

Criterios de calidad en la investigación social: la producción de datos sociales

MANUEL FERNÁNDEZ ESQUINAS

Instituto de Estudios Sociales de Andalucía
Consejo Superior de Investigaciones Científicas
mfernandez@iesaa.csic.es

1. INTRODUCCIÓN

En cualquier tipo de investigación científica un requerimiento fundamental es la calidad de los datos disponibles. Huelga decir que los datos son el fundamento de la evidencia científica, de la contrastación de hipótesis, de la generalización de los resultados de la investigación y de la relevancia de las posibles teorías y marcos analíticos¹. Si esto es así en cualquier parcela de la ciencia, incluida la sociología, suele ser frecuente que en el ámbito de las ciencias sociales existan problemas derivados de la validez y fiabilidad de los datos que se producen y utilizan. Los datos para la investigación sociológica son habitualmente pobres en lo referido a sus niveles de medición, son difíciles de adaptar a marcos poblacionales extensos y presentan la limitación de que los diseños escasamente pueden seguir los esquemas utilizados en las ciencias naturales, fundamentalmente los de diseños experimentales. Aunque esta es una limitación que proviene de la «naturalidad» de los objetos de estudio de la sociología que no es motivo para negar su estatus como ciencia (Beltrán, 1991), existe otro problema que trasciende los requerimientos de validez y fiabilidad que se derivan de las particularidades de los objetos y problemas de estudio, y que consiste en la calidad técnica de los datos que la sociología produce y emplea en sus investigaciones empíricas. A saber, las condiciones sociales y económicas en las que se realiza gran parte de la investigación sociológica suelen resultar en un sistema de producción de datos sujeto a procedimientos de trabajo que dan lugar a datos sociales poco rigurosos,

¹ Sobre el papel de los datos en la investigación científica en general ver, por ejemplo, ZIMAN (1986). Sobre el proceso de investigación científica en sociología puede verse WALLACE (1977) y VERBA, et. al. (2001).

o que meramente están limitados por contener «errores» tanto de medición como de recogida y proceso. A partir de los avances técnicos y tecnológicos que se han producido en la investigación empírica es posible utilizar herramientas que permiten aplicar estándares de calidad, pero también es necesario disponer de herramientas organizativas que tengan en cuenta los procesos reales de trabajo, de forma que sea posible acotar las limitaciones de la investigación en los problemas mismos que estudia la sociología y no en su proceso de trabajo.

La tesis general que se defiende en este escrito sostiene que es necesario aunar los componentes científicos y técnicos en el proceso de investigación con las condiciones sociales objetivas en las que se desarrolla. La decisión sobre la calidad de una investigación no puede basarse únicamente en la formulación y el diseño de un problema de investigación. Esto es, en la formulación de un problema de acuerdo con un marco teórico relevante, en el diseño ajustado al objeto de estudio y en la contrastación de las hipótesis contenidas en el marco teórico. Además de ello, es necesario tener en cuenta el contexto social en el que se produce la investigación, que reúne componentes de tipo material, tecnológico, humano y organizativo. A partir de la consideración de estos elementos, y de las divergencias que ocurren en algunas de las fases de la investigación sociológica, es posible establecer los medios para garantizar unos niveles de calidad para los datos sociales que se generan y se utilizan en la investigación empírica.

El contenido del trabajo se centra en la producción de datos estadísticos en el ámbito de la sociología, principalmente en aquellos que proceden del empleo de la técnica de la encuesta, al ser ésta una de las más habituales en la sociología empírica y también una técnica que presenta numerosas fuentes de error en su proceso de producción debido a la complejidad técnica y organizativa que revisita. De esta forma, las técnicas y los datos de tipo cualitativo se nombran sólo como complemento a la producción de datos cuantitativos, aunque debido principalmente a razones de espacio. Se considera que la investigación cualitativa dispone de un proceso de producción tan complejo o más complejo incluso que la de carácter cuantitativo, si bien de otra naturaleza, proceso que también puede definirse por criterios de calidad de acuerdo con las posibilidades técnicas y sociales para llevarlo a cabo².

2. EL CONTEXTO TÉCNICO Y SOCIAL DE LA PRODUCCIÓN DE DATOS

Antes de comenzar con las particularidades del contexto conviene especificar qué se entiende por criterios de calidad. Al ser éste un trabajo de metodología general que no se ocupa de investigaciones concretas, la definición tiene que ser necesariamente amplia. Los criterios de calidad consisten en la obtención de resultados empíricos cuantitativos adecuados para los objetivos de una investigación, de modo que sean relevantes en términos de los conocimientos socioló-

² Sobre los requerimientos de la investigación cualitativa en sociología ver, por ejemplo, ALONSO (2001) y CALLEJO (2000).

gicos aportados al problema que se define, que dispongan de la mayor fiabilidad y validez posible y estén exentos de errores técnicos y metodológicos. A partir de la anterior definición es posible especificar varios de sus componentes. En primer lugar, cuando se comienza una investigación es necesario realizar una definición adecuada de objetivos en el contexto del conocimiento sociológico que se dispone y del nuevo conocimiento que se persigue. En segundo lugar, es necesario establecer un proceso para llevar a cabo el objetivo de la investigación, esto es, hay que decidir un diseño de investigación que especifique las unidades de análisis, las técnicas, las variables a medir y su adecuación a los posteriores análisis. Hasta ahora, los requerimientos especificados corresponden al ámbito de la metodología de la investigación social, cuestión que en todo caso remite al conocimiento teórico y técnico de la sociología y a su aplicación con criterios de «imaginación sociológica». Pero aún existe un tercer elemento sin el cuál la definición de los objetivos y el diseño difícilmente pueden desembocar en resultados relevantes, y dicho elemento consiste el proceso de obtención de datos de forma exenta de errores.

2.1. El proceso de producción de datos de datos sociales

La obtención de datos sociales responde a lo que se suele denominar un proceso de producción, esto es, un sistema ordenado para la generación de un bien determinado que combina la utilización de recursos materiales, humanos y organizativos en un periodo de tiempo acotado³. Las especificidades de un proceso de producción consisten fundamentalmente en que las características del bien que se produce están «predeterminadas», respondiendo a unos atributos y una cantidad de unidades que se definen de forma externa a la producción, para lo cuál el proceso está sometido a un sistema de control, tanto interno como externo, que garantice el resultado de la combinación de recursos en la obtención del bien establecido. Bien es verdad que pensar en la investigación sociológica en términos de proceso de producción puede resultar excesivo en algunos casos, al concebirse la investigación como una tarea en la que la imaginación y la creatividad tienen un papel preponderante. Sin embargo, a pesar de que dichos elementos sean consustanciales a la tarea científica, la investigación en general está sujeta en cierto modo a un proceso de producción: en el nivel de la concepción metodológica y de la contrastación de resultados el proceso es el del método científico, mientras que en el nivel de la generación de datos el proceso corresponde a lo establecido fundamentalmente en el diseño de la investigación. Quiere esto decir que la producción de datos reúne las características principales de este tipo de proceso, al estar sus objetivos y procedimientos reglados, y al no considerarse válidos criterios que sean externos a los contemplados en el diseño en particular y a la metodología científica en general.

³ Las características sociales y económicas de los procesos de producción en las sociedades industrializadas pueden verse en algunos trabajos clásicos desde varias perspectivas teóricas, por ejemplo, BRAVERMAN (1974) y LITTER (1982).

Además, la producción de datos sociales reúne casi todas las características del proceso mencionado, lo cuál convierte a esta tarea en algo opuesto a la realización individual o personal de la investigación social, de una forma parecida a lo que ocurre cuando se contraponen la producción industrial con la producción artesana. A saber, intervienen en su realización una importante cantidad de recursos, tanto económicos como materiales, y es necesaria la intervención de un equipo humano y tecnológico relativamente amplio, siendo extremadamente difícil que una investigación social que se base en datos representativos pueda llevarse a cabo por un investigador individualmente o incluso por un grupo de trabajo reducido. En suma, la investigación social cuantitativa es una tarea colectiva, al igual que la mayor parte de la investigación realizada en el ámbito de las ciencias naturales, y es colectiva no sólo por el operativo social interno que se requiere, sino también porque debido a sus altos costes es necesario contar con un apoyo y una justificación social adecuada (Gibbons, et. al., 1994)⁴.

El proceso de producción de datos sociales se define por los recursos materiales, incluidos los tecnológicos, los recursos humanos y el aparato organizativo. Estas son las herramientas que intervienen en la mayoría de las investigaciones sociológicas, las cuales definen las condiciones en que se generan las estadísticas. En primer lugar, los recursos, además de los económicos, suelen comprender los equipamientos necesarios para el diseño, la recogida y el proceso de datos, que cada vez más utilizan tecnologías de la información al ser éstas herramientas que reducen la cantidad de trabajo necesaria y que en algunas ocasiones realizan operaciones casi imposibles de ejecutar sin ellas. Por ejemplo, desde hace relativamente poco tiempo se utiliza software especializado para el diseño y, sobre todo, para el estudio de la fiabilidad de las muestras, tanto las probabilísticas como las que, por necesidades del diseño frecuentes en las encuestas sociales, se realizan de forma semi-probabilística, y también se utilizan profusamente sistemas para la grabación y el proceso de datos estadísticos, lo que reduce la cantidad de horas de trabajo necesarias y minimiza los posibles errores.

Pero quizá el componente más importante del proceso de producción en cuestión es el humano debido a que los procedimientos de la sociología empírica difícilmente se adaptan a procesos de automatización en los que sea prescindible la intervención directa de mano de obra, y esto es así debido a que los objetos de la sociología son sobre todo personas, lo cuál requiere un importante esfuerzo para su localización y obliga a utilizar la interacción humana para cualquier proceso de captación de una realidad traducible a datos. Y toda vez que normalmente la cesión de datos es voluntaria, a excepción de algunos registros administrativos que recogen información normalizada con carácter obligatorio, una de las claves fundamentales de la generación de datos es una adecuada interacción entre el equipo que los recoge y las personas que los ofrecen, lo cuál trasciende el ámbi-

⁴ En este sentido, es importante señalar que la «colectivización» de la ciencia afecta a notables disciplinas de las ciencias sociales, para las que se requiere un alto componente tecnológico y organizativo y una división del trabajo en términos de una alta especialización. Ver ZIMAN (1983). Además de la investigación sociológica, se puede considerar en esta situación, por ejemplo, a la investigación empírica en economía o la arqueología en yacimientos de grandes dimensiones.

to de las herramientas para introducirse en el componente social de la producción de datos.

