

Construyendo las bases para una comparación fiable: la Encuesta Social Europea¹

Clara Riba y Anna Cuxart²

Departament d'Economia i Empresa

Universitat Pompeu Fabra

C. Ramon Trias Fargas, 25-27

08005 Barcelona

Fax: 93 542 17 46

clara.riba@upf.edu, anna.cuxart@upf.edu

Keywords: European survey, comparative studies, sampling design, design effect, effective sample size

Palabras clave: encuesta europea, estudios comparados, diseño muestral, efecto de diseño, tamaño muestral efectivo

Journal of Economic Literature Classification: C42

¹ Este trabajo ha sido posible gracias a las ayudas SEC2002-03363 y BFM2000-1064 del MCyT

² Las autoras son miembros del equipo que implementa la Encuesta Social Europea en España

Resumen

Las encuestas constituyen un instrumento valioso de conocimiento de la realidad social y política de nuestro entorno pero la labor de los investigadores queda en muchas ocasiones limitada por una serie de deficiencias en dos grandes aspectos. Por un lado, las muestras acostumbran a ser de escasa calidad técnica y el trabajo de campo no se realiza en las mejores condiciones. Por otro, las encuestas no han sido expresamente diseñadas para permitir la realización de estudios comparados, objetivo frecuente en la investigación politológica. El artículo presenta la Encuesta Social Europea y reflexiona sobre las bases metodológicas en las que se sustenta. Dicha encuesta, promovida por la European Science Foundation y la Comisión Europea, nace con el objetivo explícito de establecer unos requisitos de calidad en el diseño muestral y en la ejecución del trabajo de campo que garanticen la calidad de los datos y permitan la comparabilidad entre países.

Palabras clave: encuesta europea, estudios comparados, diseño muestral, efecto de diseño, tamaño muestral efectivo

Abstract

Surveys are a valuable instrument to find out about the social and political reality of our context. However, the work of researchers is often limited by a number of handicaps that are mainly two. On one hand, the samples are usually low technical quality ones and the fieldwork is not carried out in the finest conditions. On the other hand, many surveys are not especially designed to allow their comparison, a precisely appreciated operation in political research. The article presents the European Social Survey and justifies its methodological bases. The survey, promoted by the European Science Foundation and the European Commission, is born from the collective effort of the scientific community with the explicit aim to establish certain quality standards in the sample design and in the carrying out of the fieldwork so as to guarantee the quality of the data and allow each comparison between countries.

Keywords: European survey, comparative studies, sampling design, design effect, effective sample size

INTRODUCCIÓN

Las ciencias sociales en Europa tienen una larga tradición empírica. La gran diversidad cultural, política y social existente en este continente lo convierten en un escenario privilegiado para el análisis comparado de valores, actitudes y comportamientos. Sin embargo, la mayor parte de la investigación realizada hasta el momento no puede considerarse comparativa puesto que o bien no se dispone de los indicadores de interés para todos los países o bien los datos existentes han sido recogidos de formas tan diferentes que resulta difícil su estandarización. Si bien es cierto que la mayor parte de los países recogen de manera sistemática datos individuales y datos relativos a los hogares, a través de sus institutos de estadística o de otras agencias, en general, estos datos son de difícil acceso para los investigadores y acostumbran a ser poco comparables. Por ejemplo, la diversidad de sistemas de educación vigentes en Europa y los distintos criterios de clasificación usados por los países en la recogida de datos impiden comparar de manera clara las distribuciones de los niveles educativos de sus poblaciones³.

La Encuesta Social Europea (ESS) tiene como objetivo cubrir estas limitaciones mediante la realización de una encuesta bianual que facilite el estudio comparado y longitudinal de los valores, actitudes y comportamientos de los ciudadanos europeos en los ámbitos económico, social y político. La ESS también pretende facilitar el análisis de la

³ Merece la pena mencionar aquí el esfuerzo realizado en este sentido por la UNESCO, institución que ha homogeneizado la definición de los niveles educativos de los distintos países a través de la definición de los International Standard Classification of Education Levels (ISCDE-97), aprobada en Ginebra en 1997.

interacción entre el contexto y las instituciones sociales de cada país con los valores, las actitudes y patrones de comportamiento de sus ciudadanos.

