar las películas que se exhiben en los cines de algunos países latinoamericanos, se siguen varios criterios. Si las cintas muestran:

- Lenguaje soco y procaz
- Violencia física, psicológica y verbal
- Escenas de violación, parricidio, incesto, sexo explícito
- Abuso de alcohol o drogas (escenas de consumo)

- Homicidio doloso o intencional
- Agresión con premeditación, alevosía y ventaja; no se autorizan para ser vistas por menores de edad.

Por otra parte, el investigador debe asignar códigos a las categorías y etiquetarlas, con la finalidad de que el análisis sea más manejable y sencillo de realizar. Los códigos son números, letras, símbolos, palabras, abreviaturas, imágenes o cualquier tipo de identificador (Baptiste, 2001); por ejemplo,

<i>Tipo de violencia</i>	
1: Violencia física	VF: Violencia física
2: Violencia verbal	VV: Violencia verbal
3: Violencia psicológica	VP: Violencia psicológica
	 : Violencia física
	 : Violencia verbal
	 : Violencia psicológica

Los códigos identifican a las categorías, y permiten asignarles un código que indique dimensión y categoría. Por ejemplo:

TVF: Violencia física
TVV: Violencia verbal
TVP: Violencia psicológica

donde *T* nos señala que es la dimensión “Tipo de violencia” y *VF*, *VV* y *VP* las diferentes categorías.

A veces tenemos codificaciones más complejas:

- PASJE-** Pariente que abusó sexualmente de la joven y que estaba bajo el influjo de un estupefaciente.
- PASJA-** Pariente que abusó sexualmente de la joven y que estaba bajo el influjo del alcohol.

- PASJNS- Pariente que abusó sexualmente de la joven y que no estaba bajo el influjo de ninguna sustancia.
- PASP- Profesor autocrático que sanciona a los alumnos con permanencia después de que terminaron las clases.
- PASR- Profesor autocrático que sanciona a los alumnos ridiculizándolos en público.
- PDNS- Profesor democrático que no ejerce sanción.

Hay quienes separan la secuencia; por ejemplo, P-A-S-J-E, etcétera.

Cuando las categorías son personas (por ejemplo, al analizar relaciones entre miembros de una familia o un grupo de pandilleros), suelen asignarse como códigos las siglas de cada quien (MMG-Mónica Morales Gálvez). También suelen identificarse secuencias de acción a través de códigos (E VF E A-El Esposo abusó mediante Violencia Física de la Esposa bajo el influjo del Alcohol).

Las codificaciones son como “apodos” o “sobrenombres” de las categorías. Permiten que sean identificadas más rápidamente.

Algunos programas computacionales para análisis cualitativo sirven para manejar códigos por categoría de hasta 12 dígitos; por ejemplo, The Ethnograph (“El Etnógrafo”).

Una vez categorizadas todas las unidades, se revisan las reglas de clasificación. Así, nos damos cuenta de si estamos o no capturando el significado que buscan transmitir los sujetos investigados o que las categorías que anticipamos no se presentan; de ser así, es momento de volver al campo e incrementar el número de entrevistas, observaciones, sesiones de grupo, etcétera.

Ante problemas e inconsistencias, habrá que revisar dónde estamos fallando: si en la definición de la unidad de análisis, en las reglas de categorización o clasificación, en las categorías que se van estableciendo por comparación de unidades o en los nombres de las categorías; de igual forma, hay que corregir o volver al campo, y reportarlo todo en la bitácora de análisis.

Si todo va bien, los datos se reducirán a categorías y se transformarán, sin perder su significado (lo cual es imprescindible en la investigación cualitativa), además de encontrarse codificados. A veces tenemos como resultado unas cuantas categorías, y en otras diversidad de dimensiones y categorías. Todo depende del planteamiento del problema de investigación, el método de recolección de los datos, la cantidad de material obtenido y los tópicos de la investigación.

4. Codificar los datos en un segundo nivel o plano

La codificación de los datos en un segundo plano implica ir refinando la codificación e involucra la interpretación del significado de las categorías obtenidas en el primer nivel (Grinnell, 1997). Berg (1995) sugiere que el investigador genere

ejemplos detallados de las transcripciones para apoyar cada interpretación. Esto es, que ejemplifiquemos la forma como categorizamos.

Entonces comienza la tarea esencial de la codificación en un segundo plano: identificar diferencias y similitudes entre categorías. Una vez más, utilizamos la comparación constante, sólo que aquí no equiparamos unidades, sino categorías.

En este punto del análisis, la meta es integrar las categorías en temas y subtemas, basándonos en sus propiedades. Descubrir temas implica localizar los patrones que aparecen repetidamente entre las categorías. Cada tema que se identifica recibe un código (como lo hacíamos con las categorías). Los temas son la base de las conclusiones que emergerán del análisis. Además, podemos encontrar temas y subtemas.

Grinnell (1997, p.519) ejemplifica la construcción de temas con las categorías: “Asuntos relacionados con la custodia de los hijos”, “procedimientos legales de separación o divorcio” y “obtención de órdenes restrictivas”, las cuales pueden constituir un tema (categoría más general): “Asuntos relacionados con el sistema legal”.

En el estudio sobre la moda y la mujer mexicana, se les preguntó a las participantes de los grupos de enfoque sobre los factores que intervieran para elegir su tienda favorita de ropa. Surgieron, entre otras, las categorías “variedad de modelos”, “surtido”, “existencia de ropa”, “mucho ropa”, “abundancia de modelos”. Tales categorías se agruparon en el tema “abundancia”. “Precio”, “promociones” y “ofertas” fueron categorías que se integraron en el tema “economía”. “Calidad”, “buenos artículos o ropa” y “productos bien hechos” se incluyeron en el tema “calidad de producto”, lo mismo ocurrió con otras categorías.

Hemos codificado el material en un primer plano (encontrar categorías, evaluando las unidades de análisis y ubicándolas en las categorías mediante reglas, además de asignarles un código a cada categoría) y en un segundo plano (encontrar temas o categorías más generales). Estamos listos para la interpretación.

5. Interpretar los datos

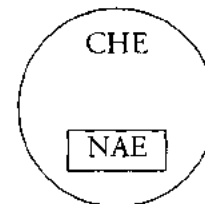
En el análisis cualitativo resulta fundamental darle sentido a:

- a) Las descripciones de cada categoría. Esto implica ofrecer una descripción completa de cada categoría y ubicarla en el fenómeno que estudiamos. Por ejemplo, La “violencia física” por parte del esposo, ¿cómo es?, ¿cuánto dura?, ¿en qué circunstancias se manifiesta?
- b) Los significados de cada categoría. Ello quiere decir analizar el significado de la categoría para los sujetos. ¿Qué significado tiene la “violencia física”

para cada esposa que la padece, y que nos la narra en una entrevista o una sesión de grupo (en sus propias palabras y de acuerdo con el contexto)? ¿Qué significado tiene para tales mujeres ver al marido en estado de ebriedad? ¿Qué significado tiene cada palabra soez que escuchan de los labios de su cónyuge?

- c) La presencia de cada categoría. La frecuencia con la cual aparece en los materiales analizados (cierto sentido cuantitativo). ¿Qué tanto es mencionada cada categoría por los sujetos? La mayoría de los programas de análisis cualitativo efectúa un conteo de categorías, frases y palabras, además de expresarlo en porcentajes. Por ejemplo, es interesante conocer cuál es la palabra con la que nombran al esposo más frecuentemente o se refieren a él y ¿qué significado tienen las designaciones más comunes?
- d) Las relaciones entre categorías. Encontrar vinculaciones, nexos y asociaciones entre categorías. Algunas relaciones comunes entre categorías son:
 - Temporales: Cuando una categoría siempre o casi siempre precede a otra. Por ejemplo, VF → E/A (cuando hay Violencia Física del esposo hacia su pareja existe generalmente abuso de Estupefacientes o Alcohol).
 - Causales: Cuando una categoría es la causa de otra. Por ejemplo, MNCE → AM (las Mujeres que No Contactan a sus Esposos después de que se han separado como consecuencia de la violencia física generalmente se Autoevalúan Mejor).
 - De conjunto-subconjunto: Cuando una categoría está contenida dentro de otra. Por ejemplo,

CHE-Chantaje del Esposo a la mujer para que regrese a la vivienda
 NAE-Negativa de Apoyo Económico del esposo para que regrese la mujer a la vivienda.



Para ayudarnos a establecer relaciones entre categorías, podemos dibujar diagramas causales, elaborar matrices y cuadros, sobre la base de nuestra lectura de los datos y las categorías descubiertas. Un ejemplo lo vemos en la figura no. 10.17

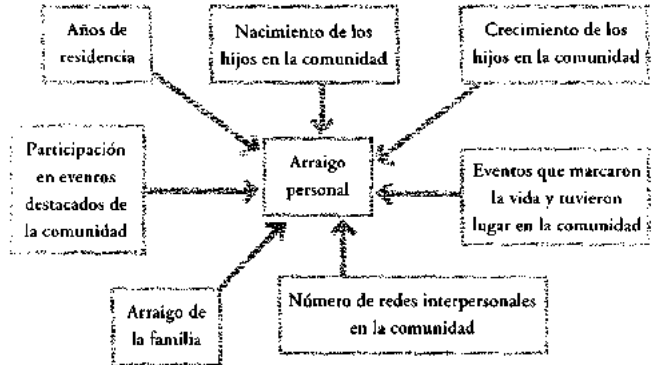
En las celdas de la matriz se anotan las interacciones entre categorías. Los diagramas causales, las matrices y los esquemas también se producen en el análisis cuantitativo, solamente que en este último eran producto de análisis estadístico, en tanto que en el análisis cualitativo se producen sobre la base de la codificación y, sobre todo, de la reflexión y los juicios del investigador. Por ello,

Ejemplos de establecimiento de relaciones entre categorías de manera gráfica

Figura 10.17

EJEMPLO 1: DIAGRAMA O MAPA CONCEPTUAL

Factores relacionados con el arraigo personal a una comunidad.



EJEMPLO 2: MATRIZ

Categorías de los padres/ Categorías de los hijos	Padres adictos al consumo de drogas	Padres adictos al consumo de alcohol	Padres divorciados	Ausencia del padre	Ausencia de la madre
Tendencia a ejercer la prostitución					
Consumo de drogas					
Consumo de alcohol					
Vagancia, pertenencia a pandillas juveniles					
Abandono de la educación formal					

resulta fundamental la profundidad con la que se codifique, revise y evalúe el material para llevar a cabo un análisis completo y exhaustivo. Por ejemplo, los mapas no se obtienen de aplicar un análisis por conglomerados o un análisis multivariado de varianza, sino a través de construir el modelo analizando los resultados de la codificación y los datos en sí.

Asimismo, es posible agregar a nuestro análisis el material adicional que recolectamos en el campo como fotografías, dibujos, artefactos (si estudiamos a un grupo de pandillas, podemos incluir piezas de su vestuario, armas, accesorios, etcétera; es muy común en la investigación policiaca cuando se analiza la escena del crimen), escritos (no las transcripciones, sino, por ejemplo, anotaciones de los sujetos —en servilletas, notas suicidas, diarios personales, etcétera—) y otros materiales. En ocasiones, estas piezas son material de análisis en sí; pero otras veces son elementos adicionales complementarios para la labor de análisis.

Utilizar metáforas ha sido una herramienta muy valiosa para extraer significados o captar la esencia de relaciones entre categorías. Muchas veces estas metáforas surgen de los mismos sujetos estudiados o del investigador. Son los casos de “Eres un paracaídas cuando la tormenta arrecia” (en una relación romántica nos sirve para establecer el tipo de vínculo entre la pareja), “Con ése no juego ni a las canicas” (desconfianza), “Eres el típico jefe que manda bajo la técnica del limón exprimido” (una manera de decir: cuando obtienes todo lo que quieres de un subordinado, cuando ya lo exprimiste y se le acabó el jugo, lo desechas, ya no te es útil). “Todos los caminos llevan a Roma” (diversas alternativas conducen a lo mismo), “Cuanto más negra es la noche significa que pronto va a amanecer”, “Siempre habrá un mañana”, etcétera.

6. Describir contexto(s), eventos, situaciones y personas sujetos de estudio; encontrar patrones; explicar sucesos, hechos, y construir teoría.

Una vez que hemos descrito las categorías y los temas, así como su frecuencia, significado y relaciones, ayudándonos de esquemas gráficos, debemos enmarcar el contexto, la situación o el evento en el cual ocurren.

¿Dónde suceden los hechos observados o registrados? ¿Qué características tiene el contexto? ¿En qué situaciones se relacionan las categorías? Por ejemplo, la violencia física expresada en el abuso de los esposos sobre sus mujeres tiene lugar en un barrio sumamente pobre de una comunidad de pescadores de Tekendama. *Las tradiciones son ancestrales y consisten en... El papel de los hombres es... Los matrimonios se dan así y... Los consumos de alcohol son... Las mujeres cuando se embarazan...* Aquí, por cuestiones de espacio, sólo damos unas pautas,

pero las descripciones del contexto deben ser completas y profundas, además de estar vinculadas por el investigador con las relaciones entre categorías (que podríamos denominar variables, aunque algunos fundamentalistas cualitativos no estarán de acuerdo).

También habrá que describir a los sujetos del estudio, los actores del contexto. En el caso de las mujeres golpeadas, describiríamos cómo viven, cuál es su ideología, su manera de ver el mundo, sus costumbres, sus creencias, sus características sociodemográficas (estado civil, edad, ocupación, ingresos, etcétera); en fin, la mayor cantidad de datos.

La calidad del análisis cualitativo depende del grado en que profundicemos en tales aspectos y los vinculemos con las categorías encontradas.

En esta labor, las observaciones, las notas de campo y la bitácora que escribimos o dictamos durante la recolección de los datos se convierten en herramientas muy valiosas. La razón: si no fuimos cuidadosos, perderemos mucha información, pues con el tiempo las memorias y los recuerdos se van perdiendo.

Ejemplos de cómo debe describirse un contexto son los trabajos de diversos etnólogos y antropólogos como Howard Carter (1977) y George Foster (1987); el primero describe el contexto del descubrimiento de la tumba de Tutankhamón en Egipto y el segundo la vida en una aldea mexicana: Tzintzuntzán. Pero además de estos ejemplos podemos encontrar muchos en colegas de acuerdo con sus disciplinas.

Asimismo, en este punto mediante una seria introspección e integración de materiales para el análisis (mapas, esquemas, etcétera), las descripciones del contexto, los actores y la revisión de notas, apuntes, bitácoras, etcétera; comenzamos a encontrar “patrones”, “tendencias”, “explicaciones”; a ligar hechos, generar hipótesis y/o probarlas y formular nuevas preguntas de investigación; lo más importante, a construir teoría.

En este sentido, las teorías son explicaciones de lo que hemos vivido, observado, analizado y evaluado en profundidad. Es como se conoce en el lenguaje del enfoque cualitativo: *Grounded theory* (teoría fundamentada). La teoría la obtenemos no de aplicar técnicas estadísticas, sino sobre la base de profundizar experiencias.

Baptiste (2001) expresa que los estudios cualitativos deben ir más allá de simples glosarios de categorías o temas y descripciones (lo cual es útil, pero insuficiente); tienen que proporcionar un sentido de entendimiento profundo.

En ocasiones, el análisis nos produce conclusiones sencillas y, en otras, descubrimientos complejos. A continuación mostramos algunos segmentos de conclusiones de análisis cualitativos.

Descripción del contexto: actividad en la cual se retoman las anotaciones, los memos y la bitácora de campo, obtenidas durante la recolección de datos.

**Estudio sobre los obstáculos para la puesta en marcha de la tecnología
en el ámbito escolar**

Episodio o situación : Reunión comunidad educativa
25 de abril de 2002

Lugar: Primaria Pública General Simón Bolívar

1. Temas principales. Impresiones (del investigador). Resumen de lo que sucede en el evento, episodio, etcétera.

El director no apoya las mociones del Ministerio de Educación para integrar la tecnología al ámbito escolar mediante el impulso a centros tecnológicos en las instalaciones de la escuela. Piensa que el cambio obstaculizará la labor del docente, en vez de apoyarla. Desconfía de pasadas intervenciones del Ministerio, donde promeren muchas innovaciones y recursos y después "no sucede nada".

Profesores jóvenes están entusiasmados con la idea de centros tecnológicos. Piensan que sí ayudarán a la calidad educativa y a la mejor preparación de los jóvenes.

Tema recurrente: Se mejorarán oportunidades futuras para el estudiante. Se integrarán a un mundo más global. Director: Piensa en otros gastos

2. Explicaciones o especulaciones, hipótesis de lo que sucede en el lugar.

El director está en una etapa de retiro, no en una etapa de búsqueda. Quiere terminar su periodo tranquilamente; literalmente dijo "sin hacer olas". Piensa que el proyecto del Ministerio puede ser algo potencialmente peligroso y no deseable. Una situación que no se reflejará en su desempeño, sino que le creará más problemas. Proposición o hipótesis: la edad del director y su antigüedad en el puesto tendrán un impacto negativo en su grado de innovación o actitud hacia programas tecnológicos

3. Explicaciones alternativas. Reportes de otros que viven la situación.

Algunos docentes informan que el director tuvo una experiencia negativa con innovaciones tecnológicas en otra institución, donde fue saboteado por los docentes.

Su aparente "experiencia" está bloqueando la incorporación de la escuela a un mundo global. Segmentos jóvenes de la docencia muestran insatisfacción. Los jóvenes están temerosos de que su institución se vea rezagada.

4. Sigüientes pasos en la recolección de datos. Considerando lo anterior, qué otras preguntas o indagaciones hay que hacer.

Entrevista con él para confirmar percepciones. Indagar con colegas si la proposición es válida.

Entrevista profunda con directivos. Grupo de enfoque con docentes.

Tema: Discutir bondades y amenazas de la tecnología. Propiciar tormenta (lluvia) de ideas sobre percepción de otras necesidades de la institución. Analizar situaciones similares en la literatura sobre tecnologías emergentes.

5. Revisión, actualización. Implicaciones de las conclusiones.

Considerar si fuerzas jóvenes de las instituciones pueden contrarrestar efectos estabilizadores de directivos.

Considerar enlaces en las fases de implementación-análisis de las nuevas tecnologías en el ámbito escolar. Implicar dinámicas de grupo para cambio de actitudes...