2.2. El componente social de la sociología empírica

A cualquier proceso de producción corresponde un proceso social. Y el proceso social de la sociología empírica está directamente relacionado con las condiciones sociales y económicas de la investigación en curso, especialmente las de las personas que participan en su ejecución. Ahora bien, ¿en qué consiste dicho proceso social? Y sobre todo, ¿cuáles son los elementos de dicho proceso que influyen en la calidad de los datos que se obtienen desde la perspectiva que estamos manejando? La estrategia que aquí se emplea parte de delimitar las fases de la investigación sociológica, las operaciones concretas que corresponden a cada fase y las condiciones sociales que actúan como posibles condicionantes de la investigación.

Lo que aquí se entiende por el componente social que afecta al proceso de producción de datos se quiere diferenciar expresamente de los contextos sociales e históricos que afectan a la sociología empírica. Una peculiaridad de la investigación sociológica, especialmente cuando se emplea la técnica de la encuesta, es que necesariamente se realiza en una situación concreta que puede influir de diversas maneras en la observación de los hechos. La encuesta es un dispositivo de indagación en el que el investigador es representado por varias personas —los encuestadores—, utilizando herramientas para guiar y consignar las observaciones —el cuestionario—, en una serie de situaciones en las que se trata de reproducir una actividad de comunicación cotidiana —las entrevistas—. Además de los sesgos provenientes del diseño del cuestionario, en dicho proceso de comunicación la información que proporciona el encuestado y que finalmente recoge el entrevistador puede estar influida por factores que provienen tanto del contexto social en el que se realiza la investigación como de la interacción que se produce en cada entrevista.

De un lado, las respuestas que se emiten pueden estar influenciadas por la imagen que se tiene de la investigación social por encuesta. Intervienen aquí la importancia y la utilidad atribuida a los estudios sociológicos, la confianza que ofrecen los resultados de este método, la credibilidad de las instituciones que realizan los estudios, la percepción de las garantías de anonimato o la legitimidad que se atribuye a la recogida de información de carácter personal y a su posterior uso (García Ferrando y García, 1976; Justel, 1981). De este modo, cualquier estudio sociológico de este tipo puede verse seriamente condicionado si su contenido o su justificación no obtiene legitimidad suficiente, o si en la opinión pública existe rechazo a los métodos de encuestación. De otro lado, son conocidos los efectos que producen los encuestadores, debido a que las personas entrevistadas reaccionan ante rasgos visibles, como la edad, el sexo, la manera de vestir o el estatus, o ante cuestiones menos manifiestas, como pueden ser los estilos de hacer las entrevistas, la simpatía o el tipo de lenguaje utilizado. También por parte de los encuestados pueden ocurrir variaciones que proceden de rasgos de la

personalidad, como la aquiescencia, o la tendencia a la deseabilidad social en las respuestas, con la que se tienden a citar en mayor medida los asuntos que son valorados positivamente en un contexto social (Reinecke y Schmidt, 1993). Pero además, la encuesta es un dispositivo en el que se produce una interacción entre el encuestador y el encuestado, donde ambos tienen expectativas y percepciones que pueden actuar como determinantes del contenido de la comunicación que se transmite, con lo cual siempre existe la posibilidad de que surja un efecto de tipo contextual en el que ambas partes reaccionan en función del desarrollo de la entrevista, de las actitudes que se manifiestan o del ambiente donde se producen. Incluso, hay autores que sostienen que no existen muestras de personas con las que se consigan respuestas auténticas, sino que en realidad existen muestras de situaciones en las que se establecen patrones de pregunta y respuesta en función de cada situación⁵.

El objetivo de este trabajo no es describir este tipo de problemas, que son consustanciales al trabajo empírico en la sociología, y que en todo caso hay que tener en cuenta en el diseño y la realización de una investigación y en su posterior interpretación. El componente social al que aquí se hace referencia consiste más bien en un conjunto de factores inherentes a la forma en que se realiza el trabajo de investigación sociológica, que de alguna forma se pueden calificar como «internos» de este proceso debido a que corresponden a decisiones organizativas, a las características de los trabajadores que intervienen en el proceso y a los recursos disponibles. Nos referimos, por tanto, a las condiciones de trabajo en que las que se realiza la producción de datos sociales. Sin embargo, no hay que descuidar el hecho de que dichas condiciones pueden influir decisivamente cuando se tratan de controlar las interferencias antes especificadas. La disposición de unas condiciones de trabajo adecuadas, especialmente en lo referido a la formación de los trabajadores, puede contribuir a que la presencia de dichos sesgos se haga más explícita y, por tanto, a que se reduzcan las influencias provenientes de las características del entrevistador y de la forma en que se produce la interacción con los entrevistados en el proceso de recogida de datos.

A continuación nos centraremos en los factores intervinientes en el proceso de generación de datos a partir de las fases típicas que comprende una investigación, desde la formulación de un problema hasta el análisis de los resultados, pasando por el diseño y la recogida de datos. Para cada una de las fases se establecen las principales operaciones que se realizan y se delimitan las necesidades de conocimiento especializado que se requieren en las distintas operaciones. En dicho contexto, se hace referencia al componente social que proviene de las condiciones de trabajo propias de las tareas que predominan en cada una de las fases, teniendo en cuenta que dichas condiciones sociales pueden actuar como factores

⁵ Un trabajo empírico en nuestro país en el que se contrasta la hipótesis de la reactividad en el contexto de interacción que se produce en una encuesta puede verse en ATIENZA y NOYA (1999). En concreto, a través de cuestionarios aplicados tanto a los entrevistadores como a los entrevistados tras la realización de una encuesta, los autores muestran cómo desde ambas partes se tienen en cuenta las distintas características de la interacción, entre otras, el aspecto, los gestos, las actitudes o el desarrollo de las entrevistas.

críticos que afectan al proceso de producción de datos y a la calidad con que éstos se generan.

Si tenemos en cuenta las fases típicas de una investigación, se pueden delimitar cuatro tipos de condicionantes que, a efectos analíticos, se interpretan como los requisitos que plantean las operaciones que habitualmente se realizan en cada una de las fases (ver figura 1). Éstos son condicionantes de tipo «teórico», «técnico», «tecnológico» y «organizativo». El primero de ellos se entiende como los conocimientos relativos a los fundamentos de los problemas de la investigación sociológica. Así, los condicionantes teóricos se refieren a la disposición de un conocimiento adecuado de la disciplina en cuando a su modo de proceder y en cuanto a los conocimientos acumulados en una parcela de la realidad, la que concierne o da sentido a los problemas de investigación que se formulan cuando se establece un objetivo concreto. Los condicionantes teóricos corresponden a las fases de establecimiento de los objetos de estudio, junto a los objetivos, a la definición literaria y al consiguiente establecimiento de dimensiones que sean relevantes para el problema y los objetivos establecidos⁶. Los condicionantes técnicos se refieren a los conocimientos específicos para llevar a cabo unas fases de la investigación, las que corresponden al proceso de operacionalización, a la construcción de los instrumentos de medida y recogida de datos, al diseño muestral y, en parte, a la realización del trabajo de campo, la preparación y el análisis de los datos. Los condicionantes tecnológicos se entienden como el manejo y la disposición de herramientas especializadas, que en este caso corresponden a tecnologías de la información específicas para el proceso y el análisis de datos estadísticos. Por último, los organizativos consisten en el manejo de recursos materiales y, sobre todo, recursos humanos para la ejecución de algunas fases de la investigación, y a su combinación con el resto de los elementos, en particular en las fases de recogida y proceso de datos.

Es relevante señalar que cada tipo de condicionante, si bien puede estar presente de forma implícita o explícita en cualquier fase de la investigación, opera de forma preponderante en unas fases concretas, aunque es posible que en algunas fases existan varios que operan de forma conjunta. Por ejemplo, el conocimiento y las operaciones intelectuales de carácter teórico se presentan cuando se formula un problema de investigación, cuando se realiza su definición en términos literarios, esto es, cuando se describen las características del problema y las hipótesis relevantes, e igualmente cuando se realiza la definición operativa para traducir en datos las observaciones que se realicen. Así mismo, una operación de carácter teórico es la que se realiza cuando se analizan los datos, se exponen los resultados y se valoran de acuerdo con los presupuestos de partida. Por otra parte, el conocimiento técnico opera en la fase de operacionalización, al ser ésta una cuestión de aplicación de un método y, a la vez, al basarse de manera importante en el conocimiento tácito que proviene de la experiencia, así como de la familiaridad con técnicas de otras disciplinas adyacentes. El conocimiento técnico también interviene en las subsiguientes fases del proceso, aunque con distintas

⁶ Para la secuencia operativa de la investigación se toman como referencia los aportes de LAZARSELD (1954).

características. La construcción de instrumentos, el diseño muestral y el análisis de datos corresponden a técnicas especializadas sujetas a procesos en gran parte estandarizados, mientras que el diseño y la ejecución del trabajo de campo suelen estar sujetos a técnicas de organización y gestión de recursos humanos. Por otra parte, el componente organizativo es fundamental en la recogida y preparación de los datos, debido a que en estas fases es donde se requiere el empleo de mayores recursos. Finalmente, los condicionantes de tipo tecnológico operan en las fases en las que es susceptible el empleo de herramientas especializadas. Nótese que es en algunos puntos del proceso de investigación donde intervienen varios condicionantes, por ejemplo, la preparación y el proceso de datos, la recogida de datos o la definición operativa, lo que emplaza a estas fases como puntos críticos en lo referido al establecimiento de criterios de calidad.

Por otra parte, el componente social tiene distinta presencia en cada uno de los condicionantes, funcionando como «agravante» o «atenuante» de las condiciones de producción de datos al estar presente con ciertas características en algunas de las fases de la investigación. Y estas fases son sobre todo aquellas para las que se requieren herramientas organizativas, al ser las que reúnen una mayor cantidad de recursos y trabajadores que incorporan distintas características de cualificación, de actitud y motivaciones en el trabajo o de recursos y prestigio obtenidos o esperados en la ejecución de su cometido laboral. Las fases citadas corresponden a la recogida y la preparación de los datos, aunque los condicionantes sociales también están presente en el diseño de instrumentos, por ejemplo, en la prueba y validación de los cuestionarios, pero lo que suelen realizarse *pretest* en condiciones reales con grupos amplios de trabajadores de campo. Es por ello que en estas fases es necesario especificar los criterios de calidad y establecer los procedimientos específicos para la obtención de niveles satisfactorios, aunque, paradójicamente, es en ellas donde las condiciones sociales imperantes provocan que ocurran con mayor frecuencia divergencias organizativas e incluso conflictos sociales que dan lugar a numerosas fuentes de error en la ejecución del trabajo.