La ESS nace con una clara orientación académica, queriendo convertirse en un instrumento de referencia en la investigación europea en ciencias sociales. En su diseño ha participado un grupo internacional de académicos de reconocido prestigio que han puesto a disposición de la ESS su experiencia en la elaboración y tratamiento de diversas encuestas nacionales e internacionales de solvencia contrastada. La metodología seguida, el texto del cuestionario y los resultados serán de acceso libre para toda la comunidad científica (www.europeansocialsurvey.com).

La Encuesta es la culminación de un trabajo metodológico de casi cinco años promovido por la European Science Foundation. Dicha fundación, junto con la Comisión Europea, es la que financia la estructura organizativa del proyecto, mientras que cada país participante se hace cargo de los gastos de implementación y del trabajo de campo en su propio país. La Encuesta se ha administrado por primera vez en otoño-invierno de 2002, de manera conjunta en 23 países. Éstos son:

Alemania	Austria	Bélgica	Dinamarca	Eslovenia
España	Finlandia	Francia	Grecia	Holanda
Hungría	Irlanda	Israel	Italia	Luxemburgo
Noruega	Polonia	Portugal	Reino Unido	República Checa
Suecia	Suiza	Turquía		

La coordinación de todo el proceso de diseño y ejecución de la ESS recae en un equipo internacional de expertos dirigido por el profesor Roger Jowell del National Center for Social Research de Londres. Este grupo constituye el Central Coordinating Team (CCT), encargado de establecer los criterios, definir los estándares de calidad y marcar las directrices de ejecución de todo el proceso.

El CCT dispone de una serie de grupos asesores a los cuales consulta sobre aspectos específicos de la encuesta, como son la traducción del cuestionario y el proceso de muestreo. Entre estos grupos, debe mencionarse por su especial relevancia el Comité Científico Asesor formado por un representante de cada agencia nacional más un representante de la European Science Foundation y dos de la Comisión Europea.

Además, en cada uno de los países participantes existe un Coordinador Nacional con su correspondiente equipo que se encargan de aplicar las directrices acordadas por el CCT.

El Norwegian Social Science Data Services (NSD) es el organismo encargado del archivo de las series de datos de cada país, de su integración en un archivo común y de su posterior divulgación.

En primer lugar se expone cuáles son las características de la ESS y se reflexiona acerca de las razones metodológicas que las justifican. En segundo lugar se explica cuál ha sido el proceso de implementación de la primera ola en España. Se finaliza con una reflexión crítica y una valoración de sus resultados.

CARACTERÍSTICAS DE LA ESS

Población. La población objeto de estudio está formada por todos los residentes en domicilios privados de edad igual o superior a quince años.

Periodicidad. La ESS será una encuesta bianual.

Cuestionario. La encuesta consta de dos cuestionarios de características y duración distintas. El primero se administra cara a cara y tiene una duración aproximada de una hora, mientras que el segundo se deja en el domicilio del entrevistado para que éste lo complete personalmente y tiene una duración de siete minutos. La forma de recoger este segundo cuestionario depende de cada país. En España, por ejemplo, la opción que se ha tomado es que sea el entrevistador el encargado de recoger el cuestionario auto administrado una vez que ha sido completado.

El cuestionario cara a cara consta de una primera parte de contenido fijo, que se replicará en cada ola, y de una segunda parte de contenido variable. La parte fija consiste en una serie de preguntas referentes al individuo y su contexto, donde se abordan temas que incluyen actitudes y atributos individuales, opiniones y actitudes políticas, posición social, participación en redes, contexto social y entorno del individuo. Con ello se pretende medir los comportamientos y opiniones de los europeos en un determinado momento y su evolución a lo largo del tiempo, así como posibilitar el estudio de la

interacción entre individuo y contexto. La parte variable consiste en dos módulos específicos sobre preocupaciones académicas o políticas en Europa que, al mismo tiempo, han de permitir la experimentación de nuevas propuestas metodológicas. Es la propia comunidad científica la que, a través de un concurso público convocado por el CCT, propone los temas a incluir en cada ola y se encarga de la redacción preliminar de las preguntas correspondientes. Para la ola 2002, los módulos específicos seleccionados han sido sobre inmigración y sobre capital social, participación y democracia. El tema de inmigración ha sido elaborado por el equipo dirigido por Ian Preston (Dustmann and Preston, 2001), y el diseño del segundo tema ha sido responsabilidad del equipo del proyecto europeo “Citizenship, Involvement and Democracy” que dirige Jan van Deth (van Deth, 2000), de la Universidad de Manheim, y en el que participan investigadores españoles vinculados al equipo de la ESS en España.