ALGUNOS COMENTARIOS DE MUJERES MEXICANAS SOBRE LA MODA

MUJERES MAYORES DE 18 AÑOS

- La mayoría de los segmentos de todas las ciudades coinciden en que hablar de moda es muy relativo, pero señalaron que para ellas significa vestirse de acuerdo con su personalidad, buscando comodidad y usando los colores de temporada.
- Lo importante es que los diseños se adecuen a ellas y que se sientan “a gusto” con la ropa.
- Demandaron que la ropa se adecue a la complexión de las mujeres mexicanas, ya que las confeccionadas para damas “más llenitas” (gordas) opacan su belleza pues los colores son oscuros y no existen ni variedad ni buenos estilos.
- Respecto de las tallas, manifestaron no encontrar ropa acorde con su cuerpo, manifestando que en general “viene muy estrecha” y les ocasiona problemas en la zona de las cadera y las piernas. Asimismo, opinaron que el largo del pantalón no es suficiente en ocasiones.
- La percepción en los segmentos de mujeres jóvenes es que “almacenes XXXX no cuenta con marcas de moda”. Recomiendan incorporar marcas exclusivas como XXXX, dirigidas al público femenino joven, preocupado por estar a la moda.
- Solicitan que ellas mismas puedan formar sus coordinados y que hubiera tallas intercambiables.
- En el interior, recomendaron que en el departamento se disponga una área para que los niños se entretengan mientras ellas se prueban modelos y compran.

Jóvenes de 15 a 17 años.

- La mayoría de las jóvenes compran su ropa en tiendas juveniles (llamadas por los adultos *boutiques*, término que a muchas de ellas les hace gracia).
- Las tiendas preferidas son XXXX, XXXX y *boutiques* locales.
- En segundo término, acuden a tiendas departamentales, principalmente...
- Fundamentalmente compran por impulso, es decir, no planean sus compras.
- Solamente planean sus compras cuando tienen un evento social.
- Se gufan por sus sentidos al ver las prendas, más que por una marca.
- Acuden a las tiendas departamentales y si les gusta una prenda normalmente regresan con sus padres para adquirirla.

ESTRATEGIAS DE CONFORTACIÓN A PACIENTES TRAUMATIZADOS

Janice M. Morse (1999, p.15)

Las estrategias y el estilo de atención de las enfermeras deben ser apropiados al estado de los pacientes. Por ejemplo, si se emplea una estrategia incorrecta en el caso

de un paciente atemorizado más que uno aterrorizado, entonces su nivel de fortalecimiento habrá de aumentar. Si el estado del paciente se deteriora, o si no existe mejoría en diez segundos, la estrategia habrá de cambiarse de inmediato.

Una vez que los pacientes hayan obtenido un nivel tolerable de confortación, entonces ellos se sentirán seguros, confiarán en el personal y aceptarán la atención. Por ejemplo, en traumatología los enfermos que están en control o han aceptado la atención responden, son cooperadores y receptivos. A pesar de su dolor, tratan de salir adelante. Un paciente que ha mejorado por completo se percata de que el cuidado es necesario y acepta cualquier medida que se requiera. El resultado es que la atención se da en forma más rápida y segura.

Frecuentemente los comentarios son apoyados por texto, diagramas, fotografías o algún otro elemento. Por ejemplo:

INVESTIGACIÓN SOBRE CENTROS COMERCIALES

Están de acuerdo con que los centros comerciales son como los zócalos de antes en donde la gente va a ver y a ser vista; **“son los centros de reunión entre jóvenes para conocerse”**; **“también los adultos, al exhibirse, sentirnos un rato a gusto; a lo mejor es importante andar entre gente de muy diversa forma de ser, de vestir; inclusive, uno algunas veces copia modas”** (en negritas comentarios textuales de participantes a un grupo de enfoque).

Desde luego, son tan sólo segmentos de ejemplos, pues los comentarios suelen ser mucho más abundantes.

7. Asegurar la confiabilidad y validez de los resultados

Con el propósito de respaldar nuestras investigaciones, es importante asegurar la confiabilidad y validez de nuestros análisis, ante nosotros mismos y ante los usuarios del estudio.

A diferencia del análisis cuantitativo, en el cual aplicamos métodos estadísticos que determinan la confiabilidad y validez del estudio, en el análisis cualitativo no se utilizan coeficientes ni pruebas. Más bien, habrá que efectuar una valoración del proceso de análisis.

En primer término debemos evaluar si obtuvimos suficiente información de acuerdo con nuestro planteamiento del problema. Es decir, si en la recolección de los datos (entrevistas, sesiones, documentos, biografías, etcétera) incluimos

los casos que nos interesaban —de acuerdo con perfiles— y si profundizamos hasta donde era posible. Esto último resulta subjetivo, pero mencionaremos que la profundidad mínima en una investigación cualitativa consiste en haber realizado las descripciones detalladas del contexto, los sujetos y el fenómeno analizado, así como descubrir los significados de las categorías.

En segundo lugar, es recomendable llevar a cabo un ejercicio de triangulación del análisis. Lo anterior implica que otros(as) investigadores(as) utilicen nuestras reglas para ver si toman las mismas decisiones (o similares) respecto de las unidades, las categorías y los temas de análisis (Grinnell, 1997). Esto no significa que si no hay una réplica del proceso de codificación, los resultados que obtuvimos no serán confiables ni válidos; pero sí implica que las categorías principales los patrones generales, los temas y las conclusiones más importantes deben coincidir en lo fundamental. No estamos hablando de que intenten replicar el estudio, sino de que con nuestros mismos datos el análisis llevado a cabo por dos o más investigadores arroje conclusiones similares en lo esencial. Desde luego, en el enfoque cualitativo la libertad para interpretar es mayor que en el cuantitativo.

Un tercer elemento para asegurar la confiabilidad y validez de los resultados (que no siempre es posible aplicar), consiste en obtener retroalimentación directa de los sujetos de la investigación (al menos una muestra de ellos), lo cual significa pedirles que confirmen o refuten nuestras interpretaciones y ver si capturamos los significados que ellos trataban de transmitir, así como si nuestras descripciones fueron lo suficientemente completas y profundas.

Baptiste (2001) plantea una serie de preguntas de cara al análisis cualitativo, de las cuales retomaremos las siguientes en un ejercicio de autoanálisis, con la finalidad de asegurar la confiabilidad y validez:

- a) ¿Nos familiarizamos a fondo con todos los datos recolectados (transcripciones, notas, bitácoras, etcétera)?
La respuesta, a conciencia, debe ser un “sí”. La contundencia de esta respuesta es en parte medida de la confiabilidad y validez.
- b) ¿Codificamos todos los datos (unidades) o sólo aquellos que nos parecieron directamente relacionados con nuestras preguntas de investigación?
La respuesta debe ser que codificamos todos los datos, ya que de lo contrario pudimos haber dejado material importante. Después de la codificación es posible eliminar la trivía, no antes.
- c) ¿Utilicé diversas herramientas de análisis?
Es decir, no me limité a puras descripciones ni a la codificación, sino que elaboré diagramas, mapas, matrices, etcétera; y analicé materiales complementarios. La respuesta más conveniente será la positiva.
- d) ¿Validé con otros investigadores (o dentro del equipo de investigación) los criterios y las reglas utilizadas para codificar?

- e) ¿Las categorías incluyeron de manera exhaustiva los datos más sobresalientes?
- f) ¿Las definiciones de las categorías son tan claras como para distinguir estas últimas entre sí, así como para asignar adecuadamente las unidades a las categorías que les corresponden y apoyar los propósitos del estudio?
- g) ¿Cada categoría está adecuadamente respaldada por datos codificados?
En estas últimas preguntas, la respuesta debe ser afirmativa, es decir, “sí”.
- h) ¿De qué modo las unidades codificadas y agrupadas en una misma categoría son similares y diferentes?
Esto, con fines de autoanálisis.
- i) ¿La bitácora y las notas de campo son lo suficientemente explicativas?
- j) ¿La bitácora de análisis es muy explicativa?
Ambas preguntas requieren una respuesta positiva.

8. Retroalimentar, corregir y regresar al campo

Una vez que hemos completado y evaluado el análisis, nos encontraremos con que éste ha cumplido con los objetivos y, en general, con nuestras expectativas sobre el estudio. Es decir, las conclusiones responden al planteamiento del problema inicial o nos ayudaron a modificarlo, pero en aras del avance del conocimiento.

Desde luego, también llega a ocurrir que durante el análisis nos demos cuenta de que no estamos alcanzando nuestros objetivos ni respondiendo a nuestras inquietudes. Entonces, en esos momentos debemos hacer “un alto en el camino” (metáfora), pausar y evaluar qué nos hace falta, dónde no hemos procedido adecuadamente, por qué no logramos alcanzar nuestras metas o qué obstáculos hemos tenido. Es recomendable hacerlo en conjunto con todo el equipo de investigación.

Como resultado de la evaluación a la obtención de retroalimentación (que también puede ampliarse consultando a otros colegas), detectaremos deficiencias, errores, falta de datos, etcétera; ello implica desde realizar más entrevistas, sesiones, estudios de caso u observaciones hasta la necesidad de recavar más materiales, recodificar, agregar nuevos esquemas o elaborar otros análisis. Lo anterior no debe preocuparnos, pero siempre y cuando hayamos sido cuidadosos en la recolección y el análisis de los datos. Tal vez el fenómeno sea tan complejo que requiere de nuestro regreso al campo por lo menos una vez. De hecho, la obtención de retroalimentación tiene que durar todo el análisis.

Planteamiento del problema

En todo el proceso de análisis debemos tener en mente el planteamiento original del problema de investigación, no para “poner una camisa de fuerza” a nuestro análisis, sino con la finalidad de que no se nos olvide encontrar las respuestas que

buscamos. Asimismo, recordemos que dicho planteamiento puede sufrir cambios o ajustes conforme avanza la investigación. Modificaciones en el planteamiento del problema, habrán de justificarse.

Programas computacionales para análisis cualitativo

En la actualidad han sido desarrollados diferentes programas —además de los procesadores de textos— que sirven de auxiliares en el análisis cualitativo. De ninguna manera sustituyen el análisis creativo y profundo del investigador. Simplemente facilitan su tarea.

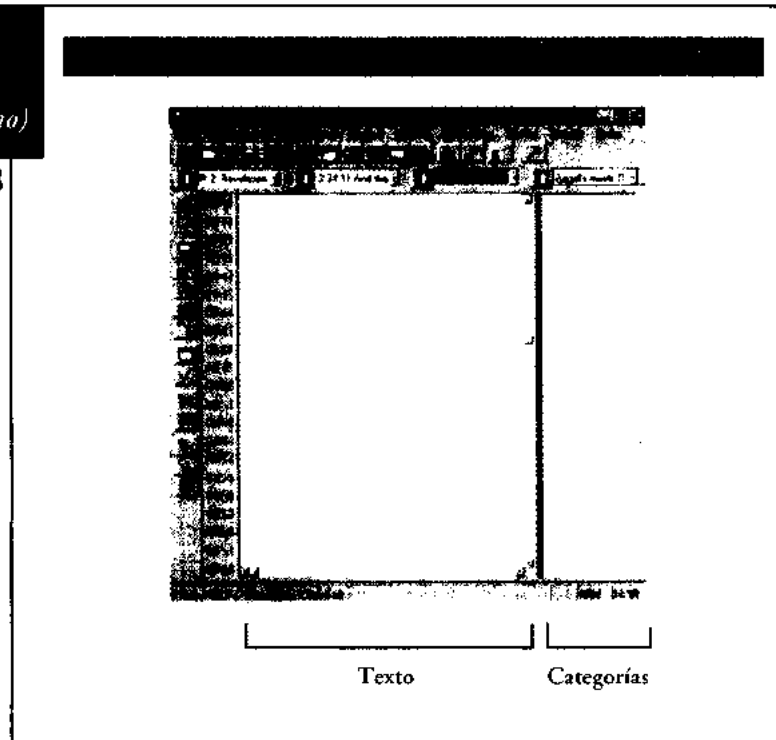
Algunos de los nombres de programas que más se escuchan en el análisis cualitativo (Fielding, 1993; Weitzman y Miles, 1995, Baptiste, 2001) son:

1. Atlas/ti®

Es un excelente programa desarrollado en la Universidad Técnica de Berlín por Thomas Muhr para codificar datos y construir teoría fundamentada. Uno introduce los datos (que pueden ser textos, fotografías, segmentos de audio o video, diagramas, mapas y matrices) y el programa los codifica automáticamente, de acuerdo con el esquema que hayamos diseñado. Las reglas de codificación las

*Ejemplo en
pantalla del
Atlas/ti (demo)*

Figura 10.18



establece el investigador y el programa las aplica. Realiza conteo (frecuencia de aparición de palabras, unidades, categorías y temas; en términos absolutos y relativos o porcentajes) y vincula las categorías (análisis de relaciones entre éstas o redes entre categorías, incluso vinculaciones entre textos). Ofrece diversas perspectivas o vistas de los análisis (diagramas, datos por separado, etcétera).

2. **The Ethnograph®**

Es un programa muy popular para identificar y recuperar textos de documentos. La unidad básica es el segmento. Asimismo, codifica las unidades partiendo del esquema de categorización que haya establecido el investigador. Los segmentos pueden ser “anidados”, entrelazados y traslapados en varios niveles de profundidad. Las búsquedas llegan a efectuarse sobre la base de códigos expresados en un carácter, una palabra o en palabras múltiples. Los esquemas de codificación suelen modificarse. Guarda memos, notas y comentarios. También los incorpora al análisis.

3. **NUD*IST®**

Un excelente programa de análisis útil para construir grandes bases de datos estructuradas jerárquicamente, que puede agregar documentos para ser analizados. También, al igual que los dos anteriores, codifica unidades de contenido (texto y otros materiales), basándose en el esquema diseñado por el investigador. Localiza texto por carácter, palabra, frase, tema o patrón de palabras; incluso, por hojas de cálculo de variables.

4. **SONAR®**

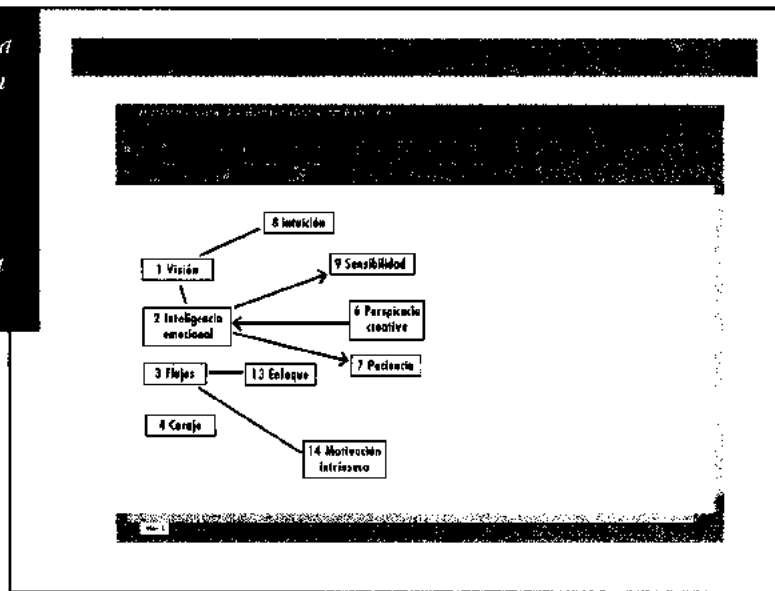
Un software que codifica unidades de contenido de texto, que localiza frases, y adjunta a la vista del texto, comentarios o notas. Puede desplegar varias vistas del texto simultáneamente. Vincula y agrupa categorías. Su capacidad de edición de documentos es notable.

5. **Decision Explorer®**

Este programa inglés resulta una excelente herramienta de mapeo de categorías. El investigador puede visualizar relaciones entre conceptos o categorías en diagramas. Como en todo programa, en ésta es el investigador quien introduce las categorías y define sus vinculaciones, Decision Explorer las muestra gráficamente. Asimismo, realiza conteo de la categoría con mayor número de relaciones con otras categorías. Cualquier idea la convertimos en concepto y la analizamos. Muy útil para visualizar hipótesis y la asociación entre los componentes más importantes de una teoría. Recomendable para análisis cualitativo de relaciones entre categorías (causal, temporal...). En la figura 10.19, se incluye un ejemplo de cómo Decision Explorer visualiza la relación entre conceptos y categorías establecidos(as) por el investigador.

Ejemplo de la visualización que hace Decision Explorer de un mapa o un diagrama causal

Figura 10.19



Existen otros programas, tales como HyperQual®, HyperResearch®, QUALPRO® y QUALOG® para fines similares.

Al igual que en el caso de los programas de análisis cuantitativo, el software cualitativo evoluciona con vertiginosidad (surgen nuevos programas, los actuales evolucionan y expanden sus posibilidades). Prácticamente todos sirven para las etapas del análisis: codificación en un primer plano, y en un segundo plano, interpretación de datos, descubrimiento de patrones y generación de teoría fundamentada; además de que nos ayudan a establecer hipótesis. Asimismo, todos recuperan y editan texto, lo mismo que numeran líneas o unidades de contenido. La tendencia es que logren incorporar todo tipo de material al análisis (texto, video, audio, esquemas, diagramas, mapas, fotografías, gráficas —cuantitativas y cualitativas—, etcétera).

La única limitación en estos momentos de los programas de análisis cualitativo es que imponen la capacidad y la memoria del disco duro o el disco compacto (CD).

ANÁLISIS CUANTITATIVO Y ANÁLISIS CUALITATIVO: CUESTIÓN DE ENFOQUES

Tanto el análisis cuantitativo (estadístico) como el análisis cualitativo son herramientas muy útiles y desarrolladas en manos del investigador. Ninguno es mejor que el otro, simplemente son diferentes caminos para llegar a un mismo fin:

lograr los objetivos de investigación y responder a las preguntas de estudio para generar conocimiento. Con frecuencia son complementarios; sin lugar a dudas, siempre que elegimos uno de ellos incorporamos elementos del otro. Por ejemplo, un histograma de barras requiere de una “interpretación” y comentarios cualitativos (ahondar en el significado de las cifras). Lo mismo un coeficiente de correlación, un valor F o una proporción. También, una observación usa conteo (cuántas veces ocurre) y en ocasiones análisis estadístico, aunque sea incipiente (tal sería el caso de un análisis cruzado de categorías que evaluara sus intersecciones: categorías de tipo de violencia con categorías de abuso de enervantes, una ji-cuadrada que nos aclarara si cierto abuso de estupefacientes está relacionado con determinado tipo de violencia). O bien, analizar de manera comparativa los porcentajes de las categorías que se presentan en textos autobiográficos o transcripciones de sesiones de enfoque.

Es obvio que los estudios cuantitativos centran sus análisis en métodos estadísticos, mientras que los cualitativos lo hacen en análisis etnológicos, en tanto que las modalidades que mezclan ambos enfoques realizan análisis cuantitativo y análisis cualitativo.