TABLA I
Características del proceso de investigación

Fases	Operaciones	Condicionantes
Formulación del Problema	Delimitación del objeto de estudio Definición de objetivos	Teórico
Definición literaria	Especificación de las características Descomposición analítica	Teórico
Definición operativa	Establecimiento de: — Dimensiones — Variables — Indicadores/índices — Unidades de análisis	Técnico/teórico

TABLA 1
(*continuación*)

Fases	Operaciones	Condicionantes
Construcción de Instrumento	Formulación de preguntas/variables Pretest-validación Diseño de indicadores	Técnico
Diseño de recogida de datos	Concreción unidades de análisis Diseño de Muestra	Técnico/tecnológico
Recogida de datos	Ejecución de trabajo de campo Control de veracidad de los datos	Organizativo/técnico
Preparación / proceso de datos	Codificación Grabación Depuración	Organizativo/técnico/ tecnológico
Análisis	Construcción de Indicadores Tabulación Cálculos	Técnico/tecnológico
Desarrollo y conclusiones	Contrastación empírica Valoración de resultados	Teórico

3. LAS DIVERGENCIAS ENTRE LOS REQUERIMIENTOS DEL TRABAJO Y EL CONTEXTO SOCIAL EN LA SOCIOLOGÍA EMPÍRICA

Hasta ahora se ha descrito esquemáticamente la forma en que se suele realizar la investigación social cuantitativa, así como los elementos que intervienen en su realización. Para centrar la exposición en el problema que se trata en este trabajo es necesario tener en cuenta las divergencias que se producen en el proceso de producción entre algunos de sus componentes, y que dan lugar a la existencia de anomalías que resultan en datos de escasa calidad o de calidad mejorable. No se quiere decir que todos los datos de la investigación sociológica presenten esta problemática, sino que, dadas unas determinadas condiciones, existe una probabilidad razonable de que los datos resulten en defectos de calidad y, de hecho, si tenemos en cuenta las condiciones en que se producen gran parte de las investigaciones empíricas, al menos en nuestro país, es razonable pensar que muchos de ellos surjan con errores derivados de falta de

recursos, falta de cualificación de algunos de los participantes o escasa capacidad de organización⁷.

Respecto a los tipos de errores a los que nos referimos, consisten en errores de validez, los que afectan a que lo que se mide de cuenta de lo que se quiere medir, y de fiabilidad, los que dan cuenta de que lo que se mide representa fielmente lo que se quiere medir⁸. Aquí se tratan principalmente los errores del segundo tipo. La recogida y proceso de datos puede responder a un diseño adecuado, diseño que recoge criterios de validez en los instrumentos y de fiabilidad en las dimensiones y extensión de los datos, pero que pueden verse afectados por el proceso de producción de dichos datos debido a la presencia de factores que son ajenos a los contemplados por el investigador o que están alejados de los que el investigador planifica y controla. A saber, merman la representatividad de los datos al ser las unidades muestrales no elegidas de forma adecuada, por ejemplo, no siguiendo criterios de aleatoriedad, y merman su contenido en las unidades de análisis al realizarse las medidas de forma no exhaustiva, lo cual resulta en problemas de «no respuesta», «no consignación» o «consignación incorrecta», como se especifica de manera resumida en la figura 2.

TABLA 2
Fuentes de error en la producción de datos sociales

LOCALIZACIÓN DE ERRORES		
TIPOS DE ERRORES	UNIDADES DE ANÁLISIS	VARIABLES
AUSENCIA DE DATOS	No elección de unidades Ausencia de contactación Ausencia de reemplazos (Muestra incompleta)	Preguntas no realizadas Respuestas no emitidas Respuestas incoherentes Respuestas no consignadas Respuestas no procesadas (Falta de información en las variables)
SESGOS EN LOS DATOS	Elección incorrecta Reemplazamiento incorrecto (Sesgo en la composición de la muestra)	Preguntas no formuladas adecuadamente Preguntas no entendidas adecuadamente por el entrevistado Respuestas no entendidas adecuadamente por el entrevistador Respuestas consignadas erróneamente Respuestas procesadas erróneamente (Sesgo en la información contenida en las variables)

⁷ Por ejemplo, si tenemos en cuenta los proyectos de investigación en sociología financiados por el Plan Nacional de I+D+I observamos que la gran mayoría dispone de financiación inferior a los 1500 euros, incluso aquellos que plantean en sus objetivos y metodología la obtención de un conjunto amplio de datos. Ver CICYT (1997, 2001).

⁸ Sobre criterios de validez y fiabilidad en la sociología puede verse GARCÍA FERRANDO, et. al (1986).

La idea que se sostiene en este apartado es que existe una divergencia entre el componente técnico y el componente social de la investigación empírica que se deriva de las condiciones de trabajo en las que se desarrolla el proceso de investigación. Y dicha divergencia se manifiesta en la ruptura que existe entre dos partes del proceso que están diferenciadas principalmente en su vertiente social, esto es, en las condiciones de trabajo que presentan, tanto en lo referido a los recursos asociados a los puestos de trabajo como en las condiciones materiales y organizativas en que se realizan las tareas, así como en la valoración profesional de los participantes y de sus cometidos. La ruptura consiste en que, de un lado, en un conjunto de fases de la investigación existen unas condiciones de trabajo⁹ definidas por elementos caracterizados por ventajas relativas, mientras que otra serie de fases se distinguen por características que corresponden prácticamente a lo contrario, en términos de desventajas o impedimentos en lo relacionado con la realización de las tareas concretas dentro del proceso productivo.

Los elementos que se distinguen (ver figura 3) se agrupan en:

- a) Aquellos que corresponden al estadio de conocimientos en la fase referida al proceso de investigación, a saber, la amplitud de los problemas y objetivos de la investigación, la disposición de conocimientos acumulados, tanto en problemas concretos como en la disciplina en general, así como la complejidad en términos teóricos, es decir, la diversidad de planos de observación y análisis.
- b) Los que corresponden a las posibilidades de discusión en las que se barajan alternativas metodológicas y técnicas para la ejecución del trabajo en función de los objetivos propuestos.
- c) Los requisitos de carácter organizativo que se plantean en función de la complejidad de los recursos empleados, así como las necesidades tecnológicas.
- d) La cualificación disponible en los trabajadores que intervienen en las distintas fases.
- e) Los recursos relativos disponibles teniendo en cuenta los recursos utilizados en cada fase en relación con las necesidades materiales, humanas y temporales para la ejecución del trabajo.

En el esquema se especifica la dirección en la que se decantan las condiciones en cada grupo de fases de la investigación, a partir de lo cuál es posible delimitar las divergencias entre los requerimientos de la investigación empírica y la forma real en que se producen. La divergencia consiste, pues, en una ruptura entre objetivos y medios, debido a que a objetivos complejos y extensos suelen corresponder medios escasos, caracterizados además por la ausencia de conocimientos reglados y acumulados en problemas concretos de forma que sea posible establecer métodos válidos para una generalidad de problemas de investigación.

⁹ Las condiciones de trabajo se entienden como el conjunto de recursos económicos, materiales, de poder y de status que corresponden a los puestos de trabajo, junto a las limitaciones del proceso productivo en el que interviene cada trabajador. Ver CASTILLO y PRIETO (1983).

TABLA 3
Condiciones de trabajo del proceso de investigación sociológica

DIMENSIONES EN LAS CONDICIONES DE TRABAJO											
FASES DE LA INVESTIGACIÓN	Definición de Objetivos	Discusión de estado de la cuestión	Acumulación de conocimientos	Discusión de propuestas	Complejidad teórica	Complejidad técnica/organizativa	Equipo de trabajo	Tiempo disponible	Cualificación de los trabajadores	Recursos relativos	Necesidades tecnológicas
Formulación del Problema	Amplios	Extensa	Alta	Extensa	Alta	Baja	Reducido	Extenso	Alta	Altos	Bajas
Definición literaria	Homogéneo										
Definición operativa	Homogéneo										
Construcción de Instrumento	Homogéneo										
Diseño de recogida de datos	Homogéneo										
DIVISIÓN ORGANIZATIVA											
Recogida de datos	Concretos	Reducida	Baja	Reducida	Baja	Alta	Amplio	Reducido	Baja	Bajos	Altas
Preparación de datos	Heterogéneo										
Proceso de datos	Heterogéneo										

Como se observa, en el extremo superior se encuentran las condiciones de trabajo relativamente ventajosas, esto es, definición de objetivos de una forma abierta y amplia, posibilidad de discusión teórica y metodológica, acumulación de experiencias previas, cualificación de los trabajadores que intervienen, equipos de trabajo reducidos y homogéneos, recursos relativamente altos en función de las necesidades de las fases del proceso y de la envergadura del equipo de trabajo y necesidades organizativas y tecnológicas relativamente bajas. En el extremo inferior se encuentran los procesos que conllevan condiciones de trabajo caracterizadas por objetivos muy concretos, escasa oportunidad de discusión de alternativas, acumulación baja de experiencia o, en todo caso, acumulación escasamente formalizada, baja cualificación de trabajadores, escasos recursos relativos en comparación con los objetivos planteados y los elementos materiales y humanos que intervienen, junto a necesidades altas de organización y, en algunas fases, de herramientas tecnológicamente avanzadas.

Quizá esta discusión puede aclararse con la exposición de un ejemplo correspondiente a una «encuesta tipo» en la sociología, dirigida a población general con un marco poblacional y muestral amplio. A saber, cuando se lleva a cabo una investigación que requiere la realización de una encuesta, las primeras tareas son realizadas por los integrantes del equipo de investigación, que son los que deciden los contenidos, los métodos y los ritmos, y posteriormente se encargan de los análisis. Dicho equipo suele estar compuesto por un número reducido de profesionales, o incluso por un solo profesional, que disponen de credenciales específicas que les reconocen como integrantes de la profesión y que además realizan el trabajo de forma regular y con remuneración reglada por su tarea¹⁰. De otro lado, una serie importante de tareas corresponden a los trabajadores de campo y a los que realizan el proceso de datos o tareas específicas, como diseños y análisis concretos. Son éstos los grupos de encuestadores, coordinadores de campo, supervisores, grabadores y trabajadores especializados en tecnologías de proceso de datos, junto a los trabajadores que realizan tareas administrativas. Si bien es cierto que en las sociedades desarrolladas suelen existir organizaciones públicas o empresariales especializadas en estas tareas, también es cierto que frecuentemente esta parte de la investigación sociológica se realiza *ad hoc*, formándose operativos a propósito para cada estudio que desaparecen o cambian a otro ámbito una vez que el trabajo ha finalizado. En suma, lo que se encuentra en este tipo de organización del trabajo es una división laboral entre investigadores y técnicos, esto es, de un lado los profesionales, y de otro lado los trabajadores semicualificados o no cualificados que constituyen la mano de obra subordinada de la investigación social.