El cuestionario auto administrado consta de dos partes. La primera incluye 21 ítems de la escala de valores humanos de Schwartz (Schwartz *et al.*, 2001). La segunda consta de un total de 36 preguntas, de las cuales a cada entrevistado se le plantean solamente seis, orientadas a la evaluación de la validez y fiabilidad de una selección de preguntas e indicadores del primer cuestionario.

Requisitos técnicos

- La selección de la muestra debe ser probabilística en todas sus etapas.
- No se admiten sustituciones de casos de no contacto y de no respuesta.

- Se exige el conocimiento y documentación de la probabilidad de cada individuo seleccionado de formar parte de la muestra.
- El número de entrevistas completadas ha de ser como mínimo de 2000 y el tamaño efectivo de la muestra debe ser como mínimo 1500.
- Se exige una tasa de respuesta del 70% en el primer cuestionario. En el segundo cuestionario la tasa de respuesta mínima ha de ser del 90% sobre el total de los que han contestado al primer cuestionario.
- El cuestionario, elaborado originalmente en inglés británico, se traduce a la lengua de cada país. En aquellos países en los que alguna lengua minoritaria sea utilizada como primera lengua por un mínimo del 5% de la población, se deberá traducir también el cuestionario a esa lengua.

Diseño muestral. Cada país participante elabora su propio diseño muestral en función de la disponibilidad o no de listas de individuos o de domicilios que puedan ser usadas como marco de muestreo. Antes de su aplicación, cada país debe especificar por escrito su propuesta de diseño muestral y enviarla al CCT para su aprobación. El CCT remite dicho proyecto al grupo de expertos en muestreo para que lo supervisen y den su conformidad o bien sugieran alternativas de mejora.

Trabajo de campo. El trabajo de campo deberá ser realizado simultáneamente en todos los países participantes, con una duración mínima de un mes y máxima de cuatro. Las especificaciones dadas por el CCT pretenden garantizar la preparación, la realización, el control y la documentación de todo el proceso:

- *Preparación.* Se requiere un esfuerzo especial en la formación de los agentes que han de llevar a cabo el trabajo de campo. El coordinador nacional de la ESS ha de participar personalmente en las sesiones de formación.
- *Realización.* Cada entrevistador debe realizar un máximo de dos bloques de 24 entrevistas. El primer contacto con el entrevistado ha de ser personal. Una vez realizado este primer contacto, se puede concertar la entrevista por teléfono. La entrevista siempre ha de ser cara a cara. No se puede considerar un caso como no contacto mientras no se haya intentado un mínimo de cuatro contactos, de los cuales uno ha de ser forzosamente en horas vespertinas y otro en fin de semana.
- *Control.* Se han de realizar controles de calidad sobre cómo se han efectuado las entrevistas, contactando con un 5% de los entrevistados, con un 10% de los que han rechazado ser entrevistados y con un 10% de los no contactados. La evolución del trabajo de campo ha de ser supervisada de cerca por el equipo de la ESS de cada país. Se ha de elaborar un informe quincenal que incluirá, entre otras cosas, la tasa de respuesta conseguida hasta el momento y que desglose las “negativas claras” de las “negativas suaves” a ser entrevistados. Las “negativas suaves” han de ser asignadas de nuevo a otro entrevistador que reiniciará el proceso de contacto.
- *Documentación.* Durante la realización del trabajo de campo deben documentarse todos los intentos de contacto y todas las incidencias ocurridas para su posterior plasmación en un fichero que deberá incluir todos los casos de la muestra (y no solamente los correspondientes a entrevistas completadas). Además, el entrevistador deberá recoger información sobre el entorno físico del edificio donde tiene su domicilio el individuo seleccionado. Asimismo, la empresa encuestadora ha de

elaborar un informe técnico final sobre el desarrollo del trabajo de campo incluyendo las tasas de respuesta alcanzadas.