Por ejemplo, un experimento que usa un cuestionario estandarizado y cuantitativo para medir efectos de la variable independiente sobre la(s) dependiente(s), que incluye un componente cualitativo para evaluar a los sujetos en entrevistas abiertas, llevará a cabo análisis cuantitativo sobre la base de las respuestas al cuestionario, aunque también efectuará análisis cualitativo del material obtenido en las entrevistas.

Dentro del enfoque mixto, la combinación de ambos tipos de análisis y el grado en que se utiliza cada uno de éstos dependen de las circunstancias que rodean la investigación.

UNA NOTA FINAL

El ideal de la investigación, por experiencia propia, es recolectar datos cuantitativos y cualitativos respecto de un problema de investigación, utilizando diversas herramientas (por ejemplo: escalas estandarizadas, observación cualitativa, entrevistas abiertas, sesiones de enfoque y análisis cuantitativo del contenido). Desafortunadamente, los recursos son limitados casi siempre; pero en la medida de lo posible es recomendable triangular los métodos cuanti-cuali.

- Existen hasta ahora dos grandes vertientes para analizar datos: el análisis cuantitativo y el análisis cualitativo.
- El análisis cuantitativo de los datos se efectúa utilizando la matriz de datos, la cual está guardada en un archivo.
- El tipo de análisis o pruebas estadísticas depende del nivel de medición de las variables, las hipótesis y el interés del investigador.

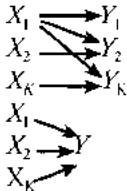
- Los principales análisis estadísticos que pueden hacerse son: estadística descriptiva para cada variable (distribución de frecuencias, medidas de tendencia central y medidas de la variabilidad), la transformación a puntuaciones z , razones y tasas, cálculos de estadística inferencial, pruebas paramétricas, pruebas no paramétricas y análisis multivariados.
- Las distribuciones de frecuencias contienen las categorías, los códigos, las frecuencias absolutas (número de casos), las frecuencias relativas (porcentajes), las frecuencias ajustadas y las frecuencias acumuladas (absolutas o relativas).
- Las distribuciones de frecuencias (particularmente hablando de las frecuencias relativas) pueden presentarse en forma gráfica.
- Una distribución de frecuencias puede representarse a través del polígono de frecuencias o de la curva de frecuencias.
- Las medidas de tendencia central son la moda, la mediana y la media.
- Las medidas de la variabilidad son el rango (diferencia entre el máximo y el mínimo), la desviación estándar y la varianza.
- Otras estadísticas descriptivas de utilidad son la asimetría y la curtosis.
- Las puntuaciones z son transformaciones de los valores obtenidos a unidades de desviación estándar.
- Una razón es la relación entre dos categorías; una tasa es la relación entre el número de casos de una categoría y el número total de casos, multiplicada por un múltiplo de 10.
- La estadística inferencial sirve para efectuar generalizaciones de la muestra a la población. Se utiliza para probar hipótesis y estimar parámetros. Se basa en el concepto de distribución muestral.
- La curva o distribución normal es un modelo teórico sumamente útil; su media es 0 (cero) y su desviación estándar es uno (1).
- El nivel de significancia y el intervalo de confianza son niveles de probabilidad de cometer un error, o de equivocarse en la prueba de hipótesis o la estimación de parámetros. Los niveles más comunes en ciencias sociales son .05 y .01.
- Los análisis o las pruebas estadísticas paramétricas más utilizados son:

<i>Prueba</i>	<i>Tipos de hipótesis</i>
— Coeficiente de correlación de Pearson	Correlacional
— Regresión lineal	Correlacional/causal
— Prueba t	Diferencia de grupos
— Contraste de la diferencia de proporciones	Diferencia de grupos

- Análisis de varianza (ANOVA): unidireccional y factorial. Unidireccional con una variable independiente y factorial con dos o más variables independientes. Diferencia de grupos/causal
- Análisis de covarianza (ANCOVA) Correlacional/causal
- En todas las pruebas estadísticas paramétricas las variables están medidas en un nivel por intervalos o razón.
- Los análisis o las pruebas estadísticas no paramétricas más utilizados son:

<i>Prueba</i>	<i>Tipos de hipótesis</i>
— ji cuadrada (χ^2)	Diferencias de grupos para establecer correlación
— Coeficientes de correlación e independencia para tabulaciones cruzadas: phi, C de Pearson, V de Cramer, lambda, gamma, taub, D de Somers y Eta. Coeficientes de correlación de Spearman y Kendall	Correlacional
- Las pruebas no paramétricas se utilizan con variables nominales u ordinales.
- Los análisis multivariados más usados son:

<i>Prueba</i>	<i>Variables involucradas y niveles de medición</i>	<i>Tipos de hipótesis</i>
— Regresión múltiple	Una dependiente (intervalos o razón) y dos o más independientes (cualquier nivel de medición)	Correlacional/causal $X_1 \rightarrow Y$ $X_2 \rightarrow Y$ $X_k \rightarrow Y$
— Análisis lineal <i>path</i>	Varias, secuencia causal (cualquier nivel de medición)	Correlacional/causal $X \rightarrow W \rightarrow L$ $Y \rightarrow N \rightarrow H$
— Análisis de factores	Varias (intervalos o razón)	Correlacional/causal $X_1 \rightarrow X_2 \rightarrow X_k$

<ul style="list-style-type: none"> — Análisis multivariado de varianza (MANOVA) y correlación canónica — Análisis discriminante 	<p>Varias independientes y varias dependientes (intervalos o razón)</p> <p>Varias independientes (intervalos o razón) y una dependiente (nominal u ordinal)</p>	
---	---	--

- Los análisis estadísticos se llevan a cabo mediante programas computacionales, utilizando paquetes estadísticos.
- Los paquetes estadísticos más conocidos son: SPSS y Minitab. Estos paquetes se utilizan consultando el manual respectivo.
- En los estudios cualitativos, el análisis de los datos no está totalmente predeterminado, sino que es “prefigurado, coreografiado o esbozado”. Es decir, se comienza a efectuar bajo un plan general, aunque su desarrollo va sufriendo modificaciones de acuerdo con los resultados.
- El análisis cualitativo tiene propósitos centrales:

Darle orden a los datos; organizar las unidades, las categorías, los temas y los patrones; comprender, en profundidad, el contexto que rodea a los datos; describir las experiencias de las personas estudiadas bajo su óptica, en su lenguaje y sus expresiones; interpretar y evaluar unidades, categorías, temas y patrones; explicar contextos, situaciones, hechos, fenómenos; generar preguntas de investigación e hipótesis; reconstruir historias; relacionar los resultados del análisis con la teoría fundamentada; o construir teorías.

- El proceso fundamental para efectuar análisis cualitativo se resume en:
 - Revisar que el material (datos) esté listo para el análisis.
 - Establecer un plan inicial de trabajo o una coreografía de análisis.
 - Codificar los datos en un primer plano.
 - Codificar los datos en un segundo plano.
 - Interpretar los datos.
 - Describir contexto(s), eventos, situaciones y personas, sujetos de estudio; encontrar patrones; explicar sucesos, hechos y contextos; construir teoría.
 - Asegurar la confiabilidad y validez de los resultados.
- Durante el proceso, se obtiene retroalimentación, con la finalidad de realizar las correcciones o adecuaciones pertinentes, incluso de volver al campo, si fuera necesario.

- Además, siempre tenemos en mente el planteamiento del problema de investigación.
- Es conveniente que el proceso de análisis sea documentado mediante una bitácora con: 1. comentarios acerca del método de análisis; 2. comentarios sobre los problemas durante el proceso, 3. comentarios sobre la codificación; 4. ideas y comentarios de los investigadores (incluyendo diagramas, mapas conceptuales, dibujos, esquemas, matrices); 5. material de apoyo localizado (fotografías, videos, etcétera); y 6. significados, descripciones y conclusiones preliminares.
- En la mayoría de los estudios cualitativos se codifican los datos para tener una descripción más completa de éstos, resumirlos, eliminar información irrelevante, realizar análisis cuantitativo y generar mayor sentido de entendimiento del material analizado.
- La codificación es clasificar y, en esencia, implica asignar unidades de análisis a categorías de análisis mediante reglas. Algunos autores la denominan *categorización*.
 - Las unidades de análisis pueden ser segmentos de texto o documentos (original o transcrito), material de audio o video, diagramas, mapas, fotografías, artefactos y objetos.
 - Las unidades de análisis son unidades de contenido.
 - Las categorías son “cajones” conceptuales que se crean a partir de la revisión de los datos, por comparación de una unidad respecto de las unidades que le precedieron.
 - La codificación tiene dos planos o niveles: en el primero se codifican las unidades en categorías; en el segundo, se comparan las categorías entre sí para agruparlas en temas y buscar posibles vinculaciones. Es decir, en el primer nivel obtenemos categorías a partir de unidades; y en el segundo, temas a partir de categorías.
 - En ambos casos, a las categorías y los temas (categorías generales) se les asignan códigos para manejarlas(os) más fácilmente.
- En el análisis cualitativo, resulta fundamental darle sentido a las descripciones de cada categoría, los significados de cada categoría, la frecuencia con que aparecen las categorías y las relaciones entre categorías (temporales, causales o de conjunto-subconjunto).
- Para ayudarnos a establecer relaciones entre categorías, podemos dibujar diagramas causales, elaborar matrices y cuadros, sobre la base de nuestra lectura de los datos y las categorías descubiertas.
- Una vez que hemos descrito las categorías y los temas, así como su frecuencia, significado y relaciones, ayudándonos de esquemas gráficos, debemos enmarcar el contexto, la situación o el evento en el cual ocurren.

- En la descripción del contexto retomamos las anotaciones, los memos y la bitácora de campo, que son materiales que elaboramos durante la recolección de datos.
- Así llegamos en el análisis cualitativo a la interpretación de los datos, la generación o prueba de hipótesis y la construcción de teoría fundamentada.
- Con el propósito de respaldar nuestras conclusiones, es importante asegurar la confiabilidad y validez de nuestros análisis ante nosotros mismos y ante los usuarios del estudio.

En el análisis cualitativo, la confiabilidad y la validez no son producto de pruebas estadísticas, sino que se originan mediante una valoración del proceso de análisis.

- Como resultado de esta valoración pudiera darse el caso de que necesitemos regresar al campo para ampliar nuestra base de datos o realizar ajustes en la forma de recolectarlos.
- En todo el proceso de análisis debemos tener en mente el planteamiento original del problema de investigación.
- En nuestra labor de análisis cualitativo, en la actualidad contamos con excelentes herramientas como los programas computacionales para tal tipo de análisis; entre éstos, tenemos: Atlas/ti, The Ethnograph, Nud*ist y Decision Explorer.
- Un estudio que recolecte y analice tanto datos cuantitativos como cualitativos será más enriquecedor.



Análisis de covarianza
 Análisis de datos
 Análisis de factores
 Análisis de varianza
 Análisis discriminante
 Análisis lineal path
 Análisis multivariado de varianza
 Análisis multivariados
 Análisis reflexivo
 Asimetría
 Bitácora de análisis
 Categoría
 Codificación
 Coeficiente de correlación de Pearson
 Coeficiente de Kendall

Coefficiente de Spearman
 Coeficientes de correlación e independencia para tabulaciones cruzadas
 Contraste de la diferencia de proporciones
 Coreografía para el análisis
 Curtosis
 Curva de frecuencias
 Curva o distribución normal
 Descripción del contexto
 Desviación estandar
 Distribución de frecuencias
 Enfoque cualitativo
 Enfoque cuantitativo
 Estadística
 Estadística descriptiva

Estadística inferencial
 Estadística no paramétrica
 Estadística paramétrica
 Gráficas
 Intervalo de confianza
 Ji cuadrada
 Mapeo
 Media
 Mediana
 Medidas de tendencia central
 Medidas de variabilidad
 Métodos cuantitativos
 Moda
 Nivel de significancia
 Paquetes estadísticos
 Patrón

Polígono de frecuencias	Puntuación z	Tabulación cruzada
Procesador de palabras	Rango	Tasa
Programa computacional	Razón	Tema
Programas de análisis cualitativo	Reglas de codificación	Transcripción de texto
Prueba t	Regresión lineal	Unidad de análisis
Pruebas estadísticas	Regresión múltiple	Varianza
	Relación entre categorías	

1. Construya una distribución de frecuencias hipotéticas, con todos sus elementos, e interprétela verbalmente.
2. Localice una investigación científica donde se reporte la estadística descriptiva de las variables y analice las propiedades de cada estadístico o información estadística proporcionada (distribución de frecuencias, medidas de tendencia central y medidas de la variabilidad).
3. Un investigador obtuvo, en una muestra, las siguientes frecuencias absolutas para la variable "actitud hacia el director de la escuela":

<i>Categoría</i>	<i>Frecuencias absolutas</i>
Totalmente desfavorable	69
Desfavorable	28
Ni favorable, ni desfavorable	20
Favorable	13
Totalmente favorable	6

- a) Calcule las frecuencias relativas o los porcentajes.
- b) Grafique las frecuencias relativas a través de un histograma (barras).
- c) Verbalice los resultados respondiendo a la pregunta: ¿la actitud hacia el director de la escuela tiende a ser favorable o desfavorable?

4. Un investigador obtuvo, en una muestra de trabajadores, los siguientes resultados al medir el "orgullo por el trabajo realizado". La escala oscilaba entre 0 (nada de orgullo por el trabajo realizado) y 8 (orgullo total).

Máximo = 5
Mínimo = 0
Media = 3.6
Moda = 3.0
Mediana = 3.2
Desviación estándar = 0.6

¿Qué puede decirse en esta muestra acerca del orgullo por el trabajo realizado?

5. ¿Qué es una puntuación z ? ¿Para qué es útil la estadística inferencial? ¿Qué es la distribución muestral? ¿Qué es la curva normal? ¿Qué son el nivel de significancia y el intervalo de confianza?

6. Relacione las columnas *A* y *B*. En la columna *A* se presentan hipótesis; en la columna *B*, pruebas estadísticas apropiadas para las hipótesis. Se trata de encontrar la prueba que corresponde a cada hipótesis. (Las respuestas se localizan en el Apéndice 4.)

<i>Columna A</i>	<i>Columna B</i>
— Hi: "A mayor inteligencia, mayor capacidad de resolver problemas matemáticos" (medidas las variables por intervalos).	— Diferencias de proporciones
— Hi: "Los hijos de padres alcohólicos muestran una menor autoestima con respecto a los hijos de padres no alcohólicos" (autoestima medida por intervalos).	— Ji cuadrada
— Hi: "El porcentaje de delitos por asalto a mano armada, en relación con el total de crímenes cometidos, es mayor en la ciudad de México que en Caracas."	— Spearman
— Hi: "El sexo está relacionado con la preferencia por telenovelas o espectáculos deportivos."	— Coeficiente de correlación de Pearson
— Hi: "La intensidad del sabor de productos empacados de pescado está relacionada con la preferencia por la marca" (sabor intenso, sabor medianamente intenso, sabor poco intenso, sabor muy poco intenso) (preferencia = rangos a 12 marcas).	— ANOVA unidireccional
— Hi: "Se presentarán diferencias en cuanto al aprovechamiento entre un grupo expuesto a un método de enseñanza novedoso, un grupo que recibe instrucción mediante un método tradicional y un grupo de control que no se expone a ningún método."	— Prueba <i>t</i>

7. Un investigador obtuvo un valor t igual a 3.25, con 63 grados de libertad y un nivel de confianza o significancia de .05, ¿aceptaría su hipótesis de investigación? (Respuesta en el Apéndice 4.)
8. Otro investigador obtuvo un valor de X^2 (ji cuadrada) de 6.12, con tres grados de libertad y un nivel alfa de .05, ¿aceptaría su hipótesis de investigación? (Respuesta en el Apéndice 4.)
9. Genere un ejemplo hipotético de una razón F significativa e interprétela.
10. Construya un ejemplo hipotético de una tabulación cruzada y utilícela para fines descriptivos.
11. Busque un artículo de investigación en revistas científicas que contengan resultados de pruebas t , “ANOVA”, “ANCOVA” y X^2 aplicadas; evalúe la interpretación de los autores.
12. Elija una poesía de su preferencia que tenga como mínimo dos cuartillas estándares; tome como unidad de análisis el verso o el párrafo; codifique el contenido (texto) de la poesía. ¿Qué categorías encontró? ¿Qué categoría tuvo mayor frecuencia de aparición en el texto?
13. Realice una entrevista abierta sobre la opinión de uno de sus compañeros(as) sobre la necesidad que tiene el mundo de alcanzar la paz; grabe la entrevista en casete, CD u otro medio. Transcriba la entrevista (no olvide incluir pausas y sonidos o elementos paralingüísticos como: ¡mmm!, ¡ahh!, etcétera; describa su contexto, y a su compañero(a); codifique y categorice la transcripción de la entrevista. ¿Qué categorías surgieron de la codificación? Asigne un código a cada categoría. Vuelva a leer la entrevista. Analice en profundidad —reflexionando seriamente— y responda: ¿qué piensa en realidad su compañero(a) sobre la paz y la necesidad de alcanzarla y qué piensa sobre la guerra?
Contéstese a sí mismo: ¿qué sabe hoy que no sabía antes de su compañero(a)?
14. Codifique en primero y segundo niveles o planos cinco canciones de un álbum o disco de su grupo o cantante preferido.
15. Seleccione el libro de un historiador, antropólogo, arqueólogo, etnólogo, que describa una cultura o comunidad pasada o presente, urbana o rural. Analice qué tan profunda y completa es su descripción de la comunidad o cultura y qué le hace falta a esta descripción (si no consigue un libro o un artículo de revista académica, le serviría un programa cultural de la televisión y también una biografía de algún personaje histórico).
16. Aplique todo el proceso de análisis cualitativo a los resultados de un par de sesiones de enfoque que realice con amigos o familiares suyos sobre una situación local (problema social, elecciones, impuestos, desarrollo de la economía regional, malos servicios médicos, el divorcio de una pareja conocida, la problemática ambiental, cómo son las relaciones heterosexuales entre jóvenes de la localidad, etcétera).