Cabe decir que este hecho está presente en prácticamente cualquier parcela de la estructura productiva. En todo proceso de producción existe una división del trabajo que resulta en una estratificación social en términos de recursos manejados y obtenidos, capacidad de control sobre el proceso de trabajo y estatus atri-

¹⁰ No se quiere decir que las remuneraciones de los investigadores en sociología sean de un nivel determinado, sino que realizan su tarea de una forma profesional con una remuneración relativamente estable.

buido a los puestos y tareas. Ello también existe en la investigación científica en general, y en la investigación sociológica en particular, toda vez que en la organización del trabajo sociológico cada vez más se realiza de una forma colectivizada en los términos antes expuestos. Así, en la sociología también existe una segmentación del mercado de trabajo (Doeringer y Piore, 1974) en la que, a grandes rasgos, el mercado interno está formado por el núcleo estable de la profesión, mientras que el mercado externo se constituye con los trabajadores auxiliares, que están más sujetos en sus condiciones laborales y económicas a las fluctuaciones del mercado, del mismo modo que las empresas externalizan tareas que pasan a formar parte de los mercados de trabajo externos de amplios sectores productivos.

Volviendo al asunto de la calidad de los datos, la divergencia social y de condiciones de trabajo es lo que provoca la ruptura entre las fases del proceso, de manera en unas fases existen mayores presiones respecto al coste y al ritmo de trabajo, menor planificación, menores recursos, menor cualificación, tanto exigida como disponible en el mercado, todo ello en un proceso productivo que exige unas necesidades organizativas altas y que es extremadamente complejo debido a la enorme variedad de factores que pueden introducir fuentes de error y sesgos sistemáticos. En el otro extremo se sitúan, todo ello en términos relativos, menores presiones sobre costes y ritmo de trabajo, sobre todo porque los costes no suelen ser importantes, mayor tiempo para la planificación, más capacidad de maniobra y contrastación y prueba de tareas, mayores recursos y mayor cualificación, ya que las tareas suelen realizarse casi en exclusiva por los integrantes de la profesión o por miembros de profesiones adyacentes. Además, entre ambos extremos el contacto suele estar reducido, en primer lugar por una limitación organizativa —el trabajo de campo suele estar disperso en el territorio y el tiempo disponible no suele dejar al investigador oportunidad para implicarse en todas las fases del proceso— pero también por razones de prestigio o de carga de trabajo que no está reconocida profesionalmente ni a través del sistema de recompensas de la ciencia basado en las publicaciones. Este hecho provoca, por tanto, que los diseños de investigación se hagan en una mesa de despacho, bien en solitario, bien en discusión con un grupo reducido, mientras que las responsabilidades organizativas y de control de recogida de datos corresponden a los coordinadores de trabajo de campo, que suelen trabajar bajo altas presiones organizativas, presupuestarias y temporales y que, además, son profesionales «hechos a sí mismos» debido a que sus conocimientos difícilmente se pueden adquirir por medio de la educación reglada o a través de la formación privada.

Esta ruptura en la organización del trabajo científico ¹¹ provoca también una ruptura en los niveles de calidad, concentrándose los riesgos en la parte fun-

¹¹ Cabe decir que la ruptura existe principalmente en el trabajo científico circunscrito a la sociología y a las ciencias sociales en general, lo que no suele tener equivalente en la investigación de las ciencias naturales, sobre todo en los sistemas de I+D más desarrollados. Ello se debe en cierta medida al relativo subdesarrollo de las ciencias sociales, al ser éstas mucho más jóvenes, pero también a la escasa legitimidad social en comparación con otras ciencias, lo que se traduce en una baja financiación, y a las graves consecuencias que en dichas ciencias puede tener la producción de

damental del trabajo de investigación empírica, esto es, el trabajo relacionado directamente con la producción de datos, con su obtención, su control y su preparación, lo que abunda en la existencia de múltiples investigaciones sociológicas de carácter teórico o sin un importante sustrato empírico, y también en la proliferación de investigaciones empíricas, si no desprovistas de todo componente teórico, sí carentes del rigor técnico suficiente para considerarlas integrantes centrales de los conocimientos sociológicos que se disponen sobre una sociedad. Como se puede comprobar, la cuestión citada tiene cierto paralelismo con la conocida controversia señalada por Wright Mills entre el «empirismo abstracto» de Paul Lazarsfeld y la «gran teoría» de Talcott Parsons (Mills, 1968), al primar uno u otro componente organizativo una de las vertientes de la investigación sociológica. Teniendo en cuenta contextos sociales y científicos concretos, se puede sostener que el peso de la investigación sociológica se ha decantado hacia uno de los dos extremos según el caso que se contemple¹². No es extraño, pues, que la investigación empírica de carácter cuantitativo se haya desarrollado principalmente en los EE UU, mientras que en países científicamente retrasados o sujetos a condiciones económicas y políticas poco favorables, como ha sido el caso de España, la profesión sociológica se haya decantado por la investigación de carácter teórico teniendo en cuenta las dificultades para financiación de la investigación empírica de calidad, pero también porque todo parece indicar que, a pesar el auge de los estudios cuantitativos, los requerimientos políticos o económicos de su alta financiación, el esfuerzo necesario para llevarla a cabo en dichas condiciones¹³ y el prestigio profesional que se obtiene a cambio no están del lado de los que la practican directa y principalmente.

Hasta ahora el panorama expuesto puede dar la impresión de consistir en un círculo vicioso de difícil resolución: ausencia de recursos conlleva ausencia de medios económicos y organizativos, ausencia de personal cualificado y de control de calidad, lo que conlleva falta de relevancia y, por tanto, falta de legitimidad social, lo que se traduce de nuevo en ausencia de recursos. Con esta secuencia, parece ser que la solución pasa necesariamente por aumentar los recursos. Esto es cierto en alguna medida. Del mismo modo que disponer de los recursos para sufragar el coste del proceso productivo de cualquier bien es la primera con-

datos de escasa calidad (la medicina, la física, las ingenierías, etc.), que no suele revestir el mismo carácter en la sociología.

¹² Permítaseme añadir un comentario valorativo al hilo de la controversia citada, que tiene que ver con el balance de la importancia adjudicada a una de las fases de la investigación: «Si se tiene un buen dato, aunque no se tenga una buena idea, el dato será irrelevante pero seguirá siendo válido hasta que alguien que tenga una idea lo utilice. Si se tiene una buena idea y se tiene un mal dato, la idea será importante pero quedará en el terreno de la elucubración, hasta que alguien obtenga un buen dato que la corrobore».

¹³ SÁNCHEZ CARRIÓN señala que el esfuerzo necesario para aprender en profundidad todos los aspectos de la técnica de la encuesta es tal que difícilmente deja tiempo y fuerzas para pensar sobre los problemas metodológicos y teóricos que conlleva su aplicación (SÁNCHEZ CARRIÓN, 1996). Dicho esfuerzo es consecuencia del relativo subdesarrollo de esta parte de la sociología en España, al contarse con escasos profesionales especializados y también con escasas organizaciones que realicen este trabajo de forma continuada.

dición para producirlo, para poner en marcha un proceso de producción de datos sociales es necesaria una masa crítica de recursos. Así, ante la pregunta de si es conveniente realizar una encuesta, la respuesta sería que, si no se tienen el dinero suficiente, lo más conveniente es que no se haga debido a que los resultados no garantizan que el esfuerzo de poner en marcha el proceso sea mínimamente fructífero. Pero ante la misma pregunta, también habría que responder que si no se tienen los suficientes recursos humanos en términos de formación y motivación, quizá lo mejor es que tampoco se haga, a no ser que se persiga realizar un entrenamiento para adquirir destrezas. No obstante, también es posible introducir elementos previos que, a partir de una situación de recursos escasos y escasa cualificación, permitan garantizar un resultado válido en términos de calidad, tal como se aborda a continuación.

4. ESTRATEGIAS DE CALIDAD: CRITERIOS PARA LA RESOLUCIÓN DE DIVERGENCIAS EN EL PROCESO DE PRODUCCIÓN DE DATOS

La muy abundante literatura sobre metodología de investigación en ciencias sociales está repleta de métodos, técnicas o procedimientos para obtener datos estadísticos de carácter sociológico que sean válidos y fiables¹⁴. Aunque muchos problemas de este tipo están resueltos en el ámbito del conocimiento sociológico, no lo suelen estar en los ámbitos concretos de las investigaciones empíricas, frecuentemente porque la envergadura de numerosas investigaciones limita la aplicación de los procedimientos existentes. Por ejemplo, es sabido que un *pre-test* es un componente esencial de cualquier encuesta. Sin embargo, teniendo en cuenta la metodología recomendada, es improbable que se cuente con los recursos y con el personal especializado para realizarlo, al igual como ocurre con otros procesos de la investigación cuantitativa. Por este motivo, la estrategia de este apartado no es exponer la gama de procedimientos de la sociología para garantizar la calidad de los datos, ni tampoco consiste en establecer operaciones para comprobar o para medir el nivel de calidad. Más bien se trata de exponer de forma resumida algunas de las operaciones que se pueden aplicar fácilmente al proceso de producción para reducir los errores, si se quiere, podemos decir en términos de probabilidad.

La exposición se divide en dos puntos, de acuerdo con algunos de los componentes que intervienen en las fases de la investigación, haciendo referencia en cada punto a los criterios fundamentales que se pueden seguir para evitar los errores más comunes en la sociología empírica, a saber, errores de recogida y proceso de datos —independientemente de los errores de diseño y de análisis de los datos—, dando primacía a los componentes que anteriormente se han identificado como los puntos claves de la ruptura del proceso de producción.

¹⁴ Sobre técnicas de investigación ver SELLIZ, et.al. (1980) y KÖNIG (1973).

4.1. El componente técnico

El diseño y la prueba de los instrumentos de medida

Un cuestionario no es sólo un conjunto de preguntas y respuestas predeterminadas y ordenadas en cierto grado. Es además un dispositivo para recoger razonamientos individuales que, agregadamente, se entienden como representantes de marcos poblacionales más amplios y, en su caso, como vinculados a un componente social de dichos marcos poblacionales. Así pues, junto a las cuestiones de hecho u objetivas que se recogen en los cuestionarios (educación, trabajo, características de la vivienda, etc.), cuando es necesario recoger cuestiones de tipo subjetivo, como son las opiniones y las valoraciones, es necesario seguir los principios que en la psicología básica rigen en cualquier razonamiento individual. Por tanto, un cuestionario tiene que tener una secuencia expositiva lógica, no sólo para el investigador, sino para el encuestado. Aunque este hecho se suele nombrar en la literatura especializada como un principio fundamental del diseño, aquí se quiere resaltar que es especialmente relevante en la fase de análisis debido a que la validez de los datos determina la validez de los análisis. De este modo, cuando se procede a realizar un análisis estadístico como soporte a una investigación social, hay que tener en cuenta que las composiciones de sentido se realizan habitualmente *a posteriori*, es decir, son construcciones hechas por el investigador, y que dicha construcción es mayor cuanto mayor es la utilización de respuestas sin conexión lógica en el desarrollo de una entrevista. De esta forma, la construcción lógica de un cuestionario de forma que se adapte a los posibles razonamientos que una persona puede realizar en torno a un problema es un criterio que evita posibles errores de interpretación de los datos, los que les atribuyen una base racional que difícilmente tienen cuando se introducen, bien en el cuestionario, bien en el análisis, siguiendo únicamente el proceso intelectual del investigador¹⁵.