Información sobre contexto. El fichero de datos individuales procedente de los cuestionarios irá acompañado de un fichero de datos contextuales sobre características del país y sucesos puntuales que se hayan producido durante el período de realización del trabajo de campo. Se prevé que la procedencia de la primera parte de estos datos sea EUROSTAT, OCDE y otras fuentes oficiales. Estos datos aportarán información sobre la estructura institucional, económica, política y social de cada país. El segundo conjunto de datos será expresamente recogido por los equipos nacionales que registrarán los sucesos políticos, culturales y sociales que se hayan producido durante el año en que se administra la encuesta y que puedan haber afectado las respuestas individuales.

JUSTIFICACIÓN DE LAS CARACTERÍSTICAS

Se exponen a continuación las razones metodológicas que han llevado al CCT a establecer como estándares de la ESS las características detalladas en el apartado anterior.

Se exige que la muestra sea totalmente probabilística. Esto quiere decir que no se admiten cuotas ni otros procedimientos en los que no intervenga el azar. La muestra puede ser polietápica, pero no se admite la práctica habitual de aplicar cuotas en la etapa

final del muestreo. La aplicación de cuotas rompe con el supuesto probabilístico subyacente en toda la teoría de la inferencia estadística, introduciendo sesgos de magnitud desconocida. Si la muestra no es probabilística carece de sentido aplicar las técnicas de inferencia. Sin embargo, con frecuencia no se tiene en cuenta este condicionante y se realizan inferencias sin conocer la probabilidad de cada individuo de ser incluido en la muestra.

Especificación por escrito del diseño muestral. Dadas las diferentes circunstancias de cada país, es poco factible pensar en un diseño muestral común a todos ellos. Ello no es óbice para exigir la justificación de un diseño aplicado a cada caso que aproveche al máximo las posibilidades existentes, y para su sometimiento a la evaluación del comité de expertos. Éstos pueden anticipar posibles problemas en el diseño planteado y sugerir opciones de mejora. En concreto, las especificaciones dadas por el CCT obligan a los comités nacionales a tomar en cuenta los efectos de diseño que aparecen en la aplicación de muestreos polietápicos. El *efecto de diseño* es la reducción total en la precisión de las estimaciones derivadas de la muestra escogida en comparación con la precisión de las estimaciones resultantes de una muestra aleatoria simple de igual tamaño (Kish, 1965)⁴. El uso de conglomerados introduce una pérdida de eficiencia (que se acostumbra a expresar en términos de la varianza de los estimadores y de los errores de muestreo) debido a que las circunstancias que agrupan a los individuos en una misma unidad de muestreo suelen provocar que estos individuos tengan, también, características (opiniones, comportamientos...) similares. En consecuencia, la información de los

⁴ Se sobreentiende que un muestreo aleatorio simple no requiere diseño. El diseño aparece en muestreos de mayor complejidad.

individuos que provienen de conglomerados, sobretodo si los conglomerados son pocos, suele ser redundante y menos informativa que la que se obtendría con una muestra del mismo número de individuos pero dispersa por toda la población. Así mismo, el CCT exige el cálculo de las probabilidades de inclusión en cada una de las etapas del muestreo. ¿Por qué esta exigencia del CCT? Porque se quiere conseguir una muestra equiprobabilística en la que todos los individuos tengan la misma probabilidad de pertenecer a la muestra. Para este objetivo, lo más sencillo sería escoger una muestra aleatoria simple de individuos (Silva, 2000). Pero si esto no es posible, como más adelante veremos en el caso español, se deberá corregir las estimaciones ponderando adecuadamente los casos para compensar la sobre representación de determinados colectivos. Las probabilidades de inclusión son la información imprescindible para el cálculo de dichas ponderaciones.

El tamaño mínimo efectivo de la muestra debe ser 1500. Ello significa que, independientemente del diseño muestral aplicado, del efecto de diseño resultante y de la tasa de respuesta conseguida, el número de entrevistas completadas ha de ser suficiente para garantizar que el error de estimación sea menor o igual que el que se derivaría de una muestra aleatoria simple de 1500 casos. La razón de esta exigencia se halla en que el muestreo totalmente probabilístico, con un mismo tamaño efectivo es la garantía de la homogeneidad en la precisión de los estimadores que se derivarán de cada muestra. Este requisito facilita la comparación de resultados entre países al permitir el cálculo del margen de error máximo de cada estimación sin necesidad de conocer el diseño muestral seguido en cada país. Desgraciadamente no es habitual hablar en términos de “tamaño

efectivo” cuando se trata de calcular el margen de error de las estimaciones. Los diseños muestrales de encuestas sociales suelen ser complejos, dando lugar a estimaciones menos precisas que las derivadas de un muestreo aleatorio simple. Sin embargo, cuando se utiliza el número de entrevistas completadas (en lugar del tamaño efectivo) para el cálculo del margen de error, se presupone, erróneamente, que el muestreo ha sido aleatorio simple (ver Anexo). El requisito del CCT de obtener un mismo tamaño efectivo para todos los países permite al investigador conocer el margen de error común para las estimaciones y detectar a simple vista cuáles de las diferencias observadas son realmente significativas.