- BAPTISTE, I. (septiembre, 2001). Qualitative data analysis: Common phases, strategic differences. *Forum Qualitative Sozialforschung/Forum: Qualitative social research* (on-line journal), 2(3). Disponible en: <http://www.qualitative-research.net/fqs-eng.htm> (1,1,2002).
- CARMINES, E. G. y ZEULER, R. A. (1988). *Reliability and validity assessment*. Serie Quantitative Applications in the Social Sciences, vol. 17. Beverly Hills, California: Sage Publications, Inc.
- CAMACHO, J. (2001). *Estadística con SPSS para Windows*. México: Alfaomega Grupo Editor.
- CRESWELL, J. (1998). *Qualitative inquiry and research designs. choosing among five traditions*. Thousand Oaks, CA: Sage Publications.
- DIAMOND, I. y JEFFERIES, J. (2001). *Beginning statistics: An introduction for social scientists*. CA: Sage Publications.
- FERRÁN ARANAZ, M. (2001). *SPSS para Windows: Análisis estadístico*. Madrid: Osborne McGraw-Hill.
- FOSTER, J. J. (2001). *Data analysis using SPSS for Windows versions 8 to 10: A beginner's guide*. CA: Sage Publications.
- GRINELL, R. M. (1997). *Social work research & evaluation: Quantitative and qualitative approaches* (5a. ed.). Itasca, Illinois: E. F. Peacock Publishers.
- HENKEL, R. E. (1976). *Test of significance*, Serie Quantitative Applications in the Social Sciences, vol. 4, Beverly Hills, California: Sage Publications, Inc.
- HILDEBRAND, D. K.; LAING, J. D. y ROSENTHAL, H. (1977). *Analysis of ordinal data*. Serie Quantitative Applications in the Social Sciences, vol. 8, Beverly Hills, California: Sage Publications, Inc.
- INVERSEN, G. R. y NORRIS, H. (1987). *Analysis of variance* (2a. ed.), Serie Quantitative Applications in the Social Sciences, vol. 1, Beverly Hills, California: Sage Publications, Inc.
- LEVIN, J. (1979). *Fundamentos de estadística en la investigación social*, México: Harla.
- NEWTON, R. R. y RUDESTAM, K. E. (1999). *Your statistical consultant: Answers to your data analysis questions*. CA: Sage Publications.
- PATTON, M. Q. (1990). *Qualitative evaluation and research methods*, Newbury Park, CA: Sage Publications.
- REYNOLDS, H. T. (1984). *Analysis of nominal data* (2a. ed.), Serie Quantitative Applications in the Social Sciences, vol. 7, Beverly Hills, California: Sage Publications, Inc.
- SEGEL, S. (1995). *Estadística no paramétrica aplicada a las ciencias de la conducta* (4a. ed.) México: Trillas.
- WIERSMA (1999). *Research methods in education: An introduction*, Boston Mass: Allyn and Bacon, Inc. capítulo 12.
- WILDT, A. R. y AHTOLA, O. T. (1978). *Analysis of covariance*, Serie Quantitative Applications in the Social Sciences, vol. 12, Beverly Hills, California: Sage Publications, Inc.
- WRIGHE, S. R. (1979). *Quantitative methods and statistics: A guide to social research*, Beverly Hills, California: Sage Publications, Inc.
- Análisis multivariado:**
- BLALOCK, H. (1988). *Estadística social*, México: Fondo de Cultura Económica.
- CAMACHO, J. (2001). *Estadística con SPSS para Windows*. México: Alfaomega Grupo Editor.
- FERRÁN ARANAZ, M. (2001). *SPSS para Windows: Análisis estadístico*, Madrid: Osborne McGraw-Hill.
- KERLINGER, F. N. y PEDHAZUR, E. J. (1997). *Multiple regression in behavioral research* (3a. ed.), Nueva York, N.Y.: Holt, Rinehart and Winston, Inc.

- KESSELER, R. C. y GREENBERG, D. E. (1981), *Linear panel analysis: Models of quantitative change*, Nueva York, N.Y.: Holt, Rinehart and Winston, Inc.
- KIM, J. O. y MUELLER, Ch. (1978), *Introduction to factor analysis*, Serie Quantitative Applications in the Social Sciences, vol. 13, Beverly Hills, California: Sage Publications, Inc.
- KIM, J. O. y MUELLER, Ch. (2000), *Factor analysis: Statistical methods and practical issues*, Serie Quantitative Applications in the Social Sciences, vol. 14, Beverly Hills, California: Sage Publications, Inc.
- KRUSKAL, J. P. y WISH, M. (1984), *Multidimensional scaling*, Serie Quantitative Applications in the Social Sciences, vol. 11, Beverly Hills, California: Sage Publications, Inc.
- LEVINE, M. S. (1977), *Canonical analysis and factor comparison*, Serie Quantitative Applications in the Social Sciences, vol. 6, Beverly Hills, California: Sage Publications, Inc.
- MONGE, P. R. y CAPPELLA, J. N. (eds.) (1980), *Multivariate techniques in human communication research*, Nueva York, N.Y.: Academic Press.
- NIE, N. H.; HULL, C. H.; JENKINS, J. G.; STEINBRENNER, K., y BENDT, D. H. (1975), *Statistical Package for the Social Sciences*, Nueva York, N.Y.: Academic Press.
- SCHWARTZMAN, S. (Comp.) (1977), *Técnicas avanzadas en ciencias sociales*, Buenos Aires: Nueva Visión SAIC.

LA TELEVISIÓN Y EL NIÑO (análisis cuantitativo)

Las estadísticas a obtener son:

- Distribución de frecuencias y estadísticas descriptivas de las variables "tiempo que dedican los niños a ver televisión", "programas televisivos preferidos de los niños", "funciones que cumple la televisión en los niños", "gratificaciones que obtienen los niños al exponerse a la televisión" y "control que ejercen los padres sobre el uso de la televisión que hacen sus hijos".
- Prueba t (diferencia de medias) para evaluar las diferencias entre niñas y niños respecto del tiempo que cada grupo dedica a ver televisión.
- Ji-cuadrada para correlacionar las variables sexo (nominal) y gratificaciones que derivan los niño(a)s de ver televisión.

Etcétera.

EL CONTAGIO DEL SIDA (cuantitativo)

Estadísticas que se podrían obtener:

- Estadística descriptiva del perfil del enfermo por VIH (edad, sexo, ocupación, prácticas sexuales, estado de avance de su enfermedad, etcétera), y estadísticas del contexto (número de enfermos, vías o medios por los cuales se adquirió el VIH, etcétera).

- Prueba t para ver el tiempo promedio en el que se presentaron los síntomas del SIDA entre los enfermos que adquirieron el VIH por transfusiones sanguíneas (grupo 1) y los que lo adquirieron por contacto sexual (grupo 2).
 - Correlación de Pearson (r) entre edad del enfermo y tiempo que tardó en presentar síntomas de la enfermedad (en meses).
- Etcétera.

LA MODA Y LA MUJER MEXICANA (enfoque mixto)

De las sesiones de grupo, se obtuvo:

- Categorías de la moda (percepción de la moda, marcas a la moda, tiendas a la moda); preferencia de tiendas; percepción del departamento de ropa para mujeres; reconocimiento de secciones, personal, tallas, ofertas; evaluación de las secciones; comparación con principales competidores y sugerencias.

De la encuesta:

- Estadística descriptiva de las variables, prueba t de las evaluaciones (intervalos) de cada sección por grupos de edad: jóvenes y adultos y análisis de varianza de tales evaluaciones por zonas geográficas y ocupación, así como ji cuadradas entre variables nominales según el caso.



Desde 1990 han disminuido las tensiones entre lo cualitativo *versus* lo cuantitativo, por lo que se buscó establecer una sinergia, así como ser más flexibles y eclécticos, dicho en el buen sentido, en los procedimientos.

De este modo, la investigación cualitativa ganó rigurosidad, credibilidad, verificabilidad y la posibilidad de triangular. El empleo de programas computacionales, para recolectar y analizar información, contribuyó al incremento de este tipo de estudios, principalmente de grupos sociales en situaciones de marginalidad, es decir, aquellos con carencia de recursos, que son afectados por la violencia.

La investigación cuantitativa, en tanto, ganó cuando particularizó los instrumentos y tomó en cuenta las características de los grupos a los cuales se dirige el estudio. Lo anterior propició un importante avance en la explicación de los procesos psicológicos, en especial los cognoscitivos; y en los descubrimientos neuropsicológicos, así como en el uso de software para el montaje de experimentos, demostraciones y simulaciones.

En este tipo de investigaciones, destacan las pruebas estadísticas por su utilidad en el análisis de datos categóricos de correspondencia, la ordenación de datos para conocer preferencias, el análisis factorial confirmativo, las correctas estimaciones de conjuntos de datos complejos, el manejo de resultados estadísticos de los experimentos, la validación de datos, la determinación del tamaño de la muestra y el análisis de regresión, entre otros aspectos a considerar.

A pesar de tan importantes avances en la investigación, aún hace falta financiamiento para una promoción significativa y que, además, fomente la especialización de los investigadores, lo cual les permitiría competir de manera efectiva.

CIRO HERNANDO LEÓN PARDO
Coordinador del Área de Investigación
Facultad de Psicología
Universidad Javeriana
Bogotá, Colombia

Para efectuar una buena investigación se requiere plantear de forma correcta el problema, con lo cual tenemos 50 por ciento de la solución, y también con un rigor metodológico, es decir, incluir todos los pasos del proceso.

Tal apego a la metodología implica el empleo de los recursos pertinentes; por ejemplo, en las investigaciones sociales las pruebas estadísticas proporcionan una visión más precisa del objeto de estudio, ya que apoyan o no las hipótesis para su validación o rechazo.

Los estudiantes pueden concebir una idea de investigación a partir de sus intereses personales, aunque se recomienda que elijan temas íntimamente relacionados con su carrera, y que procuren que sean de actualidad y de interés común.

Para ello, los profesores deben infundir en los alumnos la importancia de la investigación en el terreno académico y en el profesional, destacando su relevancia tanto en la generación de conocimiento como en la búsqueda de soluciones a problemas.

ROBERTO DE JESÚS CRUZ CASTILLO
Profesor de tiempo completo
Facultad de Ciencias de la Administración
Universidad Autónoma de Chiapas
Chiapas, México

Plantear un problema de investigación consiste en formular una serie de preguntas acerca de lo que se desea estudiar, lo cual también conduce a determinar la manera de llevar a cabo el proyecto.

Cuando la interrogante del estudio es cómo suceden las cosas, se utiliza el análisis cualitativo, en tanto que el enfoque cuantitativo responde a las preguntas cuántas veces o por qué sucede cierto fenómeno.

En investigación cualitativa ha habido grandes avances, como son los criterios de credibilidad, la transferibilidad de datos y los paquetes de análisis, entre los cuales *Atlas/ti* es el más completo para estudiar textos.

En cuanto a los avances en el análisis cuantitativo, encontramos las técnicas multivariantes avanzadas, el análisis de conglomerado, el escalamiento multidimensional y el análisis de series cronológicas.

DR. ANTONIO TENA SUCK
Coordinador del Posgrado de Psicología
Universidad Iberoamericana
D. F., México

Lectura 4

7

El uso de la computadora como auxiliar en el análisis de datos cualitativos

Lilia Beatriz Chernobilsky

Este capítulo está dirigido a investigadores sociales ya familiarizados con los métodos cualitativos o a aquellos que se iniciaron en el tema recientemente a través de la lectura de los capítulos anteriores, pero intentan usar un software específico que los asista en la tarea del análisis de datos.

Se intenta, por un lado, poner de manifiesto la influencia de las nuevas tecnologías de la información en la investigación cualitativa, un campo que hasta no hace mucho tiempo permanecía ajeno al uso de herramientas informáticas específicamente desarrolladas, al menos en los países latinoamericanos. En particular, observar cómo la convergencia digital, que es potencialmente uno de los desarrollos tecnológicos más significativos, puede asistir al investigador cualitativo.

Por otro lado, y a partir de algunas reflexiones que distan de posiciones extremas sobre los programas informáticos, sus capacidades e incompetencias y el conocimiento que debe tener el investigador antes de tomar una decisión con respecto a su utilización o adquisición, se detalla una lista exhaustiva de funciones propias de estos programas. Se pretende muy especialmente alertar al investigador sobre las consideraciones previas a tener en cuenta antes de incursionar en el uso de estos asistentes informáticos. Por último se acompaña el capítulo con un estudio comparativo de los principales software disponibles y sus características.

Tomada de: Lilia Chernobilsky, "El uso de la computadora como auxiliar en el análisis de datos cualitativos", en *Estrategias de investigación cualitativa*, Barcelona, Gesida, 2006, pp. 239-273.

La inclusión de ejemplos mediante la aplicación del programa *Atlas.ti*, uno de los programas más completos en cuanto a las funciones detalladas, tiene por objeto convencer al lector del potencial y las bondades de estos software.

1. Un poco de historia

Hasta la década de 1980, los investigadores sociales utilizaban las herramientas informáticas solo para el cálculo estadístico de los datos. Los métodos estadísticos, base del análisis cuantitativo, pueden ser descritos como un procedimiento algorítmico que tiene como finalidad resolver un problema determinado. Las nuevas tecnologías de la comunicación y de la información beneficiaron fuertemente a los investigadores cuantitativos, que comenzaron a contar con equipos poderosos –hardware– así como también con programas apropiados para el análisis de datos numéricos. Sin embargo, estas innovaciones daban por hecho el conocimiento adecuado por parte del investigador de la prueba o test a aplicar (estadísticas descriptivas, coeficiente de regresión, coeficiente de correlación, etc.) y de la interpretación de los resultados.

Por el contrario, el análisis de datos cualitativos es esencialmente una actividad hermenéutica que intenta interpretar vivencias relacionadas, experiencias vistas o creencias de las personas en diferentes situaciones sociales, además de poner esta exégesis a disposición de la comunidad de investigadores. Lo importante, tanto en la investigación cualitativa como en la cuantitativa, es que las técnicas usadas no distorsionen o corrompan los datos (Roberts y Wilson, 2002).

En los últimos años, las metodologías cualitativas han logrado una gran aceptación por parte de investigadores de distintas disciplinas. Esto trajo aparejado el desarrollo de una serie de herramientas informáticas que intentan auxiliar al investigador cualitativo en la ardua tarea de procesar sus datos. Si bien en un principio se conformaron con usar programas del tipo «multipropósito» como un procesador de textos o un administrador de bases de datos (que no dejan de ser auxiliares poderosos de la investigación cualitativa), a partir de la década de 1980 –y con mayor desarrollo en la de 1990– comenzó a proliferar una serie de programas específicos para el análisis de datos cualitativos, conocidos con la sigla CAQDAS (*Computer Assisted Qualitative Data Analysis Software*), utilizada por primera vez por Fielding y Lee en 1991.

Las diferencias entre la filosofía subyacente en las nuevas tecnologías de la comunicación y la información basadas en procesos algorítmicos por un lado, y el pensamiento filosófico detrás de la investigación cualitativa por otro, marcan distancias difíciles de salvar y aceptar por

los investigadores sociales cualitativos, dadas sus profundas implicancias metodológicas. No obstante, la naturaleza misma del análisis cualitativo está caracterizada por una serie de procedimientos que pueden ser sistematizados con el objeto de asistir al investigador en tareas rutinarias.

Más aun, los soportes digitales en los que se presentan los datos textuales, de audio y vídeo hoy en día tornan posibles nuevas maneras de producirlos, procesarlos y analizarlos. Paralelamente, el crecimiento de la red Internet también hace viables nuevas formas y oportunidades de recolectar los datos cualitativos y establecer nuevos contextos de interacción para su análisis.

En la actualidad existe una gama enorme de software para el análisis de datos cualitativos, y como respuesta a las demandas de los investigadores, los programas son actualizados constantemente con nuevas funciones y características que facilitan las tareas para las cuales fueron diseñados. Entre los distintos tipos de programas, vale la pena destacar aquellos específicamente desarrollados para asistir al investigador en el proceso de construcción de conceptos cualitativos, tipologías y desarrollos teóricos. Sin embargo, la creciente diversidad hace necesario crear patrones comunes para armonizar e intercambiar datos y resultados del análisis cualitativo.

2. Convergencia digital

La investigación cualitativa, como otros ámbitos científicos, también está siendo influenciada aceleradamente por el desarrollo y el despliegue de las Tecnologías de la Información. Para entender este fenómeno es necesario comprender el concepto de convergencia digital.

La convergencia digital es potencialmente uno de los desarrollos tecnológicos más significativos que pueden asistir al investigador cualitativo. Esta representa, según Covell (1999: 15),

la reconciliación entre una amplia gama de nuevas tecnologías y las tradicionales, o sea, entre las capacidades computacionales más avanzadas, las nuevas tecnologías digitales multimediales y las nuevas tecnologías digitales de las comunicaciones. Esta combinación del poder de cálculo y la funcionalidad, la red de interconectividad digital, y la capacidad multimedial permite nuevas formas de interacción humana, colaboración e intercambio de información. La convergencia digital está modificando el modo en que los individuos y las organizaciones colaboran y comparten información. El audio, el vídeo, las animaciones, entre otras formas multimediales, mejoran las comunicaciones digitales existentes y permiten nuevas formas de interacción humana.

La pregunta a responder es cómo podríamos digitalizar el proceso de conocimiento de manera que el uso de estas tecnologías represente una mejora más que una amenaza a la naturaleza ideográfica de la investigación cualitativa. Asimismo, de qué manera conseguiríamos explorar el potencial de la convergencia digital de manera que no altere la creatividad individual, la pluralidad de pensamientos y de representaciones y al mismo tiempo nos beneficie con la eficacia ganada al usar la computadora como una herramienta digital durante todo el proceso de investigación (Brown, 2002). El objetivo es, pues, saber cómo llegar a una solución que permita capturar y procesar digitalmente todos los datos con un alto grado de calidad, sin alterar las características fundamentales del paradigma cualitativo distinguidas por Maxwell (1996): el proceso abierto, la flexibilidad y una fuerte orientación hacia el razonamiento inductivo.

3. Reflexiones acerca de los CAQDAS

Mucho se ha escrito acerca del uso de las computadoras en el análisis de datos cualitativos. Algunos autores muestran preocupación con respecto a que el software pueda guiar al investigador en alguna dirección en particular (Seidel, 1991). Otros piensan que la utilización de estos programas podría distanciar al investigador de sus datos, llevándolo a efectuar un análisis cuantitativo de los datos cualitativos (Barry, 1998). Sin embargo, los más fervorosos defensores de los CAQDAS sostienen que estos permiten un proceso ajustado y transparente de análisis de los datos (Richards y Richards, 1994).