Por otra parte, este criterio también es útil en las estadísticas centradas en datos objetivos, lo que plantea otro problema adicional. A pesar que en los trabajos de producción de datos suele existir una frontera entre las estadísticas de tipo objetivo y subjetivo, a saber, entre las mediciones de la renta, el trabajo o las condiciones de vida que suelen realizar los institutos oficiales de estadística, y las mediciones de las opiniones y actitudes que suelen realizar tanto organismos oficiales para el estudio de la opinión pública, como empresas y organizaciones de investigación académica, es conveniente que las encuestas de tipo

¹⁵ Por ejemplo, los análisis dirigidos a la reducción de datos que se utilizan habitualmente en la sociología (análisis factorial, de componentes principales o de ajuste óptimo para medidas no intervalales) pueden dar lugar a interpretaciones de este tipo cuando se describen las dimensiones subyacentes a los datos. Si bien estos problemas no suelen existir cuando las dimensiones corresponden a cuestiones objetivas, sí existe el riesgo de atribuir a las opiniones y actitudes que se agrupan en una dimensión un rango más elevado del que tienen en la realidad, a no ser que se interpreten a modo de «factores latentes» que no están interiorizados o no se manifiestan por los propios encuestados. Sobre técnicas estadísticas para la reducción de datos ver GIFÍ (1990) y SPSS (1997).

objetivo incluyan preguntas de opinión o de percepción, debido a que las cuestiones de hecho suelen «filtrarse» con las categorías de entendimiento de los encuestados, que no tienen por qué corresponder en su totalidad con los conceptos plasmados en un cuestionario¹⁶. Por ejemplo, un tema clásico es la autoevaluación en el asunto de la actividad y la ocupación, en lo que las estadísticas oficiales suelen estar sujetas a controversias. En este caso, es más conveniente, por ejemplo, contrastar la clasificación que realiza una persona sobre ella misma como parado u ocupado si se dispone del entendimiento que tiene esa persona del hecho de estar parado u ocupado.

Un segundo aspecto, esta vez relacionado con la prueba de los instrumentos, es la necesidad de contrastar que el cuestionario sea válido para los objetivos de la investigación. También se puede encontrar en la literatura una amplia casuística en la realización de *pretest*, si bien el aspecto que se quiere resaltar aquí es que la prueba de un cuestionario es conveniente realizarla tanto en fase cualitativa como cuantitativa¹⁷. En un primer momento, el objetivo del *pretest* se dirige a indagar el sentido en el que se formulan las preguntas y las respuestas. Mediante el discernimiento de los sentidos otorgados a las proposiciones y conceptos plasmados en un cuestionario es posible establecer en qué grado dicha interpretación corresponde a la inicialmente formulada en el diseño de la investigación y realizar correcciones dirigidas a adaptar el lenguaje de la investigación al lenguaje de sentido común, tanto en los segmentos sociales a los que pertenecen los sujetos de estudio como en los trabajadores de campo. En un segundo momento, el *pretest* responde al problema de la transformación de la realidad en datos estadísticos y, sobre todo, a la adecuación de dicha transformación a las condiciones del trabajo en curso. Esto es, esta fase del *pretest* debe comprobar si el diseño se adapta a las expectativas que los trabajadores de campo tienen y, sobre todo, a sus capacidades para interactuar con los sujetos de estudio, para formular las preguntas del cuestionario y para consignar adecuadamente las respuestas que se emiten. Por ejemplo, puede ocurrir que un diseño impecable resulte fallido porque no se adapte a los hábitos de trabajo o a las capacidades de los encuestadores.

Así pues, un principio fundamental de la producción de datos sociales es que la cualificación de los trabajadores de campo define el nivel de la complejidad de los datos que se producen, así como los recursos disponibles también definen la complejidad y la extensión de los datos. Quiere esto decir que, siguiendo con el ejemplo del cuestionario bien diseñado, éste tampoco será adecuado si no es realista en términos de la correspondencia de su dificultad y duración con las capacidades, la remuneración y el tiempo que de que disponen los integrantes del equipo de trabajo para realizar su tarea.

¹⁶ Sobre construcción de cuestionarios ver DE MAIO (1993), CONVERSE y PRESSER (1986) y BRADBURN, et. al. (1991).

¹⁷ Sobre la metodología de *pretest* puede verse HUNT, et. al. (1982) y BOND (1993) Una revisión comparativa de los distintos métodos es PRÜFER y REXROTH (1996). Sobre la combinación de técnicas cualitativas y cuantitativas en la investigación social ver CONDE (1994) y CALLEJO (1988).

La adaptación del diseño de las muestras

Los criterios de calidad en lo referido al diseño de muestras para encuestas sociales tienen igualmente dos vertientes que aquí se quieren resaltar, al margen de los criterios habituales referidos a los principios del muestreo¹⁸. Una de ellas tiene que ver con la utilización de criterios técnicos adecuados para asegurar la representación de los datos en términos de adecuación óptima de los recursos con los objetivos científicos en función de las posibilidades técnicas disponibles. En este punto, los avances en el software especializado hacen posible tener en cuenta criterios estadísticos en el diseño de las muestras que ayudan a superar las tradicionales limitaciones de los diseños en las encuestas sociales¹⁹. Una limitación importante es la que adjudica niveles de error correspondientes a muestreos probabilísticos a diseños que, por imposibilidades técnicas u organizativas, se realizan de forma no probabilística, como es el caso de los muestreos polietápicos mediante la elección previa de conglomerados. Así, los niveles de error que se manejan en las estimaciones no son los reales debido a que es extremadamente difícil realizar los cálculos que tienen en cuenta los efectos del número, el tamaño y la dispersión interna de los conglomerados. También es cierto que si el software especializado permite calcular los errores reales para la estimación de cada variable, introducen requerimientos de exactitud en el diseño debido a que los resultados que ofrecen las nuevas herramientas pueden resultar inadecuados en términos estadísticos para los presupuestos de representatividad que se emplean en diseños simples²⁰. Una segunda vertiente es la referida a la optimización del trabajo de campo en términos estadísticos, teniendo en cuenta la dispersión de variables clave en segmentos o estratos poblacionales. En el caso de las muestras a población general que se distribuyen en partes amplias del territorio las nuevas herramientas posibilitan elegir puntos muestrales de determinado tamaño y situados en lugares concretos del territorio, de forma que sea posible optimizar los costes que suelen ir asociados al esfuerzo organizativo y presupuestario necesario cuando las unidades de análisis están muy dispersas, garantizando al mismo tiempo unos niveles de error aceptables.

La codificación y el proceso de datos

Una fase a la que no se suele prestar atención en la literatura de la investigación empírica de carácter sociológico es a la preparación de los datos que se reco-

¹⁸ Ver COCHRAN (1996) y KISH (1995).

¹⁹ En los años 90 ha surgido software especializado para el diseño de muestras complejas, para la estimación de errores en muestreos por conglomerados (por ejemplo Wesvar), así como procedimientos de «remuestreo» (Bootstrapp) que, entre otras cosas, intentan solventar el problema del sesgo muestral mediante la replicación artificial de un número elevado de muestras a partir de la original obtenida. Ver EFRON y TIBSHIRANI (1993).

²⁰ Por ejemplo, para obtener un margen de error de $\pm 2\%$ que en un supuesto de aleatoriedad simple corresponde a 2400 unidades, en un muestreo a población general mediante conglomerados la cantidad muestral suele ascender a 3700 unidades para la obtención de un nivel de error equivalente. Ver TRUJILLO (2000).

gen en función de su adaptación a los objetivos y las posibilidades del análisis²¹. La tarea principal que corresponde a la codificación es el tratamiento de la información que se suele llamar «abierta», es decir, la que se recoge de forma no estandarizada, bien a través de la observación de un hecho por parte del encuestador, bien a través de la respuesta que corresponde a una pregunta cuyas posibilidades de respuesta no vienen predeterminadas. En estos casos, utilizados habitualmente en las encuestas sociales, es necesaria la intervención de un equipo de trabajo especializado en la codificación de los datos debido al gran volumen que es necesario procesar. La codificación es el proceso de atribución de signos, normalmente códigos numéricos, correspondientes a una relación predeterminada de posibilidades exhaustivas y mutuamente excluyentes en la que se clasifica la amplia variedad de respuestas o de observaciones particulares, todo ello de acuerdo con un esquema que responde a un sentido específico de clasificación u ordenación de la realidad observada. El proceso de codificación se puede equiparar al proceso de traducción de una lengua, al tener que realizarse una comprensión de un lenguaje específico en el que hay que transformar una información recogida con unos criterios de clasificación distintos, es este caso marcados por la diversidad de códigos lingüísticos utilizados por los encuestadores y encuestados. Así, la tarea de los codificadores consiste en un proceso de traducción de una determinada realidad a signos, proceso en el que pueden producirse notables distorsiones que provienen del aprendizaje del lenguaje especializado y de la atribución diferencial de significado que se realiza tanto en la información original como en la relación de códigos. Los problemas relativos al nivel de calidad relacionados con el proceso de codificación tienen que ver con dicha tarea de traducción. El primero de ellos es que difícilmente la información de que se dispone es original. El segundo y subsiguientes consisten en que en el proceso de transformación de observaciones o expresiones verbales a códigos de carácter estadístico está «filtrado» por las atribuciones de sentido que realiza cada uno de los intervinientes en el proceso.