Utilización de dos formatos de entrevista. La ESS ha optado por el formato de entrevista cara a cara para el primer cuestionario debido a su longitud y para garantizar la cobertura de la población en todos los países. Por razones sobradamente conocidas, se descartó la entrevista telefónica (falta de cobertura de la red telefónica en algunos países o creciente uso de telefonía móvil en otros), la encuesta por correo ordinario (baja tasa de respuesta) o la encuesta por correo electrónico (escasa implantación de nuevas tecnologías). En cambio, se ha optado por un formulario auto administrado para el segundo cuestionario. Esta decisión se justifica por la privacidad del primer bloque de preguntas de este cuestionario (escala de valores de Schwartz). Además, la cumplimentación de este segundo cuestionario está prácticamente asegurada ya que el entrevistado ha aceptado previamente participar en la encuesta.

Medida de la validez y la fiabilidad de una selección de preguntas. El segundo bloque de preguntas del cuestionario auto administrado ha sido diseñado básicamente con el objetivo de contrastar la validez y fiabilidad de algunas preguntas seleccionadas. Realizar este contraste es importante por dos motivos. En primer lugar, porque muchas de las preguntas son indicadores de conceptos que no son directamente medibles. En ocasiones estas preguntas han sido utilizadas repetidamente en encuestas de ámbito nacional, considerándose que tienen probada validez en su correspondiente país, pero no existen garantías de que también sean válidas en otros contextos. En segundo lugar, porque un mismo concepto puede ser medido, no solamente a través de diferentes indicadores (preguntas), sino también a través de diferentes instrumentos (escalas de medida). En tales situaciones conviene conocer la capacidad que tiene cada uno de ellos para medir el concepto (validez) y la estabilidad de los resultados que proporciona cuando se aplica repetidamente (fiabilidad). En el ámbito de las ciencias políticas y sociales, donde muchos de los conceptos son variables latentes (no directamente medibles) se está realizando un esfuerzo importante para conocer la validez y fiabilidad de las baterías de indicadores que se utilizan para medir tales conceptos. En los últimos años, sobretodo en los países con una mayor tradición estadística, empieza a ser una práctica habitual la inclusión de los coeficientes de fiabilidad de las escalas en la documentación de las encuestas (Saris y Andrews, 1991; Saris, y Van Meurs, 1989; Litwin, 95). En este sentido, la ESS proporciona una magnífica oportunidad para conocer los valores de estos coeficientes en nuestro país.

Protocolo para la traducción de los cuestionarios. La traducción del cuestionario, en el caso de España al castellano y al catalán, ha merecido también especial atención por parte del CCT que ha establecido un protocolo específico consistente en:

- Dos equipos de traducción independientes
- Una reunión para contrastar versiones y discutir la mejor alternativa
- Un árbitro para decidir la formulación final
- Documentación y justificación de todas las decisiones tomadas

La complejidad de este protocolo se justifica por la necesidad de asegurar al máximo la equivalencia funcional del redactado de las preguntas en los distintos idiomas y contextos. Por ejemplo, en el caso catalán, un equipo de traducción estaba formado por filólogos traductores profesionales y el otro equipo estaba formado por politólogos con un conocimiento correcto de la lengua pero que no eran expertos traductores.

Ciertamente, las dos versiones resultantes contenían formulaciones no coincidentes.

Ambos equipos realizaron un esfuerzo importante para aunar la corrección lingüística con el uso del lenguaje propio del ámbito político-social y la utilización de un registro lingüístico comprensible para el ciudadano. Así, pues, la experiencia ha justificado sobradamente la existencia de este protocolo.