En lo que respecta, por ejemplo, a la Argentina actual, el uso de herramientas informáticas para el análisis de datos cualitativos (CAQDAS) aún está circunscrito a un número reducido de investigadores en ciencias sociales y humanas. Tal como sucede en otros países de América Latina, siguiendo la clasificación propuesta por Cisneros Puebla (2003) podemos decir que nuestra realidad muestra hoy la coexistencia de tres tipos de investigadores cualitativos: 1) los que prefieren las técnicas artesanales de trabajo (registros manuales de campo, lápices de colores, tijeras, fichas, goma de pegar, etc.); entre ellos se encuentran los más «románticos» del trabajo manual –para ellos esta es la verdadera forma de «palpar» los datos y hacer análisis interpretativo–, quienes rechazan categóricamente el uso de técnicas informáticas, con frecuencia por su poca afinidad con las computadoras o su falta de experiencia con ellas; 2) los que privilegian el uso de programas generales o «multipropósito» (procesadores de textos, bases de datos, planillas de cálculo, etc.), quienes, satisfechos por el uso y los resultados obtenidos con alguna técnica, no quieren incursionar en otras, bien por cierto temor a enfrentarse con algún software más complejo o por falta

de tiempo para aprender a manejarlo; y 3) los que, en efecto, utilizan los programas específicamente diseñados para el análisis de datos cualitativos, entre quienes se cuentan los grandes defensores del CAQDAS, quienes declaran no querer volver a ninguna de las etapas anteriores, y los –pocos, por razones de escasez de recursos– investigadores que van al trabajo de campo acompañados de su computadora portátil, su cámara digital para captar fotografías o imágenes de vídeo o su grabadora digital, o que poseen soporte técnico para transcribir digitalmente materiales cualitativos que habían sido originalmente almacenados en formato analógico (Chernobilsky y D'Onofrio, 2004).

4. Capacidades e incompetencias de los CAQDAS en el análisis de datos cualitativos

¿Qué permite hacer un CAQDAS?
¿Qué no permite hacer un CAQDAS?

La investigación cualitativa genera una considerable cantidad de material, fundamentalmente textual. Los investigadores que utilizan métodos manuales para registrar, organizar y manipular esa gran cantidad de datos pasan horas tratando de clasificarlos, utilizando lápices de colores, cortando y pegando papeles. Existe la creencia entre este tipo de investigadores de que cuanto más tiempo dediquen a estas tareas manuales y más «metan las manos» en los datos, mejores serán los resultados alcanzados. Para ellos, es una manera de estar literalmente en contacto físico con sus datos y de conectarse con el objeto de estudio. Sin embargo, para convencer al investigador «artesanal» de cuáles son las posibilidades de la computadora en la investigación cualitativa es necesario saber qué pueden y qué no pueden hacer estas herramientas en el proceso de investigación.

En principio, podemos afirmar que las computadoras pueden ser programadas para realizar tareas mecánicas de análisis, tales como identificar similitudes, diferencias y relaciones entre distintos fragmentos de texto, pero de ninguna manera se las puede preparar para la actividad de creación conceptual. Y aunque se hable del software para el análisis de datos cualitativos, no puede reemplazar al investigador en su actividad analítica; *el software no puede analizar los datos en lugar del investigador*. Si bien por un lado las técnicas de manejo y almacenamiento de datos se han desarrollado y vuelto cada vez más complejas y matemáticamente muy sofisticadas, por el otro nuestros esquemas de interpretación, que son los que dan el verdadero sentido a nuestros datos, no han sido estructurados eficazmente.

En general, podemos asegurar que los CAQDAS permiten:

- Integrar y estructurar los datos de manera que el investigador pueda acceder a las distintas partes de su proyecto, constituido como una sola unidad.
- Utilizar herramientas para la búsqueda de texto.
- Facilitar la tarea operativa de codificar.
- Recuperar segmentos codificados.
- Consultar la base de datos buscando relaciones preestablecidas por el investigador entre los códigos.
- Facilitar la escritura de memos, comentarios y anotaciones.
- Escribir informes finales como resultado de la investigación.

Es importante destacar que el rol de la computadora no consiste en influir en el tipo y calidad de los datos. Por más poderoso que sea el programa informático o por más habilidoso que sea el investigador en su manejo, si los datos tienen un contenido pobre eso se reflejará en los resultados. De la misma manera, es responsabilidad del investigador y no de la computadora seleccionar qué partes de los documentos textuales ilustrarán las categorías o los temas en estudio, así como también determinar qué método utilizar para analizar los datos. Sería muy ingenuo suponer que un software puede ser capaz de realizar estas elecciones o mejorar datos irrelevantes.

5. Principales usos de la computadora en el análisis de datos cualitativos

¿En qué actividades usaría la computadora durante el análisis de datos cualitativos?

La lista a continuación ilustra varios usos posibles de las computadoras en el proceso de análisis de datos cualitativos, muchos de ellos ya presentados por Miles y Huberman (1994: 44). Algunas de estas funciones son muy simples y pueden ser ejecutadas con un procesador de textos, mientras que otras más complejas requieren de programas más sofisticados. Queda a criterio del lector determinar cuándo usar uno u otro.

1. Tipiar transcripciones de notas de campo «en bruto» del investigador.
2. Corregir, ampliar y editar notas de campo.
3. Tipiar, corregir y editar transcripciones de observaciones, grabaciones de entrevistas, gráficos, vídeos y otros documentos para su utilización como material de análisis.

4. Almacenar los materiales del análisis cualitativo en una base de datos organizada.
5. Marcar fragmentos del material considerados relevantes por el investigador.
6. Codificar, es decir, registrar los temas y las interpretaciones emergentes durante la lectura y reflexión sobre los datos, adjuntando una o múltiples categorías de análisis (códigos, etiquetas, palabras clave o breves frases) a cada uno de los fragmentos de material juzgados como relevantes.
7. Buscar y recuperar los fragmentos de material identificados como relevantes y/o categorizados analíticamente.
8. Refinar progresivamente la categorización analítica a través de la modificación, la agrupación y/o la eliminación de códigos.
9. Tipiar, corregir, ampliar y editar memos o apuntes analíticos del investigador (comentarios reflexivos acerca de temas emergentes, ideas, conceptos y proposiciones teóricas elaboradas a partir de los datos y/o la lectura y relectura de material bibliográfico, referidas a aspectos focalizados de los datos y que sirven de base para análisis más detallados).
10. Buscar y recuperar palabras o frases para su análisis.
11. Vincular datos, es decir conectar unos fragmentos relevantes de datos con otros, formando agrupamientos o redes de información.
12. Separar datos correspondientes a los diversos casos estudiados para su análisis comparado.
13. Organizar aspectos de los datos en matrices y/o cuadros-resúmenes para su visualización e interpretación.
14. Vincular categorías, es decir conectar categorías de análisis con otras, formando redes semánticas que especifiquen relaciones jerárquicas o no jerárquicas entre conceptos vinculados a fragmentos de datos.
15. Probar hipótesis de trabajo mediante búsquedas complejas que recuperen fragmentos de datos pertenecientes a redes semánticas, para la construcción de teoría fundamentada en los datos.
16. Mapear, es decir elaborar diagramas que muestren redes semánticas y desarrollos teóricos.
17. Analizar contenidos mediante el conteo de frases.
18. Tipiar informes de resultados de la investigación.

Los puntos 11, 14, 15 y 16 claramente no pueden, hoy en día, realizarse con un procesador de textos.

6. Consideraciones previas a la utilización de un CAQDAS

¿Qué es necesario hacer antes de usar un CAQDAS?

Antes que nada y como punto de partida, se requiere *comprender qué es el análisis de datos cualitativos*; o sea, conocer a fondo la *tradicción* o estrategia metodológica a aplicar, y en vinculación con ella: a) las características del proceso de la preparación de los datos y las formas de su transcripción; b) los pasos del análisis, interpretación y/o codificación; c) el proceso de escritura durante el análisis; y d) la escritura del reporte final.

6.1. Determinar la tradición metodológica a considerar

Supone investigar las diferentes estrategias metodológicas disponibles e identificar cuál es la más conveniente para el proyecto en cuestión. Varios de los programas específicos para el análisis de datos cualitativos parecen estar fuertemente influenciados por estrategias de análisis como la «teoría fundamentada» (*Atlas.ti, N6*). Sin embargo, según los desarrolladores de software estos están basados en diferentes concepciones metodológicas de cómo se produce el conocimiento de la realidad social (Kelle, 1997). John Seidel ha desarrollado y usado su paquete *Ethnograph* en varios proyectos, en que aplica métodos de análisis de discurso enraizado en concepciones fenomenológicas y etnográficas. Por otro lado, *Maxqda* –creado por Udo y Anne Kuckartz–, se funda en el concepto de «tipos ideales» de Max Weber. *QDA Miner* fue concebido manteniendo la integración de las metodologías cualitativas y cuantitativas. *Atlas.ti* y *Maxqda* fueron probados satisfactoriamente en análisis de contenido cualitativo (Mayring, 2000).

En definitiva, el aporte de los CAQDAS al análisis cualitativo depende de la tradición o estrategia que se intente utilizar. Si bien estos aparecen como un método en sí mismo, lo cual es erróneo, no necesariamente se adaptan a todo tipo de análisis cualitativo. Aunque están pensados como programas que se pueden utilizar siguiendo los lineamientos de cualquier estrategia de análisis cualitativo, y a pesar de las diferencias metodológicas y las diversas bases teóricas de los diseñadores de software, todos han cimentado sus programas en los algoritmos de codificación y recuperación, y particularmente en la «teoría fundamentada» (*Grounded Theory*, véase cap. 4), ya que sostienen el argumento de que representan una «tecnología abierta» aplicable en contextos teóricos y metodológicos variados (Kelle, 1997).

6.1.1. Recolectar y preparar los datos

La forma más común de recolectar datos cualitativos es la entrevista, realizada ya sea personalmente, por teléfono, por correo electrónico o vía Internet. Una vez terminada, el investigador tiene que decidir el formato que debe darle a sus datos para un análisis apropiado. Generalmente se recurre a la transcripción, aunque el análisis puede efectuarse directamente a partir de la grabación. Los datos pueden provenir también de notas de campo obtenidas de la observación participante o de grupos focales en forma de textos, gráficos, fotos, vídeos, los cuales deben ser transformados convenientemente para su análisis. En esta instancia es necesario tener en cuenta cuáles son los programas informáticos que pueden manejar los distintos tipos de datos.

6.1.2. El tipo de codificación

La codificación consiste en identificar pasajes de textos o secciones de imágenes en un documento que representan una idea o un concepto. El investigador puede organizar su codificación de distintas formas. Puede 1) utilizar una lista de códigos que le permita incluir distintos niveles de codificación o agrupamiento en subcódigos, es decir, una codificación jerárquica; 2) utilizar una lista de códigos que *no* le permita incluir distintos niveles de codificación o agrupamiento en subcódigos, o sea, codificación no jerárquica; o 3) ambas posibilidades a la vez. Los programas CAQDAS se diferencian en la forma en que operan la codificación, esto es la creación y la recuperación del segmento codificado en el contexto de los datos.

6.1.3. El proceso de escritura

La escritura de notas e impresiones del investigador durante la investigación cualitativa es una parte muy importante del análisis en sí mismo. La escritura comienza en la fase inicial de la investigación y continúa durante todo el proceso hasta el informe final. Las distintas formas de escritura pueden resumirse en registros diarios, notas de campo, memos y el mismo informe final. En este sentido, es importante considerar cuáles son los programas informáticos que permiten la creación de archivos de texto fácilmente legibles en otras aplicaciones, tales como el procesador de textos Word de Microsoft.

6.2. Otras consideraciones a tener en cuenta previas a la utilización de un CAQDAS

Si el investigador está decidido a optar por *el análisis asistido por una computadora* en lugar de *optar por un análisis manual tradicio-*

nal con lápiz y papel, deberá saber que la elección dependerá, entre otros aspectos, de:

- a) la cantidad de datos recolectados: si se ha entrevistado un número manejable de individuos probablemente no sea necesario utilizar un CAQDAS, pero sí hay que tener en cuenta que una de las potencialidades de este tipo de programas es su capacidad de manejar grandes cantidades de datos distribuidos en diferentes archivos, en forma mucho más rápida de lo que se lograría manualmente;
- b) si el software será usado sólo como una ayuda para el proceso de codificación, en cuyo caso se puede recurrir a programas de operatividad más simple y menores prestaciones;
- c) si se incorporan también datos cuantitativos;
- d) el tiempo disponible para la etapa de análisis –en el que deberá incluirse el período de aprendizaje del software–; ese tiempo puede llegar a ser más extenso según la afinidad de los usuarios con la herramienta informática en general;
- e) la afinidad de los usuarios con determinados entornos informáticos (PC, Mac);
- f) la disponibilidad de equipamiento informático en la institución donde se efectúa el trabajo y de presupuesto para la adquisición de estos materiales; y
- g) si la investigación será ejecutada en forma individual o en equipo.

7. Programas CAQDAS disponibles

<p style="text-align: center;">¿Qué tipo de programas se consideran CAQDAS? ¿En qué se basan los programas CAQDAS?</p>

No es una tarea sencilla advertir a simple vista qué nos ofrece un CAQDAS cuando incursionamos en él por primera vez. Muchas veces, ante la duda de estar eligiendo el programa adecuado, el investigador se queda con el que tiene a mano; pero cuando debe comprarlo para su propio proyecto de investigación, se encuentra ante una verdadera encrucijada. Es conveniente para ello que esté informado sobre qué tipo de programas se consideran CAQDAS y, por otro lado, en qué se basan los programas CAQDAS.

Principalmente, se consideran CAQDAS a aquellos programas que permiten una aproximación cualitativa a los datos cualitativos, es decir que intentan asistir al investigador para interpretarlos, mediante la identificación y codificación de temas, o sea, atribuir categorías, crear

conceptos, etc., para llegar a explicar localmente un fenómeno, adaptar, modificar, profundizar o desarrollar una teoría. Estos programas asisten en el acceso, la manipulación y el análisis de datos textuales y/o multimediales eliminando la tarea manual de clasificar, ordenar y preparar la información. Además, están constituidos por una serie de herramientas que ayudan a clarificar el significado de los datos.

Los programas que utilizan los investigadores cualitativos hoy en día se dividen en dos tipos: 1) genéricos o «multipropósito», y 2) específicos (Richards y Richards, 1994). Dentro del tipo genérico se incluyen los procesadores de textos y administradores de bases de datos. Dentro de los específicos, si bien Miles y Weitzman (1995) sugieren dos tipos de programas, a) codificación y recuperación, y b) codificación y generación de teoría, en la actualidad la mayoría de los programas está comprendida en esta última categoría.

Los programas CAQDAS están diseñados para permitir al analista el marcado de segmentos de datos al adosarles palabras clave, y luego facilitar su búsqueda mediante la recuperación de todos los segmentos identificados por el mismo código o por alguna combinación de códigos, mostrándolos junto con el texto al que están asociados. Los códigos pueden ser adosados de varias formas, en el sentido de que puede asociarse un código o varios a un segmento o a segmentos superpuestos.

Este tipo de software facilita y robustece desarrollos teóricos gracias al tratamiento de códigos adosados a segmentos de texto mediante su agrupamiento y la formación de un conjunto de categorías conceptuales interrelacionadas. Es decir, el uso de las herramientas informáticas apropiadas permite al analista ir más allá de la simple codificación de los temas relevantes en los datos. De hecho, los códigos se transforman en categorías teóricas que surgen de los datos, relacionadas de forma compleja.

Cabe aclarar, sin embargo, que algunas aproximaciones cualitativas podrían ser asistidas mejor por un software no específicamente desarrollado como CAQDAS; si el investigador requiere solamente algunas funciones específicas podría utilizar programas de índole más general. De hecho, Microsoft Word ha sido sugerido en varios casos como una alternativa de los CAQDAS.

8. ¿Cuáles son las funciones destacables de los CAQDAS?

En este apartado se describen las funciones más destacables de los CAQDAS. Si bien no todos los programas ofrecen la totalidad de las funciones, es conveniente que el investigador conozca cuáles son las herramientas que puede tener a su alcance. Como complemento se muestra, en el anexo 4, un cuadro comparativo de las principales fun-

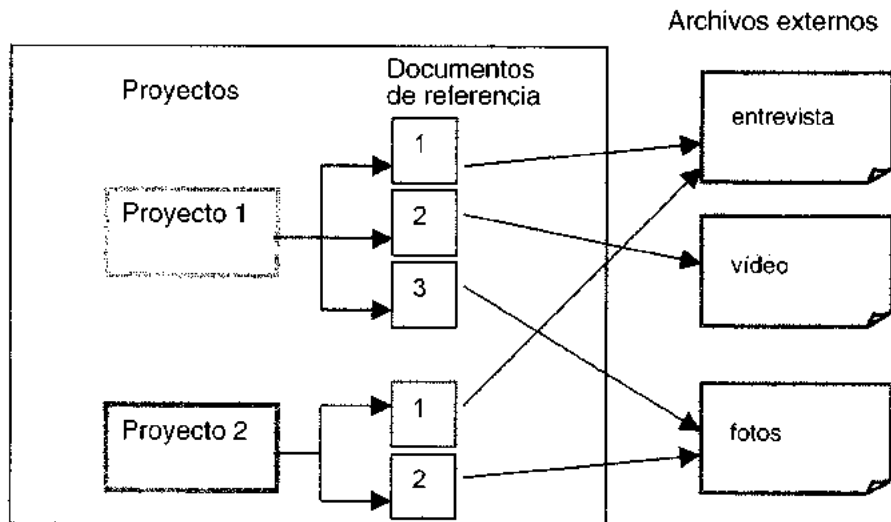
ciones desarrolladas a continuación entre los principales software que se comercializan actualmente.

8.1. Organización del trabajo (proyecto y datos)

Este tipo de software permite básicamente la creación de un «contenedor» de los diferentes archivos de datos que constituyen el proyecto de investigación. Este es nombrado como «proyecto», «unidad hermenéutica», «estudio», etc., según los distintos paquetes de programas. El proyecto es considerado como un «todo» en sí mismo. En algunos casos, estos archivos de datos primarios (entrevistas, notas de campo, vídeos, etc.) pueden permanecer dentro del contenedor (*MAXQDA*, *WSR NB*, *Nvivo*, *QDA Miner*) o afuera (*Atlas.ti*, *HyperRESEARCH*, *QUALRUS*, *TRANSANA 2*).

Los programas con contenedores «internos» engloban las bases de datos en su totalidad, mientras que los externos simplemente guardan en él las referencias a los archivos de datos (fig. 7.1). En ambos casos, el usuario mantiene una estrecha relación con sus datos, sin importar si los archivos están dentro del contenedor propiamente dicho o en su lugar de origen. La ventaja de las bases internas es que pueden ser transportadas de una computadora a otra fácilmente; la virtud de las bases externas consiste en que permiten manejar un mayor número de datos, particularmente si se trata de datos multimediales, y el tamaño del archivo del proyecto generado por el software será mucho menor que si incluyera «físicamente» los archivos de datos. Esto posibilita tener un gran número de datos (material textual o de otros tipos)

Figura 7.1. Archivos externos al proyecto



relacionado con el proyecto sin resentir la velocidad de procesamiento o consumir una enorme cantidad de recursos de memoria. Esta organización de los datos ayuda al investigador a estar en estrecho contacto con ellos y concentrarse en distintos subconjuntos de datos para su análisis.