La secuencia esquemática para el caso de una pregunta abierta en un cuestionario corresponde a lo siguiente: el origen del proceso es la respuesta deseada por el respondente en función de la pregunta formulada. Al margen del posible sesgo que se puede introducir en dicha formulación, existen una serie de posibles desviaciones en cada uno de los pasos que se siguen. Puede existir una desviación de la respuesta deseada respecto a la respuesta emitida. Una segunda desviación va desde la respuesta emitida a la respuesta entendida por el que la consigna y una tercera desde la respuesta entendida a la respuesta escrita. Estas son las desviaciones habituales en la fase de trabajo de campo, que son las que determinan el tipo de información disponible. En la fase de codificación pueden existir igualmente otras desviaciones que responden a la siguiente secuencia: i) desviación entre la respuesta escrita y la respuesta categorizada o entendida en función del esquema de codificación, ii) desviación entre la categorización y la consignación del código,

²¹ No es así en la literatura estadística en general, en la que sí existe abundante metodología para el proceso y la depuración de datos, igualmente útiles para la sociología, por ejemplo, VILLÁN CRIADO (1990).

iii) desviación entre la lista de posibles categorías y la agregación de categorías en función de su adaptación operativa —la recodificación—, y iv) a su vez desviación entre la categoría, agregada o no, y la interpretación de su significado. Todas estas desviaciones, a las que se pueden añadir las del proceso de grabación, pueden multiplicarse por tantas personas como intervienen, primero en la recogida de datos y segundo en el proceso de los mismos, debido a que cada uno de los participantes puede tener distinto entendimiento y distintos criterios en las operaciones que realiza, lo que amplifica las posibles fuentes de error en el proceso de generación de datos. Así pues, si no se controlan adecuadamente todas estas fuentes de variación puede ocurrir que cuando se utilizan cuestionarios abiertos la correspondencia entre la información manejada en los análisis y la información original pueda ser más resultado de la coincidencia que del propio diseño de investigación.

A pesar de que estas fuentes de error puedan ser controladas en los análisis, y de que se suponga que los errores aleatorios o no sistemáticos se puedan asumir en algunos casos como «no distorsionantes», la realidad es que los analistas o investigadores difícilmente disponen de información suficiente sobre desarrollo de la investigación para poder interpretar los datos en este sentido. Además, los errores suelen ser en gran medida sistemáticos, es decir, repetidos y realizados con escasas posibilidades de cambio que no sea inducido. Esto es debido a que en los trabajos de tipo subordinado, en los que los trabajadores no tienen conocimientos de muchos de sus fundamentos, y en los que difícilmente conocen los resultados —como son los trabajos de encuestación y codificación—, las rutinas que se aprenden no tienen posibilidades de ser corregidas a partir de la contrastación que cada trabajador puede realizar de su propia tarea en términos de aciertos o errores.

4.2. El componente organizativo y tecnológico

Si bien los componentes técnicos tienen un papel preponderante en los niveles de calidad de los datos sociales, lo cuáles tienen una resolución principalmente de tipo formativo, como se verá a continuación, los componentes organizativos tienen un papel fundamental. A partir de ellos se conjugan los criterios científicos y técnicos con los recursos disponibles, siendo estas operaciones las que permiten realizar una adecuación a las condiciones sociales en las que se realiza el trabajo. Entre los factores organizacionales se contemplan los que tienen que ver con el proceso de trabajo de campo, con la adecuación del diseño muestral al trabajo de campo y con los recursos humanos tanto en su formación como en su articulación interna. Entre los factores tecnológicos se citan los derivados de la utilización de las tecnologías de la información y las comunicaciones (TIC) en el proceso de los datos y en la realización de entrevistas.

El proceso de recogida de datos y el equipo de trabajo

La recogida de datos es uno de los puntos clave al concentrarse en ella limitaciones económicas, temporales, humanas, técnicas y de organización. El princi-

pal problema que da lugar a la mayor parte de las fuentes de error es el relativo a la variedad de criterios que se pueden emplear en función de la cantidad y la diversidad de las personas intervinientes en el trabajo. Al ser la materia prima que se procesa información de tipo verbal, las fuentes de variación provienen de las posibilidades que tiene el procesamiento intelectual de dicha información. Al igual que lo especificado en la fase de codificación, las posibilidades de aplicar procedimientos distintos son múltiples, e igualmente son múltiples las posibilidades de conseguir datos con distintos niveles de procesamiento. Si bien un problema clásico en este ámbito es la introducción de sesgos de carácter ideológico o valorativo por parte de los entrevistadores tanto en la selección de entrevistados como en la generación de respuestas de una determinada tendencia ²², es igualmente relevante el problema de la formulación de las preguntas y la consignación de las variables, si no con criterios ideológicos, sí con criterios personales o de concepción técnica del trabajo.

Ante la posibilidad de existencia de múltiples criterios personales, el mecanismo de resolución suele ser el control de la variación. Al igual que en cualquier proceso de producción estandarizado, los criterios de calidad se realizan primordialmente mediante la introducción de controles de calidad que consisten sobre todo en normalizar las tareas que se realizan, anulando en lo posible la utilización de criterios arbitrarios en la ejecución del trabajo ²³ y primando las iniciativas individuales que mejoran la producción, si bien con el objetivo de extender los cambios resultantes al conjunto de los trabajadores. En el caso de la producción de datos sociales uno de los requisitos fundamentales es evitar la introducción de criterios dispares en la recogida de datos. Es decir, los criterios de los participantes deben de adaptarse a los establecidos en la formulación del problema y en el diseño de investigación, de ahí la importancia que se suele adjudicar a las normas que corresponden a cada estudio y a realización de *briefings* en los que se realiza una transmisión directa de indicaciones por parte del investigador al equipo de trabajo de campo.

No obstante, dicha transmisión dista frecuentemente de realizarse en términos óptimos, primero debido a la complejidad organizativa del trabajo de campo y segundo debido al perfil profesional y a la cualificación de estos trabajadores. Respecto a lo primero, ya se ha especificado las dificultades tanto objetivas como ideológicas que frenan dicho contacto. Respecto a lo segundo, las condiciones laborales y la formación de estos trabajadores es el limitador principal de la cali-

²² Sobre la introducción de sesgos provenientes de la ideología política de los encuestadores ver ALVIRA y MARTÍNEZ (1982). Sobre organización del trabajo de campo ver, por ejemplo, DÍAZ DE RADA (2001).

²³ Esto es uno de los conocidos principios del taylorismo. Ver CASTILLO (Comp.) (1988). Si bien aquí no se está de acuerdo con sus presupuestos, sí es cierto que la organización de la producción contemporánea sigue utilizando muchas de las nociones originales de la producción estandarizada, aunque introduciendo los elementos de la capacidad y la colaboración de los trabajadores en la mejora de los procesos de producción, y teniendo en consideración el «clima laboral» de las unidades productivas al comprobarse que la consideración del trabajo, las condiciones y las actitudes personales condicionan enormemente la rentabilidad y la calidad de cualquier bien que se produce industrialmente.

dad de los datos que se recogen. Así pues, el operativo fundamental para el trabajo de campo se concentra en las condiciones que favorecen dicha transmisión, en la interiorización homogénea por parte de los participantes y en la aplicación uniforme de los instrumentos que se diseñan, sujeto todo ello al control del proceso de trabajo. Al no participar en esta actividad los responsables del diseño, uno de los puntos claves es la disposición de un equipo de trabajo adecuado para llevar a cabo cada una de las tareas, esto es, los coordinadores de trabajo de campo que transmitan las instrucciones y distribuyan el trabajo de acuerdo con las capacidades y actitudes de los participantes, y los supervisores que detecten las posibles anomalías cuando surgen divergencias en la interpretación y consignación de los datos que se recogen. También en este sentido una cuestión fundamental es que el diseño muestral, realizado o no con las herramientas antes mencionadas, debe de ser operativo en términos de la organización concreta del trabajo. Es decir, el diseño debe de adaptarse a los medios disponibles y a las cualificaciones de los trabajadores, ofreciendo herramientas adicionales para la identificación de los puntos y las unidades muestrales y, sobre todo, considerando los medios disponibles en el conjunto de trabajadores, tanto en los desplazamientos como en la cantidad de trabajo que se adjudican ²⁴.

Las cualificaciones de los trabajadores de campo

Es relativamente sabido que la situación laboral y profesional de los trabajadores que intervienen en la recogida de estadísticas suele ser de carácter subordinado en términos de prestigio y condiciones económicas. A pesar de que esta figura está profesionalizada en las grandes organizaciones gubernamentales de estadística y en las corporaciones de investigación social y de mercados, la realidad de la mayoría de este sector productivo es que el trabajo se realice por parte de pequeñas y medianas empresas o por equipos de investigación reducidos en el ámbito académico. Dicha situación provoca que, excluyendo a organizaciones especializadas, las condiciones laborales de dichos trabajadores correspondan a una situación de subempleo, al realizarse los trabajos de forma estacional o a tiempo parcial, que la composición social esté formada por personas con una situación personal más adaptada a esta secuencia laboral (amas de casa, estudiantes, trabajadores estacionales o a tiempo parcial) y que la remuneración suela estar en la base de la pirámide de la estructura de ocupaciones de la sociología ²⁵.

²⁴ Por ejemplo, un recurso habitual para hacer frente a los desplazamientos a núcleos de población alejados para los que se requiere un tiempo importante es organizar el sistema de conglomerados a modo de puntos muestrales a los que se asigna una o varias «rutas». El número de unidades muestrales o entrevistas en cada ruta se puede establecer en función de la disponibilidad de tiempo para el desplazamiento y la realización de las rutas, toda vez que la cantidad sea compatible con el criterio estadístico para el establecimiento del tamaño del conglomerado. Ver TRUJILLO y FERNÁNDEZ (1997).

²⁵ La normativa laboral sobre los trabajadores de campo suele recoger las peculiaridades de dedicación de este colectivo, junto a las necesidades productivas del sector. Por ejemplo, no se con-

Por tanto, lo que cabe esperar de este tipo de trabajadores es que sean personas con escasa experiencia laboral, escasa cualificación, poca disponibilidad de tiempo para el trabajo, alta rotación y pocas posibilidades de formación específica. Aunque en España raramente se han realizado estudios sobre las condiciones y las actitudes de estos trabajadores, la información disponible muestra que en gran medida se nutren de la situación de subempleo y desempleo que ha caracterizado a la economía española en las últimas décadas. Por ejemplo, los ya antiguos estudios del CIS sobre su propia red de campo indican que la mayoría de los encuestadores alternan este trabajo con el desempleo, o bien se dedican a este trabajo como consecuencia de estar en una situación de desempleo²⁶. En suma, al margen de los escasos profesionales que se dedican a esta tarea de forma estable, cuando se organiza un trabajo de campo es necesario contar con un contingente de personas que se enfrentan por primera vez a las tareas de campo, o incluso a cualquier tipo de tarea profesional. Por consiguiente, la formación es un componente fundamental en la ejecución de la mayor parte de los trabajos de campo, siendo extremadamente arriesgado iniciar un trabajo de envergadura y complejidad técnica sin dedicar parte del tiempo y los recursos a transmitir los conocimientos necesarios específicos del estudio y al entrenamiento en condiciones reales de forma que se prevean los posibles errores y que se establezcan criterios homogéneos de resolución. Sin embargo, la realidad paradójica es que, siendo este proceso extremadamente complejo y exigiéndose una capacitación especializada tanto en la técnica específica como en las habilidades personales, en el trabajo de campo existe una práctica ausencia de formación tanto reglada como no reglada y, de hecho, en función de los recursos disponibles en cada proyecto también existe una ausencia de formación específica para las investigaciones que no cuentan con una amplia financiación.