8.2. Manejo de distintos tipos de datos cualitativos

Las entrevistas, las notas de campo, los distintos tipos de documentos de texto, gráficos, de audio y vídeo son formas típicas de datos cualitativos. Existen distintas maneras de recolectar los diferentes tipos de datos. Si bien la transcripción de las entrevistas o notas de campo es el modo más usual de encauzar el análisis, esta no es la única opción.

Algunas formas de análisis utilizan directamente la grabación sonora. El proceso de grabación convencional con casetes incorpora ruidos que pueden desmejorar la calidad de la transcripción y provocar la pérdida de palabras. Hoy en día, los grabadores digitales reducen notablemente el riesgo de perder información. Existen programas sofisticados (algunos se pueden obtener vía Internet sin costo) que permiten editar el sonido digital y almacenar los archivos de audio en CD-ROM. Aprovechan desde los grabadores y filmadoras analógicas hasta los grabadores y videocámaras digitales que no solo ofrecen una alta calidad de sonido e imagen sino que permiten transferir rápidamente la información a la computadora, almacenándola en una base de datos para su posterior análisis. Otra forma de recolectar datos es a través de la World Wide Web, gracias a la utilización de un software que permite realizar entrevistas en profundidad en tiempo real usando Internet como medio de comunicación, grabando directamente la entrevista en un archivo (Chen y Hinton, 1999). Este método resulta esencialmente útil para facilitar la tradicional entrevista cara a cara o la entrevista telefónica cuando el costo se hace totalmente prohibitivo por las distancias. Aquí se pone de manifiesto la importancia de la *convergencia digital* entre las computadoras y las telecomunicaciones como valor agregado en la investigación social. Con la proliferación de aplicaciones asociadas con la World Wide Web, los investigadores, estudiantes y académicos comenzaron a utilizar cada vez más la tecnología de las telecomunicaciones.

Volviendo al tema de este apartado, algunos programas permiten solamente el tratamiento de archivos de datos textuales (*MAXqda*, *WSR N6*, *NVivo*, *QDA Miner*), mientras que otros pueden operar con archivos multimediales (texto, gráfico, audio y vídeo), como el *Atlas.ti*, *HyperRESEARCH*, *QUALRUS*, *TRANSANA 2* (figs. 7.2 y 7.3).

8.3. Incorporación de documentos

La entrada de datos puede realizarse directamente en el sistema y/o a través de la importación de datos vía procesadores de textos. En general, los documentos textuales pueden ser asignados al sistema de manera directa desde un procesador de textos; algunos programas requieren algún tipo de conversión previa, como transformar archivos con formato al tipo «solo texto» o «rtf» [*rich text format*]. Los archivos multimedia, en los programas que los aceptan, son asignados directamente. Los tipos de archivo más utilizados son .jpg, .jpeg, .gif, .bmp, .wav, .avi y .mp3, entre otros.

Independientemente del software a utilizar, es conveniente seguir algunas reglas básicas previas a la transcripción del material textual compuesto por distintas unidades de texto fácilmente identificables, tales como el párrafo, la oración, la frase o la palabra. Estas reglas constituyen un elemento útil, particularmente en el momento de efectuar tareas como la búsqueda de texto (véase Anexo 1).

8.4. Codificación

La codificación consiste en adosar códigos a segmentos de textos, identificando y diferenciando así unidades de significado. Los criterios para identificar estas unidades pueden ser diversos: espaciales, temporales, temáticos, conversacionales y hasta gramaticales (párrafos, oraciones, líneas y palabras). Esos criterios están profundamente ligados a la pregunta de investigación.

Todos los programas poseen la función de codificar. Los códigos llamados índices o categorías conceptuales consisten generalmente en palabras o frases que indican un concepto abstracto. En algunos casos, el segmento de texto se limita a una palabra, a líneas o a párrafos; en otros no existen limitaciones. Los códigos se muestran en el margen del texto o dentro del texto mismo.

La longitud de los códigos es variable según el programa. Estos pueden ser adosados a fragmentos de textos superpuestos o incluidos entre ellos según el nivel de codificación permitido. Más de un código puede ser asignado a un mismo fragmento, quedando en manos del investigador la estrategia de codificación a utilizar (figs. 7.2, 7.3 y 7.6). Algunos programas incluyen la codificación *in vivo* (véase cap. 4) y la codificación rápida, la cual permite asignar un código activo a distintos segmentos de textos [*quick coding*]. El proceso de codificación también comprende la recodificación y la agrupación de los códigos. Asimismo, el esquema de codificación puede ser jerárquico, no jerárquico o ambos.

El siguiente ejemplo ilustra la utilización del software *Atlas.ti* en una investigación realizada en la Universidad de Buenos Aires, dentro

del área de Sociología del Trabajo y de Sociología de la Vejez, denominada: «Las representaciones sociales de los estudiantes de la Carrera de Relaciones del Trabajo de la Universidad de Buenos Aires con respecto a los trabajadores de mayor edad» (Oddone y Chernobilsky, 2003).

Ejemplo 1: Teniendo en cuenta que el envejecimiento [*ageism*] es un problema complejo con dimensiones históricas, culturales, sociales, psicológicas e ideológicas muy estudiadas desde una perspectiva gerontológica, es que en este trabajo se considera el *ageism* como los prejuicios y discriminaciones aplicados a los trabajadores de mayor edad (45 y más), por su edad. Se asume que ellos serían portadores de características negativas (ausentismo, lentitud en el aprendizaje, incapacidad de incorporar nuevas tecnologías, etc.) habiendo pocas excepciones a la regla, y se supone que tanto los trabajadores mayores como los jóvenes internalizan esta imagen negativa de la vejez.

Las consecuencias del *ageism* en el empleo pueden ir desde el rechazo a emplear trabajadores viejos hasta su despido, pasando por la negativa a promoverlos o capacitarlos y el incentivo para que acepten el retiro anticipado. A esto se suman los datos estadísticos que nos indican que los trabajadores de mayor edad desempleados permanecen en esta situación por mucho más tiempo que los más jóvenes, y cuando consiguen nuevo empleo, si alguna vez lo hacen, es con una remuneración mucho menor.

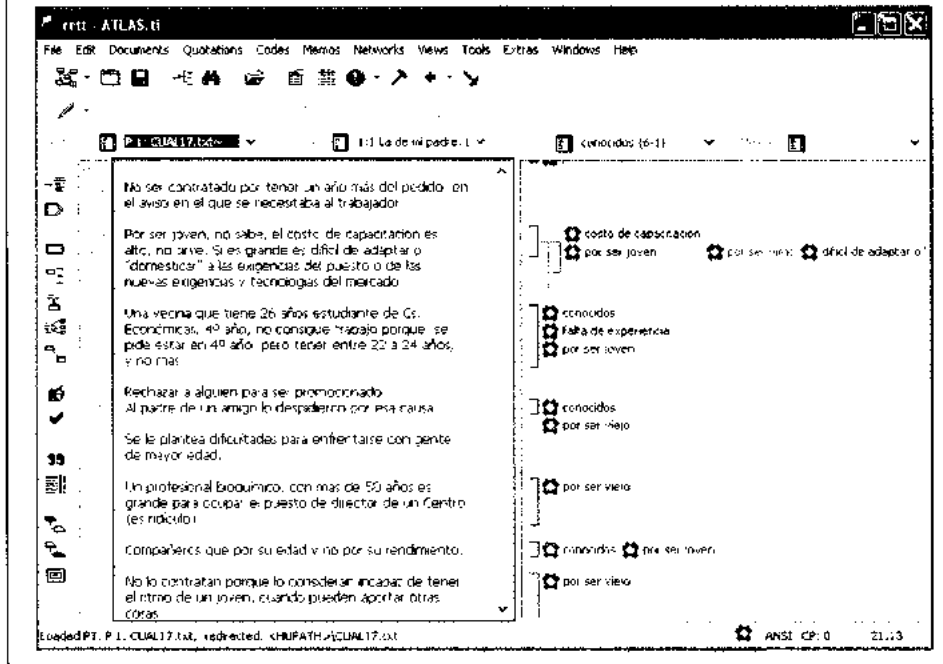
El objetivo de este estudio, basado en una combinación de estilos de investigación, es:

- conocer las representaciones sociales que tienen los estudiantes de la carrera de Relaciones del Trabajo de la Universidad de Buenos Aires como futuros decisores de políticas de Recursos Humanos, en relación a los trabajadores de mayor edad;

- establecer el grado en el cual los jóvenes experimentan *ageism* en el empleo y son en sí mismos discriminadores en sus actitudes y creencias acerca de los empleados.

Se realizó una encuesta a los estudiantes de la carrera con varias preguntas abiertas. Entre ellas se quiso saber si los alumnos conocían casos de discriminación en el trabajo por edad. Las respuestas fueron analizadas como material cualitativo.

Figura 7.2. Codificación de material cualitativo textual con el software *Atlas.ti*

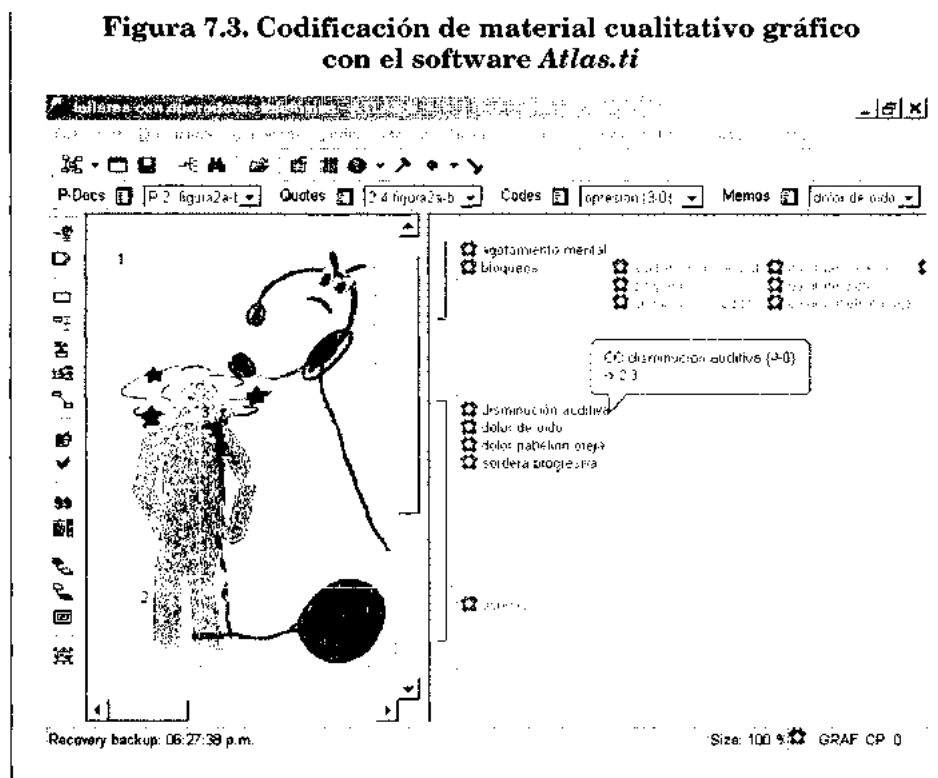


Algunos programas para el análisis de datos cualitativos, haciendo uso de la riqueza de información acumulada en la base de conocimiento del proyecto, identifican patrones en los datos durante el proceso de codificación, que luego son sugeridos como posibles códigos al investigador. Este software se basa en estrategias computacionales de inteligencia artificial. Es posible también codificar material gráfico, sonoro y de vídeo (fig. 7.3).

El siguiente ejemplo ilustra la utilización del software *Atlas.ti* en la investigación ya mencionada en el capítulo 2, *Telegestión: su impacto en la salud de los trabajadores* (Neffa et al., 2001).

Ejemplo 2: Entre las diversas técnicas utilizadas para captar cómo eran y cómo eran percibidas las condiciones de trabajo de los operadores telefónicos, se muestra a continuación el material extraído de los talleres de reflexión que se iniciaban con los interrogantes: «¿cómo siento mi cuerpo trabajando?» y «¿cómo veo mi puesto de trabajo?». Los participantes expresaron mediante representaciones gráficas sus vivencias sobre las condiciones de trabajo durante el desarrollo de sus actividades. Puede observarse en la siguiente imagen cómo perciben la carga global de trabajo y su influencia en la salud.

Figura 7.3. Codificación de material cualitativo gráfico con el software *Atlas.ti*



8.5. Autocodificación

Los CAQDAS permiten la codificación automática controlada por el investigador. Esta tarea consiste en realizar una búsqueda en los documentos, seleccionar una determinada extensión de texto cuando se cumpla la condición de búsqueda y relacionar ese texto con un código existente. La facilidad con que esta tarea es ejecutada por el software suele provocar una excesiva fragmentación de los datos, que contribuirá a dificultar luego un análisis más profundo. El investigador se ve involucrado en un esquema de codificación tan extenso, dada la magnitud de veces que se repite ese código a lo largo de su material, que «entra en pánico» y no sabe cómo obtener una explicación coherente de todo su análisis. Sin embargo, se supone que para un investigador formado y experimentado esta función no debería presentar tales problemas.

8.6. Búsqueda de texto

La recuperación de segmentos de textos constituye la principal función de los CAQDAS, ya sea a través de la búsqueda de texto o de códigos. En este sentido, es posible explorar los datos con una herramienta de búsqueda que permita encontrar un conjunto de caracteres (pa-

labras o frases) en todos los documentos de texto. Equivaldría a la función de búsqueda de un procesador de textos, pero en algunos casos permite búsquedas más complejas como la *booleana* (*and, or, not*) (véase Anexo 2), la de múltiples palabras, la de palabras con raíces comunes, etc., además de búsquedas autodefinidas como las *Globally look for Regular Expression and Print* [GREP]. Este tipo de búsqueda se utiliza cuando es necesaria una mayor flexibilidad o amplitud de criterios que la proporcionada por la búsqueda estándar.

La base de la búsqueda GREP es la inclusión de caracteres especiales dentro de la cadena de caracteres que controla el proceso de búsqueda de coincidencias. Por ejemplo, si se quieren recuperar las distintas intervenciones del entrevistador y del entrevistado buscándolas por sus respectivos nombres con la condición de que aparezcan al principio de una línea, se utilizarán símbolos específicamente definidos para la ejecución de esta tarea. Del mismo modo, en el supuesto caso de tratarse de una historia de vida (véase cap. 5), si se quisieran identificar todas las emisiones que se refieren a un mismo intervalo de tiempo, por ejemplo acontecimientos en la narrativa del actor que tuvieron lugar entre los años 1990 y 1999, es posible la utilización de caracteres específicos para identificar los distintos años.

Por otra parte, algunos programas permiten obtener un listado indexado de frecuencia de palabras mediante una herramienta de conteo.

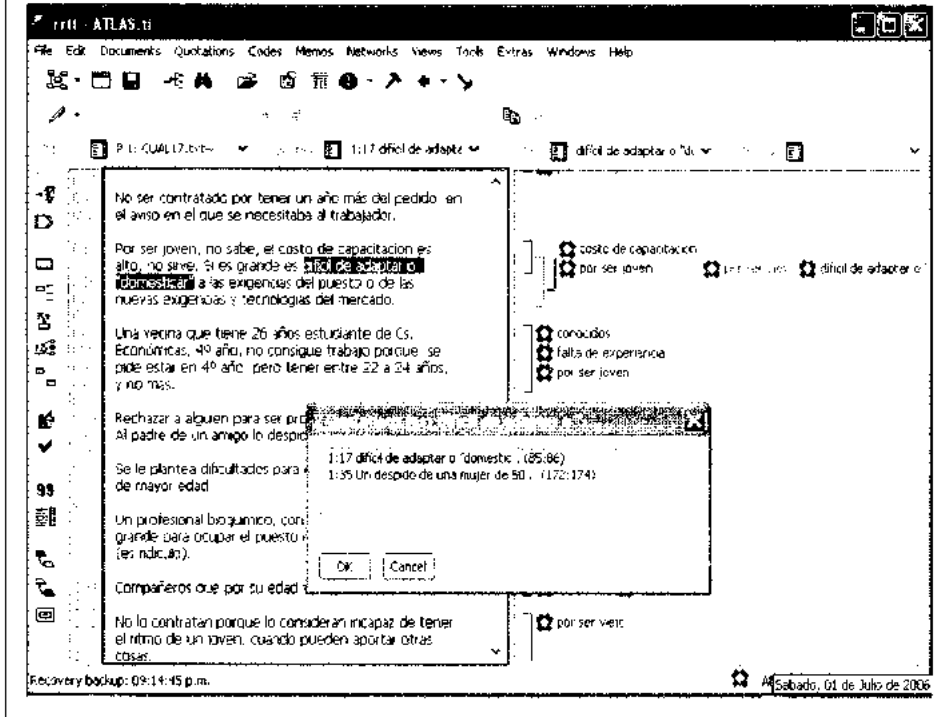
8.7. Búsqueda de códigos

Una vez codificados los datos, es posible recuperar todos los segmentos de texto de un conjunto de documentos que fueron codificados con el mismo código, para poder buscar relaciones entre fragmentos de textos y explorar las preguntas de la investigación mediante comparaciones e identificación de temas. Esta función básica de recuperación está presente en la mayoría de los software (fig. 7.4). Muchos programas complementan la búsqueda utilizando combinaciones de códigos. Se pueden combinar dos o más términos en una expresión, utilizando operadores de búsqueda que indiquen la relación que se desea imponer entre ellos. Generalmente incluyen la búsqueda *booleana* (*and, or, not*) y la de proximidad (*precede a, superpuesta a, cerca de*) (véase Anexo 2).

Una de las herramientas más potentes que contienen estos programas es la posibilidad de imprimir todos estos fragmentos de textos sin que se alteren los documentos originales.

Esta figura muestra la recuperación de citas relevantes codificadas en el proyecto mencionado en la figura 7.2, ejemplo 1.

Figura 7.4. Recuperación de segmentos de texto codificados utilizando el software *Atlas.ti*



Sobre ellos puede aplicarse la recuperación selectiva mediante la utilización de filtros (por ejemplo «respondentes del mismo sexo») o a través de agrupaciones, como se explica en el punto 8.9.

8.8. Incorporación y recuperación de memos o anotaciones

Se entiende por memos todas aquellas anotaciones que realiza el investigador durante el proceso de análisis (notas recordatorias, hipótesis de trabajo, descripciones ampliadas de los códigos, material bibliográfico, metodología propuesta, informes de avance, conclusiones, etc.), que en su conjunto pueden constituir el punto de partida en la redacción de un informe de resultados. Usados en forma sistemática, constituyen otra forma que tiene el analista de entrar en contacto con sus datos.

La interfaz para la escritura de memos generalmente consiste en documentos en blanco con las funciones de un procesador de textos. Es-

tos documentos producidos con el CAQDAS pueden ser clasificados y relacionados con códigos y/o segmentos de textos. También pueden ser codificados.

La figura 7.5 ilustra la utilización de memos en el ejemplo 2 citado anteriormente.

Figura 7.5. Uso de memos con el software *Atlas.ti*

El memo «dolor de oído» muestra las causas señaladas por los operadores –ruidos molestos, interferencias, excesivo tiempo de atención, abuso intensivo y permanente del microteléfono–, que según ellos provocan sus problemas auditivos, tales como: disminución auditiva, dolor de oído, sordera progresiva, etcétera.

8.9. Variables/atributos y agrupaciones

La función de agrupación facilitada por los CAQDAS permite crear familias de unidades de análisis que comparten el mismo valor del atributo y/o variable. Por ejemplo, en una investigación referida a la situación de extrema pobreza urbana cuya unidad de análisis es la familia, dichas unidades se pueden agrupar según sea que el jefe del hogar esté a cargo de hombres o mujeres, según que el jefe del hogar esté ocu-

pado o desocupado, según que el tiempo de desocupación del jefe del hogar y/o cónyuge sea más o menos de un año, según sea el número de hijos por hogar, etc. Tanto variables socio-demográficas como cualquier tipo de variable, sea numérica, categórica, lingüística, lógica o de fechas pueden ser usadas para categorizar casos.

8.10. Sistemas de hipertextos

El hipertexto es un vínculo entre diferentes partes de la base de datos –códigos, memos, documentos– que permite mostrar una conexión significativa y lógica entre segmentos de textos sin recurrir a los códigos. Se trata de una forma de organizar y vincular información de manera no secuencial, que permite al investigador moverse rápidamente entre distintas fuentes de información. Los CAQDAS también hacen posible definir el vínculo (explicación, justificación o contradicción) entre distintos elementos. El carácter dinámico, asociativo y no secuencial del hipertexto se asimila a los procesos heurísticos e iterativos propios de la investigación cualitativa. El ejemplo más popular de un hipertexto es la red Internet (World Wide Web) (véase Anexo 3).

8.11. Confección de mapas conceptuales o redes semánticas

Los mapas conceptuales o redes semánticas son herramientas para organizar y representar conocimiento de manera intuitiva mediante representaciones gráficas, que incluyen conceptos y relaciones entre ellos en forma de proposiciones. Se entiende por concepto a «la regularidad percibida en hechos u objetos, o registros de hechos u objetos, designada con un nombre» (Novak y Cañas, 2006). Las proposiciones son sentencias acerca de un objeto o hecho en el universo, ya sea natural o construido que contienen dos o más conceptos conectados que forman una sentencia significativa.

Los programas permiten vincular entre sí distintos elementos tales como fragmentos de textos, códigos, memos y documentos de datos, en forma gráfica mediante símbolos (líneas y flechas), atribuyéndoles diferentes significados, lo cual posibilita usar las redes para construir modelos de los datos. En general, los CAQDAS agilizan la construcción de un listado de códigos que puede ser organizado de manera jerárquica o de otras formas y la visualización de la red de vínculos mediante un diagrama.

Los investigadores pueden navegar fácilmente dentro de esta red y seleccionar diferentes vistas para destacar diversas características teóricas (figs. 7.7 y 7.8). Así, las redes semánticas facilitan la producción de predicciones teóricas y explicaciones basadas en códigos.

El siguiente ejemplo ilustra la utilización del software *Atlas.ti* en una investigación realizada en el ámbito del CONICET titulada: «Vejez,

pobreza y vida cotidiana. Un estudio de casos en el área metropolitana de Buenos Aires» (Oddone, 2001).

Ejemplo 3: El objetivo de este trabajo fue estudiar en profundidad las estrategias de supervivencia que desarrollan los ancianos que habitan en un contexto de pobreza para satisfacer las necesidades de la vida cotidiana.

Se trabajó exclusivamente con ancianos pobres, tanto estructurales como aquellos pauperizados. Se realizaron 51 historias de vida en la ciudad de Buenos Aires y el conurbano bonaerense.

Los objetivos específicos o temas estudiados fueron: datos demográficos, familia de origen y actual, migraciones, educación formal e informal, historia laboral, convivencia, vivienda, hacinamiento, hábitat y entorno, violencia familiar y social, familia a cargo/ancianos a cargo de otros, salud, servicios sociales, ingresos, alimentación, gastos, intercambios de redes sociales (capital social) e institucionales, tipo de ayudas, problemas y percepción, soledad-aislamiento-marginalidad, tiempo libre, sentimientos en relación a su situación, etcétera.

En la muestra intencional se tuvo en cuenta: edad (75 años y más; menos de 75 años), sexo, y situación de pobreza (estructurales y pauperizados). Se recogieron historias de vida mediante el abordaje de bola de nieve hasta la saturación de la muestra.

Figura 7.6. Entrevista codificada con la asistencia del software *Atlas.ti*

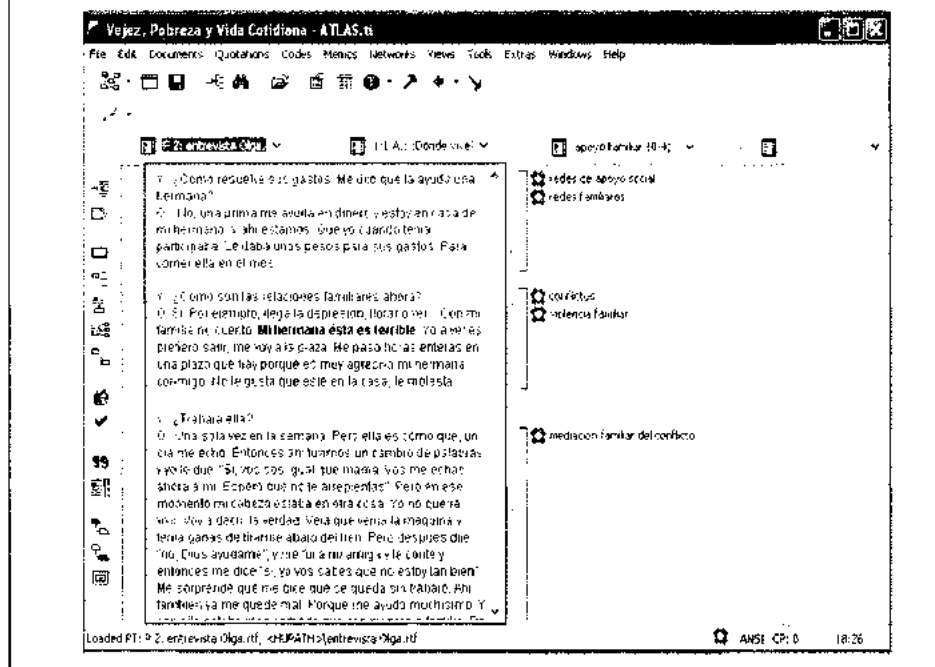
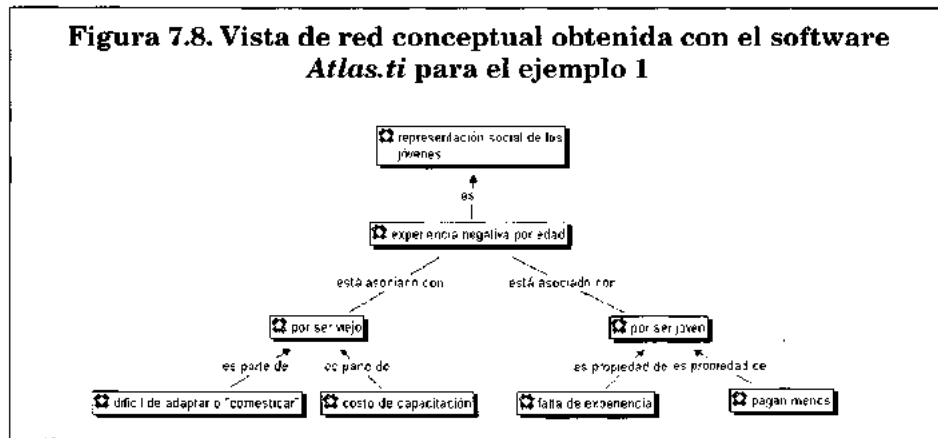


Figura 7.7. Vista de red conceptual obtenida con el software *Atlas.ti* para el ejemplo 3



Figura 7.8. Vista de red conceptual obtenida con el software *Atlas.ti* para el ejemplo 1



8.12. Construcción de teoría

El término «construcción de teoría» usado por la concepción metodológica de la «teoría fundamentada» (*Grounded Theory*) está relacionado con funciones avanzadas de codificación y recuperación. Se basa en el supuesto de que los códigos usados para organizar los datos representan las categorías teóricas que los investigadores desarrollan durante el proceso continuo de análisis. Puesto que una teoría puede ser considerada como una red de categorías, la herramienta que permite conectar códigos entre sí puede ser muy útil para mostrar la estructura de la teoría emergente, y por lo tanto el software que facilita esta función contribuirá favorablemente a su construcción. En este sentido, la herramienta para la confección de diagramas conceptuales ofrece una forma visual de representar las ideas y teorías que emergen de los datos que representan.

8.13. Interfaz con datos cuantitativos

Si bien los CAQDAS están pensados para asistir al investigador en procesos básicamente cualitativos, cuando se trabaja con una gran cantidad de datos a veces es útil su análisis estadístico. Las funciones más útiles para la integración de métodos cuantitativos en la investigación cualitativa son el conteo de palabras –que puede poner en evidencia ocurrencias útiles de códigos potenciales– y la exportación de datos, tales como frecuencias de códigos encontradas en todos los documentos, que puedan ser vinculados con datos cuantitativos y analizados estadísticamente con un software externo apropiado. La interfaz clásica permite la transformación de los datos de los programas CAQDAS para su utilización en el programa SPSS (*Statistical Package for Social Sciences*).

8.14. Reportes

La mayoría de los programas brindan la posibilidad de imprimir los resultados obtenidos durante el proceso de análisis, tales como listados de documentos de datos, de fragmentos de textos seleccionados, de códigos, memos, resultados de búsquedas y otros ítems. Estas listas se originan dentro del mismo programa, permitiendo de este modo obtener una vista previa de los resultados obtenidos antes de imprimirlos. También es factible exportar los listados para ser leídos por otras aplicaciones, como por ejemplo el procesador de textos Word.

8.15. Publicación en Internet

Esta función permite publicar resultados intermedios y/o finales en la red Internet gracias a una estructura prediseñada o *template*, lo cual permite a otros investigadores tener acceso a los archivos originales en el servidor vía la página web creada.

8.16. Trabajo en equipo y autoría

Esta función permite que más de un autor pueda trabajar en los mismos o en distintos documentos pertenecientes a un mismo proyecto, estableciendo sistemáticamente el registro de la producción de cada uno mediante su nombre, fecha y hora de realización. También facilita la posibilidad de integrar cada una de las partes del proyecto en un todo.

Conclusiones

Los CAQDAS son útiles para una amplia variedad de aproximaciones metodológicas, aunque probablemente no tan dilatada como los entusiastas de estos programas pregonan. Lo importante en cualquier tipo de análisis es ser consciente de que los CAQDAS son una herramienta más que un método en sí mismos. Para poder sacar provecho de este tipo de software, es esencial adaptar su uso a la propia metodología y no al revés.

Teniendo en cuenta las facilidades que ofrecen los CAQDAS para el almacenamiento, codificación y recuperación de grandes masas de datos textuales o de otros tipos, y en virtud de las funciones específicas detalladas, queda en manos del investigador utilizarlos de acuerdo con su propia estrategia metodológica sin someter por esto su trabajo a las particularidades de cada tecnología.

Cualquiera sea la herramienta elegida, el uso de un software específico puede beneficiar y enriquecer el proceso de análisis. Para finalizar y teniendo en cuenta que «la investigación cualitativa es la práctica de hacer preguntas simples y obtener respuestas complejas y que además el verdadero secreto para llevar a cabo exitosamente una investigación cualitativa consiste en el arte de manejar tanto la simplicidad como la complejidad» (Chenail, 1995), es importante que los investigadores reconozcan el valor de la herramienta informática sin magnificar en exceso su protagonismo en el complejo proceso de la investigación social; al mismo tiempo, es aconsejable que estén abiertos a las nuevas tecnologías para aprovechar las múltiples ventajas que estas proporcionan.

Anexo 1

Reglas sugeridas para transcribir el material textual para ser analizado por algún CAQDAS

Utilizar el procesador de palabras Microsoft Word para crear el documento o material textual que será analizado cualitativamente, asignándole al archivo un nombre descriptivo. Por ejemplo: nombre y apellido del entrevistado (Entrevista a José Pérez.doc), cargo ocupado por el entrevistado en la institución (Entrevista a Jefe de personal.doc), lugar de realización y número de entrevista (Entrevista 4 en Institución B.doc). El criterio adoptado tiene que ser relevante desde el punto de vista del diseño de la investigación y, a su vez, debe permitir la identificación fácil y rápida de los documentos que constituirán la materia prima del análisis (Chernobilsky y D'Onofrio, 2003).

En el documento de Word (.doc) creado:

- *presentar el tipo de material textual cualitativo* (señalar si se trata de una entrevista, una observación, una nota de campo, etc.), identificando su/s autor/es y la situación social en la que fue producido (según criterios normalizados). Esto es, acordar un formato común de encabezado que contenga la información considerada pertinente desde el punto de vista del diseño de la investigación. Por ejemplo: «Entrevista realizada por María Fernández a José Pérez, Jefe de Personal de la Institución B el 1 de septiembre de 2002»;
- *identificar* (según criterios normalizados) *a entrevistador/es y entrevistados/s, observador/es y observado/s* (según corresponda). Es decir, acordar un criterio de identificación homogéneo a lo largo del documento para los actores involucrados en el texto (por ejemplo, fijar si entrevistador y entrevistados se designarán por las iniciales de sus respectivos nombres y apellidos o por sus nombres de pila, etcétera);
- *separar*, mediante dos o más espacios tipográficos (↵), *los párrafos que indican distintos actores* (según criterios normalizados, tales como intervenciones de entrevistador/es seguidas de intervenciones de entrevistado/s, notas de observador/es y a continuación expresiones verbales literales de observado/s).

Anexo 2

Operadores booleanos

Gran parte de la búsqueda en las bases de datos se apoya en los principios de la lógica *booleana*. Estos hacen referencia a las relaciones lógicas existentes entre los términos de búsqueda, las cuales tomaron su nombre del matemático británico George Boole. La lógica *booleana* consiste en la aplicación de tres operadores lógicos: «o» (*or*), «y» (*and*) y «no» (*not*).

El operador lógico «o» (or)

Código A or Código B: El operador lógico «*or*» ayuda a recuperar todos los segmentos codificados con, por lo menos, uno de los dos códigos.

El operador lógico «y» (and)

Código A and Código B: El operador lógico «*and*» permite recuperar todos los segmentos codificados con ambos códigos.

El operador lógico «no» (not)

Not Código A: El operador lógico «not» excluye todos los segmentos codificados con el código especificado.

Operadores de proximidad

Los operadores de proximidad posibilitan buscar segmentos codificados en función de la relación espacial existente entre diferentes códigos. Es decir, tendrá en cuenta códigos separados por un número determinado de líneas de texto dentro de un documento.

El operador de proximidad «cerca»

El operador de proximidad «cerca» permite buscar por términos separados por un número determinado de líneas de texto en cualquier orden: *Código A cerca Código B (n líneas)*.

El operador de proximidad «seguido»/«precedido»

El operador de proximidad «seguido» permite buscar por términos separados por un número concreto de líneas de texto y en un orden predeterminado: *Código A seguido Código B (n líneas)*; *Código A precedido Código B (n líneas)*.

Anexo 3

Hipertexto

El hipertexto se basa en la escritura y en la lectura no secuencial de los documentos. Esto supone la existencia y la adopción del criterio de asociación de ideas y conceptos como principio organizador del conjunto de documentos, el conjunto informativo (Landow, 1997). Este tipo de organización no puede reflejarse adecuadamente en material impreso, ya que en este predomina la linealidad. Sin embargo, existe la posibilidad de superar esa linealidad, utilizando como herramientas sistemas informáticos capaces de establecer relaciones entre archivos.

Dos elementos clave en estos sistemas son los de «enlace» y «nodo». Un nodo es un documento o elemento informativo reproducido en un soporte informático. Entre ellos pueden establecerse relaciones y asociaciones a través de enlaces. De esta forma, dos documentos que contengan conceptos relacionados ofrecerán acceso directo entre uno y otro, o a la parte que sea pertinente. Por otra parte, es viable establecer varios caminos posibles para acceder a la misma información.

Hipermedia

El panorama esbozado es perfectamente comprensible desde una perspectiva textual, en la que los nodos corresponden a documentos con información escrita. Sin embargo, la información y la documentación actual incorporan gran cantidad de componentes gráficos (estáticos o dinámicos), sonidos u otros complementos. Gracias a la evolución de la tecnología, estos documentos no textuales pueden ser integrados en formato digital a los sistemas informáticos. Como consecuencia, los sistemas hipertextuales han evolucionado hasta disponer de las capacidades necesarias para incluir en ellos documentos e información con presentación gráfica y/o sonora. Esta expansión de las funciones de gestión de información a cualquier forma que pueda presentar, manteniendo como criterio la asociación o relación de conceptos, es lo que caracteriza a un sistema hipermedia (Tramullas Saz, 1996).

Anexo 4

Existe actualmente una gama enorme de software para el análisis de datos cualitativos y, teniendo en cuenta las demandas de los investigadores, los autores y desarrolladores de los mismos introducen, con cierta frecuencia, nuevas funciones y características a los programas que aquellos necesitan aplicar. Esta variedad da cuenta de la necesidad de establecer patrones comunes para combinar e intercambiar datos durante el proceso del análisis cualitativo.

En las siguientes páginas se muestra un estudio comparativo de los principales programas CAQDAS disponibles y sus características.