Así mismo, lo que cabe esperar de la situación descrita es un procedimiento de trabajo sujeto a notables fuentes de error con importantes limitaciones para prevenirlas y, una vez realizadas, solventarlas, toda vez que la producción de datos sociales se realiza de forma puntual, no existiendo posibilidad de replicación excepto en las estadísticas oficiales y en aquellas efectuadas periódicamente por los institutos de opinión pública. Ante esta situación, la única forma de minimizar el riesgo cuando no es posible contar con una organización especializada es dedicar parte de los recursos a la formación y el entrenamiento, incluso a costa de la extensión de los propios estudios. Si bien esto es una opción sobre

templán jornadas laborales en función de un horario, sino en función de las unidades de trabajo que es necesario llevar a cabo, estableciéndose la remuneración a partir de la cantidad de unidades. Una interpretación del marco laboral de los encuestadores puede verse en CHACARTEGUI (1996). Una colección de artículos de opinión sobre el oficio de encuestador se incluye en VV.AA. (2000a).

²⁶ Los estudios del CIS sobre los trabajadores de campo llevan por título «Coordinadores y entrevistadores del CIS» y corresponden a las referencias 1145 (año 1977), 1379 (año 1983) y 1799 (año 1989). Ver fichas técnicas en <http://www.cis.es/bd>. Por otra parte, en los últimos años dicha situación está cambiando. La mejora de las posibilidades de empleo provoca que a esta tarea se dediquen colectivos cada vez más desfavorecidos y con menos cualificación, con el agravante de que muchos de ellos tienen limitaciones para realizar el trabajo debido a que son inmigrantes que tienen escaso conocimiento de la lengua estatal o de las lenguas autonómicas.

la que pueden existir opiniones contrarias, parece un hecho constatado que la representatividad de los datos, aunque sea limitada, es preferible que sea cierta, es decir, que sea válida y fiable, en contraste con una representatividad amplia basada en grandes muestras sobre la que se tiene escasa seguridad de su certeza.

El componente tecnológico

Por último, es conveniente realizar una breve referencia a las herramientas pertenecientes a las TIC que, si bien difícilmente pueden sustituir los elementos organizativos y formativos en la producción de datos, sí pueden contribuir a mejorar y facilitar algunas de las fases del proceso en tanto que incorporan elementos automatizados que pueden reducir importantes fuentes de error que tienen su origen en el componente humano y, además, detraer recursos de algunas tareas que obligan a aumentar los controles de calidad. Dichas herramientas intervienen en la investigación social en las fases de recogida y proceso de datos de dos posibles formas. De un lado, pueden sustituir el sistema tradicional de consignación de datos en papel mediante unidades informáticas en las que es posible introducir directamente los datos, eliminando el proceso de grabación. Las ventajas de dicho sistema, además de las económicas, consisten en eliminar posibilidades de error en la consignación al poder estar las posibilidades de introducción de datos predeterminadas, e igualmente en eliminar los posibles errores en la traslación de datos desde el papel al soporte informático, aunque presentan la limitación de la cualificación para su manejo y de la falta de flexibilidad para la consignación de observaciones que no están predeterminadas²⁷. De otro lado, en los casos en los que se utiliza el soporte de los cuestionarios en papel, un sistema habitual es la grabación de datos mediante captura óptica, lo que presenta ventajas y limitaciones semejantes a las anteriores. Existe una reducción de tiempos y costes, una disminución de los errores que provoca el proceso manual y una mayor facilidad para detectar y corregir anomalías, si bien un problema frecuente suele ser la adaptación del sistema de consignación en papel a un formato compatible con los aparatos y el software disponible, junto a la relativa complejidad del proceso, lo cuál lo hace operativo cuando existe un suficiente volumen de datos de modo que haga posible las economías de escala y resulte rentable su empleo.

Mención aparte merecen dos procedimientos utilizados para la recogida de datos de encuesta, como son el sistema de entrevistas telefónicas y el sistema de entrevistas por correo, ya más electrónico que convencional. Ambos sistemas se comenzaron a utilizar debido a dos motivos: la reducción de costes y tiempos respecto a las entrevistas presenciales y su mejor adecuación a ciertos problemas de estudio. Respecto a lo primero, es cierto que el desarrollo de los sistemas de telefonía y su extensión a la práctica totalidad de la población posibilitan la rea-

²⁷ El sistema habitual es el llamado CAPI, aunque recientemente se están incorporando las agendas electrónicas tipo PDA adaptadas al proceso de recogida de datos. Sobre la utilización de nuevas tecnologías para la recogida de datos en sociología ver DÍAZ DE RADA (2000).

lización de encuestas telefónicas de una forma fácil, económica y rápida, que ofrece a la vez unos niveles de calidad aceptables dependiendo del marco poblacional y los objetivos de los estudios. Respecto a lo segundo, la dispersión de algunas poblaciones o su difícil accesibilidad, junto a la existencia de temas y colectivos sociales que por el anonimato o por el interés que suscitan los asuntos estudiados se muestran más accesibles a la entrevista telefónica, también han emplazado a esta técnica como una de las más frecuentes²⁸.

Sin embargo, la proliferación de encuestas telefónicas mediante sistemas automatizados —habitualmente el sistema llamado *Computer Assisted Telephonic Interview*, aunque existen varios sistemas similares en el mercado— ha respondido más a lo primero que a lo segundo, es decir, se realizan como forma de ahorrar los recursos casi siempre escasos en la investigación social y de solventar los problemas de las redes de campo que tienen un elevado coste de mantenimiento al tener que estar funcionando periódicamente para poder garantizar su operatividad. A pesar de ello, su adecuación es discutible en bastantes supuestos. No se quiere entrar aquí en los problemas de las encuestas telefónicas, que por otra parte coinciden bastante con las encuestas presenciales en lo referido a las condiciones sociales de trabajo, sino en una nueva circunstancia que provoca problemas de representatividad a raíz de la rápida expansión de los sistemas de telefonía móvil. Es sabido que en casi todos los países desarrollados existen más teléfonos móviles que fijos, a pesar de que la telefonía fija siga teniendo una penetración mayoritaria. Este hecho, si bien no invalida la representatividad de los marcos muestrales de teléfonos fijos, que con algunas diferencias suelen corresponder a los marcos muestrales de los hogares, sí plantea una dificultad cuando se trata de identificar a los respondentes con la población en general y con el perfil de los respondentes a los teléfonos móviles en particular. A pesar de existir más teléfonos móviles que fijos, no existe un registro de números de teléfono móvil y, si existe, es por definición incompleto, al estar el acceso a esta tecnología sujeto a cambios extremadamente rápidos. E, igualmente, a partir de la liberalización del mercado de telefonía fija es factible que cada vez se incluyan menos de los teléfonos realmente existentes en un registro centralizado, como es la guía de teléfonos.

En dichas circunstancias, la información disponible obliga a que las muestras en las encuestas telefónicas a población general tengan que basarse en los registros de teléfonos fijos disponibles, con lo que los realmente encuestados serán los que tienen más probabilidad de tener y usar estos teléfonos. Dicho factor, unido a que los usuarios del teléfono móvil son socioeconómicamente distintos a los del fijo, es lo que añade un sesgo adicional a las encuestas telefónicas, forzosamente basadas en muestras construidas a partir de marcos poblacionales limitados a los teléfonos fijos o a algunos listados de móviles. Este problema, por otra parte irresoluble con esta técnica, es usualmente minimizado a través de la calibración de las muestras y mediante la introducción de criterios de selección previos, que en

²⁸ Sobre encuestas telefónicas y por correo ver BOSH y TORRENTE (1993), BIEMER, et. al. (1988). Para encuestas vía Internet puede verse el especial de la Revista de Metodología de Encuestas (VV.AA., 2000b). Una exposición comparada de los distintos métodos de encuestación se encuentra en FINK, et. al. (1995).

todo caso añaden complejidad al trabajo, como son el empleo de cuotas de acuerdo con criterios sociodemográficos entre los que se pueden incluir el perfil de usuarios de teléfono móvil. Si las encuestas telefónicas estaban sesgadas hacia arriba, esto es, no incluían a personas de clase baja por no disponer de acceso o por no estar habituados a su uso, ahora también pueden estar sesgadas hacia abajo, es decir, no incluir a personas, si no de clase alta, sí de características educativas y culturales determinadas, las que van asociadas con la utilización primordial de las TIC en sustitución de otras tecnologías más tradicionales, a saber, mayor juventud y mayor nivel de estudios y concentración en zonas urbanas²⁹.

Los problemas de las encuestas telefónicas son en ciertos aspectos parecidos a los de las encuestas por correo, que están siendo sustituidas en algunos casos por encuestas vía Internet. A saber, las encuestas por correo se han considerado adecuadas para el estudio de sujetos que, bien por su dispersión, bien por su difícil accesibilidad, eran más factibles de contactar a través del correo, y para temas que resultaban más adecuados plantearlos en formato escrito debido a su complejidad o su extensión. A cambio, exigían que dichos sujetos estuviesen identificados en un marco poblacional de nombres y direcciones y que dispusieran de un nivel cultural homogéneo y suficiente para entender y responder las preguntas del cuestionario escrito, además de una disposición favorable a la contestación. Todas estas cuestiones son válidas para las encuestas por correo electrónico, si bien existe un problema específico además de los sesgos de la disposición de esta tecnología, como es la distinta disposición y hábito de uso. En suma, las encuestas vía Internet son adecuadas cuando los elementos de la población a muestrear dispongan de un acceso mayoritario al correo electrónico, dispongan también de una dirección propia, y el conjunto de direcciones esté disponible en un registro que permita extraer muestras aleatorias. Además, el uso de dicho correo debe ser también habitual y familiar en términos de facilidad y hábito de lectura y escritura con un teclado de ordenador.

Para el conjunto de problemas mencionados difícilmente existe una solución distinta a la de reconocer la limitación de las nuevas tecnologías a ciertos problemas y poblaciones de estudio, a la vez que su adecuación a otras muchas otras realidades para las que las técnicas de encuestación habituales resultan, si no inapropiadas, sí más costosas en términos económicos y de esfuerzo, y que consisten en los temas que resultan más receptivos por parte de los posibles encuestados y en poblaciones que son de carácter homogéneo en la disposición al uso de la tecnología en cuestión.