- *Atlas.ti* (versión 5.2), desarrollado en la Universidad de Berlín por el computador científico Thomas Muhr, <http://www.atlasti.de>.
- *HyperRESEARCH* (versión 2.6), desarrollado por la empresa ResearchWare Inc., <http://www.researchware.com>.
- *MAXqda*, desarrollado originalmente por Udo Kuckartz, <http://www.maxqda.com>.
- *The Ethnograph* (versión 5.08) desarrollado por el sociólogo John Seidel, en 1985, <http://www.qualisresearch.com/default.htm>.
- *Nvivo* (versión 7), emparentado con N6, desarrollados ambos por la empresa QSR. Su antecesor es Nudist, creado por Tom Richards, <http://www.qsrinternational.com>.
- *QUALRUS* (versión 3), desarrollado por Idea Works, Inc., utiliza estrategias de inteligencia artificial, <http://www.qualrus.com>.
- *QDA Miner* (versión 2.0), desarrollado y comercializado por Provalis Research, <http://www.provalisresearch.com/QDAMiner/QDAMinerDesc.html>.

Atlas.ti v. 5.2	HyperRESEARCH v. 2.6	MAXqda	The Ethnograph v. 5.08	QSR N6 (ex Nudist)	Nvivo v. 7	Qualrus v. 3	QDAMiner v. 2.0
Organización del trabajo							
<ul style="list-style-type: none"> • Unidad hermenéutica • Estructura de base de datos externa 	<ul style="list-style-type: none"> • Estudio • Estructura de base de datos externa 	<ul style="list-style-type: none"> • Proyecto • Estructura de base de datos interna 	<ul style="list-style-type: none"> • Proyecto • Estructura de base de datos externa 	<ul style="list-style-type: none"> • Proyecto • Estructura de base de datos interna 	<ul style="list-style-type: none"> • Proyecto • Estructura de base de datos interna 	<ul style="list-style-type: none"> • Estructura de base de datos externa 	<ul style="list-style-type: none"> • Proyecto • Estructura de base de datos interna
Tipos de datos cualitativos							
Datos textuales, vídeo digital, sonido y gráficos	Datos textuales, vídeo digital, sonido y gráficos	Datos textuales	Datos textuales	Datos textuales	Datos textuales	Datos textuales, vídeo digital, sonido y gráficos	Datos textuales
Incorporación de documentos							
Texto (.txt, .rtf, .doc), gráficos (.jpg, .jpeg, .bmp, .gif), audio (.wav, .mp3)	Texto (.txt), gráficos (.jpeg, .bmp, .gif, .png, .pict), audio (.wav, .aif, .mov, .mp3), vídeo (.avi, .mpeg, .mov, .swf, .gif)	Texto (.txt)	Texto (.txt)	Texto (.txt)	Texto (.txt, .rtf). Permite incorporar vínculos a archivos multimediales	Texto (.txt, .rtf), gráficos (.jpg, .bmp), audio (.mp3, .wav), vídeo (.avi)	Texto (.rtf, .doc, .html, .pdf, .txt)

Atlas.ti v. 5.2	HyperRESEARCH v. 2.6	MAXqda	The Ethnograph v. 5.08	QSR N6 (ex Nudist)	Nvivo v. 7	Qualrus v. 3	QDAMiner v. 2.0
Codificación							
<ul style="list-style-type: none"> • Mínima unidad de texto a codificar: un carácter. • Codificación no jerárquica, <i>in vivo</i>. • Longitud del código: sin límites. • Nivel de codificación: sin límites. 	<ul style="list-style-type: none"> • Mínima unidad de texto a codificar: un carácter. • Codificación no jerárquica, <i>in vivo</i>. • Longitud del código: 64 caracteres. • Nivel de codificación: difícil de apreciar. 	<ul style="list-style-type: none"> • Mínima unidad de texto a codificar: un carácter. • Codificación jerárquica, <i>in vivo</i>. • Longitud del código: 64 caracteres. • Nivel de codificación: hasta 10. 	<ul style="list-style-type: none"> • Mínima unidad de texto a codificar: una línea (42 caracteres). • Codificación jerárquica, <i>in vivo</i>. • Longitud del código: 10 caracteres. • Nivel de codificación: 7. 	<ul style="list-style-type: none"> • Mínima unidad de texto a codificar: una unidad de texto o oración o párrafo). • Codificación jerárquica, <i>in vivo</i>. • Longitud del código: sin límites. • Nivel de codificación: sin límites. 	<ul style="list-style-type: none"> • Mínima unidad de texto a codificar: un carácter. • Codificación jerárquica, <i>in vivo</i>. • Longitud del código: sin límites. • Nivel de codificación: sin límites. 	<ul style="list-style-type: none"> • Mínima unidad de texto a codificar: una unidad de texto. • Codificación no jerárquica, <i>in vivo</i>. 	<ul style="list-style-type: none"> • Codificación jerárquica.
Autocodificación							
Permite	Permite	Permite	No permite	Permite	Permite	Permite (pero poco flexible)	Permite

Atlas.ti v. 5.2	HyperRESEARCH v. 2.6	MAXqda	The Ethnograph v. 5.08	QSR N6 (ex Nudist)	Nvivo v. 7	Qualrus v. 3	QDAMiner v. 2.0
Búsqueda de texto							
Cadena de caracteres, categorías, GREP	Palabras o frases.	Palabras o frases.	Cadena de caracteres.	Cadena de caracteres, categorías, GREP.	Cadena de caracteres, categorías, GREP.	Cadena de caracteres.	Cadena de caracteres.
Meimos							
Permite	Permite	Permite	Permite	Permite	Permite	Permite	Permite
Búsqueda de códigos							
Operadores booleanos, semánticos y de proximidad.	Operadores booleanos y de proximidad.	Operadores booleanos y de proximidad.	Operadores booleanos y de proximidad.	Operadores booleanos y de proximidad.	Operadores booleanos y de proximidad.	Operadores booleanos y semánticos.	Operadores booleanos.
Variables/atributos y agrupaciones							
Permite	Permite	Permite	Permite	Permite	Permite	Permite	Permite
Sistema de hipertexto							
Permite	No permite	No permite	No permite	No permite	Permite (básico)	Permite (básico)	No permite
Confección de mapas conceptuales							
Permite	Permite	No permite	No permite	Limitada	Permite	Permite	No permite

Atlas.ti v. 5.2	HyperRESEARCH v. 2.6	MAXqda	The Ethnograph v. 5.08	QSR N6 (ex Nudist)	Nvivo v. 7	Qualrus v. 3	QDAMiner v. 2.0
Construcción de teoría							
Sí	Sí	No	No	No	Sí	Sí	No
Interfaz con datos cuantitativos							
Permite frecuencia de palabras y exportación a software estadístico.	Permite frecuencia de códigos y exportación a Excel.	Permite frecuencia de palabras y exportación a software estadístico.	No permite.	No permite.	Permite frecuencia de palabras y exportación a software estadístico.	No permite.	Permite exportación a software estadístico.
Reportes							
Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Trabajo en equipo							
Sí	No	Sí	Sí	Sí	Sí	Sí	Sí
Requerimiento de sistema: plataforma							
PC y Apple Macintosh	PC y Apple Macintosh	PC	PC	PC	PC	PC	PC

C. EJERCICIOS

1. Complete el siguiente cuadro con los tipos de muestra:

Muestras probabilísticas	Muestras no probabilísticas

2. La Comisión de Derechos Humanos le pide asesoría para levantar una encuesta entre sus trabajadores que permita detectar necesidades de capacitación; se desea saber qué es más conveniente, aplicar la encuesta a los 525 trabajadores o seleccionar una muestra representativa, ¿qué les aconsejaría? Argumente su respuesta de acuerdo con las lecturas realizadas en el módulo.

3. Usted está a cargo de una encuesta mediante la cual se pretende conocer las percepciones de la población de lesbianas, gays, bisexuales y transgénero de la ciudad de México sobre el respeto a sus derechos humanos. Considerando que se desconoce el tamaño de la población, ¿qué procedimiento llevaría cabo para seleccionar la muestra?

4. Un investigador obtuvo, en una muestra de los habitantes del Distrito Federal, las siguientes frecuencias absolutas para la variable "qué significa ser ciudadano":

Categoría	Frecuencias absolutas
Tener derechos y obligaciones	94
Poder votar	75
Tener responsabilidad	63
Pertenecer a un país	55
Haber cumplido 18 años	49
Tener educación política	36
Todas las anteriores	26
Ninguna/otra	18

- a) Calcule las frecuencias relativas
- b) Interprete los resultados respondiendo a la pregunta ¿qué significa para los habitantes del Distrito Federal ser ciudadano?

5. Un investigador obtuvo, en una muestra de empleados de la CDHDF, los siguientes resultados al medir el "orgullo de trabajar de la Comisión". La escala oscilaba entre 0 (nada de orgullo) y 10 (orgullo total)

Máximo: 10
Mínimo: 0
Media: 8.6
Moda: 8.0
Mediana: 8.2
Desviación estándar: 0.6

¿Qué puede decirse de esta muestra acerca del orgullo por trabajar en la CDHDF?

6. Con base en la lectura de Jacinto Rodríguez, señale las diferencias entre el muestro aleatorio simple y el muestreo aleatorio sistemático.

7. De acuerdo con Jacinto Rodríguez responda, ¿cuáles son las ventajas y desventajas del muestreo por cuotas?

8. Con base en la lectura de Roberto Hernández, Carlos Fernández y Pilar Baptista explique en qué consisten las medidas de variabilidad.

9. Según Jacinto Rodríguez, ¿cuál es la relación entre la varianza poblacional y el tamaño de la muestra?

10. ¿Por qué es importante considerar el nivel de medición de las variables al momento de analizar los datos?

D. AUTOEVALUACIÓN

1. En este tipo de muestreo la selección de los elementos de la muestra se hace en una sola etapa. Equivale a utilizar el censo de la población objeto del estudio, para obtener al azar los elementos que van a formar parte de la muestra.

- a) Muestreo aleatorio simple.
- b) Muestreo aleatorio sistemático.
- c) Muestreo aleatorio estratificado.
- d) Muestreo aleatorio por cuotas.

2. Uno de los siguientes tipos de muestro no es probabilístico, señale cuál.

- a) Muestreo aleatorio estratificado.
- b) Muestreo por conglomerados.
- c) Muestro por cuotas.
- d) Muestro aleatorio sistemático.

3. De acuerdo con Rodríguez Osuna, es recomendable hacer una selección probabilística de la muestra cuando...

- a) se quieren realizar estudios de opinión.
- b) se hacen estudios de mercado.
- c) se pretende hacer una exploración del problema.
- d) se trata de encuestas que se toman como base para decisiones importantes.

4. Con base en la lectura de Rodríguez Osuna, señale cuál es la relación entre el tamaño de la muestra y la precisión de las estimaciones.

- a) Si aumenta el tamaño de la muestra lo hace el nivel de precisión de las estimaciones.
- b) Si aumenta el tamaño de la muestra, disminuye el nivel de precisión de las estimaciones.
- c) La relación entre el tamaño de la muestra y la precisión de las estimaciones es inversamente proporcional.
- d) El tamaño de la muestra no guarda relación alguna con la precisión de las estimaciones.

5. Según Roberto Hernández, Carlos Fernández y Pilar Baptista, ¿por qué es importante considerar el nivel de medición de las variables al momento de analizar los datos?

- a) Porque el nivel de medición varía en función del enfoque cualitativo o cuantitativo del análisis.
- b) Porque el análisis de los datos no es el mismo para todos los niveles de medición.
- c) Porque las variables nominales no pueden ser analizadas en términos cuantitativos.
- d) Porque en la estadística descriptiva las variables de razón no son analizadas.

6. Al conjunto de puntuaciones ordenadas en sus respectivas categorías se le llama:

- a) Varianza.
- a) Desviación estándar.
- a) Distribución de frecuencias.
- a) Distribución muestral.

7. Las frecuencias relativas son:

- a) Los porcentajes de casos en cada categoría.
- b) Los casos acumulados en cada categoría.
- c) Las puntuaciones ordenadas en cada categoría.
- d) Los promedios ordenados a partir de la media.

8. Al promedio aritmético de una distribución se le llama:

- a) Moda.
- b) Mediana.
- c) Media.
- d) Varianza.

9. De acuerdo con la lectura de Roberto Hernández, Carlos Fernández y Pilar Baptista, la media es una medida que sólo se puede aplicar a dos tipos de variables, ¿cuáles son éstas?

- a) Nominales y ordinales.
- b) De intervalo y de razón.
- c) Ordinales y de intervalo.
- d) De razón y ordinales.

10. ¿Cómo se interpreta la desviación estándar?

- a) Como "cuánto se desvía, en promedio, de la media un conjunto de puntuaciones".
- b) De acuerdo con el número de unidades en una distribución.
- c) Como la dispersión de los datos en una distribución.
- d) Como las distancias entre las escalas de medición.

CLAVES DE RESPUESTA

MÓDULO I

1. c
2. c
3. a
4. c
5. d
6. a
7. b
8. c
9. a
10. b

MÓDULO II

1. d
2. b
3. a
4. c
5. d
6. c
7. b
8. d
9. a
10. d

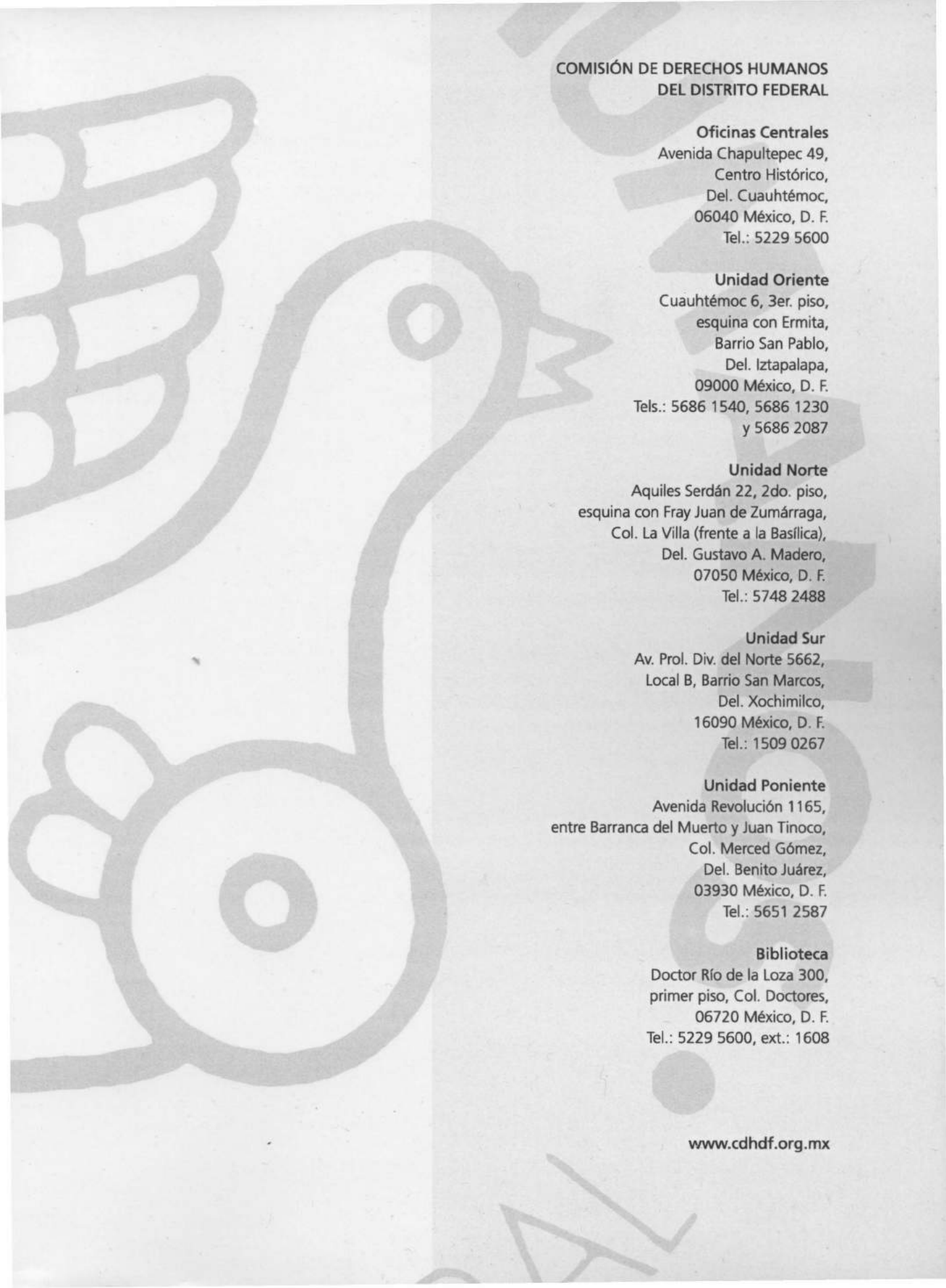
MÓDULO III

1. b
2. a
3. d
4. a
5. a
6. c
7. d
8. c
9. a
10. c

MÓDULO IV

1. a
2. c
3. d
4. a
5. b
6. c
7. a
8. c
9. b
10. a

Programa de Estudio de los Derechos Humanos, número 8 de los Cuadernos de Capacitación, se terminó de imprimir en abril de 2008 en los talleres de Solar, Servicios Editoriales, S. A. de C. V., Calle 2, núm. 21, San Pedro de los Pinos, 03800 México, D. F.
El tiro fue de 60 ejemplares impresos en papel bond de 75 g con la familia tipográfica Frutiger.



**COMISIÓN DE DERECHOS HUMANOS
DEL DISTRITO FEDERAL**

Oficinas Centrales

Avenida Chapultepec 49,
Centro Histórico,
Del. Cuauhtémoc,
06040 México, D. F.
Tel.: 5229 5600

Unidad Oriente

Cauhtémoc 6, 3er. piso,
esquina con Ermita,
Barrio San Pablo,
Del. Iztapalapa,
09000 México, D. F.
Tels.: 5686 1540, 5686 1230
y 5686 2087

Unidad Norte

Aquiles Serdán 22, 2do. piso,
esquina con Fray Juan de Zumárraga,
Col. La Villa (frente a la Basílica),
Del. Gustavo A. Madero,
07050 México, D. F.
Tel.: 5748 2488

Unidad Sur

Av. Prol. Div. del Norte 5662,
Local B, Barrio San Marcos,
Del. Xochimilco,
16090 México, D. F.
Tel.: 1509 0267

Unidad Poniente

Avenida Revolución 1165,
entre Barranca del Muerto y Juan Tinoco,
Col. Merced Gómez,
Del. Benito Juárez,
03930 México, D. F.
Tel.: 5651 2587

Biblioteca

Doctor Rfo de la Loza 300,
primer piso, Col. Doctores,
06720 México, D. F.
Tel.: 5229 5600, ext.: 1608

www.cd hdf.org.mx

COMISION