5. CONCLUSIONES

Como otras muchas ciencias, la sociología dispone de una tecnología social que utiliza experiencias y conocimientos de la propia sociología y de disciplinas

²⁹ Sobre el uso diferencias de las tecnologías de la información y las comunicaciones, incluidas telefonía e Internet, ver OECD (2001). Para España ver datos de la Encuesta General de Medios (AIMC, 2002).

y tecnologías adyacentes, como pueden ser la estadística, la psicología o la teoría de la organización. En este caso concreto, los dispositivos para obtener y analizar datos sociales representativos en particular, y la encuesta social en general, se pueden considerar como una tecnología social que provee información de carácter sociológico, que permite un autoconocimiento de las sociedades en las que se emplea y que constituye la base para la prueba y la extensión del pensamiento sociológico si los problemas son formulados en términos científicos. Dicha tecnología dispone de un proceso de producción específico con sus propios requerimientos y limitaciones que, debido al estadio de desarrollo de la investigación sociológica, se realiza en unas condiciones de producción que limitan el proceso en unas cotas distintas a las de los procedimientos científicos y técnicos propios de la sociología. En dicho contexto, una estrategia para superar las limitaciones es establecer mecanismos que prevengan y supervisen el nivel de calidad de los datos que se generan, introduciendo técnicas y dispositivos concretos de modo que se minimicen las probabilidades de que aparezcan fuentes de error, considerando para ello las condiciones de trabajo que intervienen en cada investigación concreta. Esto quiere decir que, habitualmente, la complejidad y la extensión de los datos que se pueden obtener se definen por las cualificaciones de los trabajadores que intervienen en el proceso y por la adecuación de los recursos disponibles. Por tanto, los objetivos que se pueden establecer en una investigación cuantitativa respecto a su complejidad y respecto a la representatividad de los datos sociales en los que se basa se ven circunscritos por una realidad en parte ajena a la configuración interna de la investigación científica, que viene determinada por la realidad social en la que se encuadra el proceso de producción de datos.

BIBLIOGRAFÍA

- ALONSO, L. E. (2001): La mirada cualitativa en sociología, Madrid: Fundamentos
- ALVIRA, F. y MARTÍNEZ, E. (1982): El efecto de los entrevistadores sobre las respuestas de los entrevistados en encuestas de opinión, *Revista Española de Investigaciones Sociológicas*, N.º 29, pp. 219-225.
- AIMC (2002): Estudio General de Medios —primera oleada— (Resumen de resultados), Edición en web: <http://aimc.es/>.
- ATIENZA, J. y NOYA, F.J. (1999): La encuesta como interacción social. Una aproximación empírica, *Empiria. Revista de Metodología de Ciencias Sociales*, N.º 2, pp. 73-92.
- BELTRÁN, M. (1991): La realidad social, Madrid: Tecnos.
- BIEMER, P., GROVES, R. M. y LYBERG, L. E. (1988): *Telephone survey methodology*, New York: Wiley.
- BOLTON, R. (1993): Pretesting Questionnaires: Content Analyses Of Respondents' Concurrent Verbal Protocols, *Marketing Science*, N.º12, pp. 280-303.
- BOSCH, C. L. y TORRENTE, D. (1993): Encuestas telefónicas y por correo, Madrid: CIS, Cuadernos Metodológicos, N.º 9.
- BRADBURN, N. M., SUDMAN, N., et al. (1991): Improving interview method and questionnaire design: response effects to threatening questions in *Survey Research*, Ann Arbor: Jossey-Bass.
- BRAVERMAN, H. (1974): Labor and monopoly capital. The degradation of work in Twentieth century, New York: Monthly Review Press.

- CALLEJO, J. (1988): Sobre el uso conjunto de prácticas cualitativas y cuantitativas, *Revista Internacional de Sociología* (Nueva Época), N.º 21, pp. 101-126.
- CASTILLO, J. J. (Comp.) (1988): *Las nuevas formas de organización del trabajo*, Madrid: Ministerio de Trabajo, Colección Informes.
- CASTILLO, J. J. y PRIETO, C. (1983): *Condiciones de Trabajo*, Madrid: CIS
- CHACARTEGUI, C. (1996): *¿Son los encuestadores trabajadores?*, Madrid: Tecnos. Colección *Jurisprudencia Práctica*.
- CICYT (1997, 2001): *Plan Nacional de I+D. Memoria de actividades (años 1996 y 2000)*, Madrid: Ministerio de Educación y Cultura.
- COCHRAN, W. (1996): *Técnicas de Muestreo*, Madrid: Wiley.
- CONDE, F. (1994): Una propuesta de uso conjunto de los técnicas cuantitativos y cualitativos en la investigación social, *Revista Española de Investigaciones Sociológicas*, N.º 39, pp. 213-224.
- CONVERSE, J. M. y PRESSER, S. (1986): *Survey questions: handcrafting the standardized questionnaire*, London: Sage.
- DEMAIO, Th. (1993): *Approaches to Developing Questionnaires*. Bureau of the Census. Office of Management and Budget. *Statistical Working Papers*, N.º 10.
- DÍAZ DE RADA, V. (2000): Utilización de nuevas tecnologías para el proceso de recogida de datos en la investigación social mediante encuesta, *Revista Española de Investigaciones Sociológicas*, N.º 91, pp. 137-166.
- (2001): *Organización y gestión de los trabajos de campo en encuestas personales y telefónicas*, Barcelona: Ariel.
- DOERINGER, P. B. and PIRE, M. J. (1971): *Internal labour markets and Manpower analysis*, Lexington, M.A.: Heath Lexington Books.
- EFRON, B. y TIBSHIRANI, R. (1993): *An introduction to the «bootstrap»*, London: Chapman and Hall.
- FINK, A. et. al. (1995): *The survey kit*, London: Sage.
- GARCÍA FERRANDO, M., ALVIRA, F. e IBÁÑEZ, J. (Coords.) (1986): *La realidad social*, Madrid: Alianza Editorial.
- y GARCÍA, M. (1986): Las encuestas a encuesta. *Revista Española de Opinión Pública*, N.º 45, pp. 266-292.
- GIBBONS, M., LIMOGES, C., NOWOTNY, H., SCHWARTZMAN, S., SCOTT, P., y TROW, M. (1994): *The new production of knowledge. The dynamics of science and research in contemporary societies*, London: Sage.
- GIFI, A. (1990): *Nonlinear Multivariate Analysis*, Chichester: Wiley.
- HUNT, S. D., SPARKMAN, R. D. y WILCOX, J. B. (1982): The Pretest in Survey Research: Issues and Preliminary Findings, *Journal of Marketing Research*, Vol. XIX, pp. 269-273.
- JUSTEL, M. (1981): Las encuestas a encuesta, de nuevo, *Revista Española de Investigaciones Sociológicas*, N.º 13, pp. 155-181.
- KISH, L. (1995): *Survey Sampling*, London: Wiley.
- KÖNIG, R. (1973): *Tratado de sociología empírica*, Madrid: Tecnos.
- LAZARSFELD, P. (1954): *The language of social research*, New York: The Free Press.
- LITTER, G. R. (1982): *The development of labour process in capitalist societies*, London: Heinemann.
- MILLS, C. WRIGHT (1968): *La imaginación sociológica*, México: FCE.
- OECD (2001): *Understanding the digital divide*, París: OECD Publications.
- PRÜFER, P., y REXROTH, M. (1996): Verfahren zur evaluation von survey-fragen: ein überblick, *Nachrichten-ZUMA*, N.º 39, pp. 95-115.
- REINECKE, J. y SCHMIDT, P. (1993): Explaining interviewer effects and respondent behavior: theoretical models and empirical analysis, *Quality & Quantity*, N.º 27, pp. 219-247.

- SÁNCHEZ CARRIÓN, J. J. (1996): La calidad de la encuesta, *Papers, Revista de Sociología*, N.º 48, pp. 127-146.
- SELLTIZ, C. et. al. (1980): *Métodos de investigación en las relaciones sociales*, Madrid: Rialp
- SPSS (1997): *SPSS for Windows. Categories*, release 6.0, Chicago: SPSS Inc.
- TRUJILLO CARMONA, M. y FERNÁNDEZ ESQUINAS, M. (1997): *Diseño de muestras y organización del trabajo de campo en encuestas a población general*, Córdoba: IESA, Documentos de Trabajo Ref. DT-97-01
- (2000): Optimización del error muestral en encuestas a población general: criterios de estratificación y cálculo de error con «Wesvar Complex Samples», *Metodología de Encuestas*, Vol. 2, N.º 2, pp. 199-210.
- VERBA, S. et. al. (2001): *La inferencia científica en la investigación cualitativa*, Madrid: Alianza Editorial.
- VILLÁN CRIADO, I. (1990): *Procedimientos de depuración de datos estadísticos*, Vitoria: Instituto de Estadística del País Vasco.
- VV.AA. (2000a): *El oficio de encuestador (Sección Especial)*, *Metodología de Encuestas*, Vol. 2, N.º 2.
- (2000b): *Encuestas en Internet (Sección Especial)*, *Metodología de Encuestas*, Vol. 4, N.º 1.
- WALLACE, C. (1977): *La lógica de la ciencia en la sociología*, Madrid: Alianza Editorial.
- ZIMAN, J. (1983): *The collectivization of science (The Bernal Lecture)*, *Proceedings of The Royal Society of London*, B 219, pp. 1-19.
- (1986): *Introducción al estudio de las ciencias*, Barcelona: Ariel.

RESUMEN

Cuando se pone en marcha una investigación empírica en el ámbito de la sociología la literatura especializada suele remitir a las fases consistentes en el establecimiento del objeto de estudio, la formulación del problema y el diseño de la investigación, así como a los criterios técnicos para su ejecución. Sin embargo, otros elementos fundamentales que determinan la calidad de los datos que se obtienen tienen que ver con el proceso de trabajo en la generación de dichos datos, que reúne características similares a las de un proceso de producción industrial. En este trabajo se exponen los fundamentos a tener en cuenta en el proceso de producción de datos sociales a partir de las controversias que surgen en la secuencia de la investigación sociológica. Una vez analizados los condicionantes técnicos y organizativos en las fases típicas de una investigación cuantitativa, se realiza una exposición de las estrategias que se pueden aplicar para garantizar la calidad de los datos. Para ello se tienen en cuenta las condiciones de trabajo que subyacen a las tareas concretas en las fases de recogida y procesamiento de los datos y los elementos técnicos, organizativos y tecnológicos que es posible emplear para minimizar las fuentes de error que surgen en dicho proceso.

ABSTRACT

The specialized literature in sociology usually refers to the various phases of empirical research as process which involves identifying the object to be studied, formulating the problem, designing the study and establishing the necessary technical criteria. Howe-

ver, the quality of data is determined also by other fundamental elements which have to do with the process of generating data, a process which is similar to that of industrial production. This study describes the fundamentals of producing social data from the viewpoint of specific issues that arise during the different phases of sociological research. The author analyzes the technical and organizational elements in the typical phases of quantitative research (the making of a sociological survey) and proposes several strategies to guarantee data quality. With this objective, the author considers specific tasks such as data collection and data processing and the technical, organizational and technological elements that can be employed to minimize error.