No se admiten sustituciones. La práctica de sustituir los casos no contactados y los de no respuesta por nuevos casos obtenidos en una segunda (tercera, cuarta, o...) selección tiene consecuencias importantes y peligrosas (Rubin, 1987). Citaremos un par de ellas. La primera es que la probabilidad de inclusión es mucho menor para los casos que han sido seleccionados en la muestra original que para los casos obtenidos en sucesivas

sustituciones. Por lo tanto, ya que los individuos que finalmente forman parte de la muestra no han tenido la misma probabilidad de ser escogidos, deberían ponderarse los casos para poder realizar inferencia, dando menor peso al individuo que tiene mayor probabilidad de ser seleccionado. El problema es que se desconoce esa probabilidad. La segunda consecuencia, muy relacionada con la primera, es el sesgo derivado del abandono de aquellos casos difíciles de contactar o de no respuesta. El abandono de estos casos provoca una sobre representación de las personas de fácil acceso o que no ponen reparos en opinar sobre determinados temas. Por ejemplo es conocido que en el País Vasco existe una elevada tasa de no respuesta global a las encuestas sociales y de no respuesta a preguntas específicas de contenido político aún entre aquellos que aceptan participar. Ello indica que existen individuos reacios a opinar sobre determinados temas y que, probablemente, las opiniones de éstos son distintas de las de aquellos que no han tenido reparos en expresarlas. Si sustituimos los primeros por los segundos, estamos sesgando claramente los resultados.

Tasa de respuesta mínima del 70%. Esta tasa se calcula dividiendo el número de entrevistas completadas entre el total de casos válidos. Idealmente se debería alcanzar una tasa de respuesta cercana al 100% para garantizar que la muestra fuera representativa de la población. Sin embargo, es conocida la dificultad real que existe para conseguir una tasa de respuesta elevada. Por este motivo el CCT se ha planteado para la ESS una tasa de respuesta del 70% creyendo que ésta es más realista. Ya que no se admiten sustituciones, alcanzar esta tasa requerirá un enorme esfuerzo de seguimiento de cada uno de los casos

incluidos en la muestra. De ésta manera se minimiza el sesgo derivado de excluir del estudio las personas de difícil acceso o las que rechazan ser entrevistadas.

Exigencias sobre el trabajo de campo. Los requisitos establecidos por el CCT sobre cómo ha de ser la realización del trabajo de campo y la minuciosa documentación que se exige del mismo están encaminados a la consecución de una elevada tasa de respuesta y a la minimización de las malas prácticas que, de existir, son de consecuencias imprevisibles. El énfasis puesto en la formación y el establecimiento de incentivos especiales para todo el personal involucrado en el trabajo de campo (entrevistadores y coordinadores locales) se justifica por la necesidad de que éste tenga perfectamente claros los objetivos de la encuesta, las razones por las que se exige agotar al máximo las posibilidades de contacto y respuesta y las consecuencias para la calidad de la muestra que se derivarían de su no observancia. Por otro lado, el énfasis puesto en la documentación exhaustiva de todo el proceso se justifica por ser ésta una parte imprescindible en la monitorización que ha de realizar la empresa encuestadora y en el control del trabajo de campo que debe realizar equipo nacional de la ESS. Por lo que respecta a la recogida de información sobre el entorno físico de la vivienda del encuestado, el interés estriba no solamente en posibilitar el control, sino que se trata de una información que puede ser utilizada en estudios posteriores y, en particular, en el estudio de la no respuesta.

Fichero de datos contextuales. El objetivo del fichero de datos contextuales de ámbito nacional es permitir el control de los efectos del contexto en las respuestas individuales

dadas por los ciudadanos y el estudio de la interacción entre individuo y contexto, facilitando así la realización de estudios comparados entre países (King *et al.*, 2003). La ESS, que ha sido diseñada explícitamente para posibilitar la comparación y que tiene en su ánimo impulsar la aplicación de técnicas estadísticas multinivel (Jones *et al.*, 1992; Steenberger y Jones, 2002), ha incorporado en el diseño de la encuesta los elementos necesarios para ello.

LA PRIMERA OLA DE LA ESS EN ESPAÑA⁵

Diseño muestral aplicado

Siguiendo las sugerencias de la empresa encuestadora⁶, se optó por un muestreo estratificado polietápico. Los estratos se obtuvieron de la combinación de dos criterios: uno territorial, las 17 comunidades autónomas, y otro relacionado con el tamaño de hábitat, urbano/rural. La categoría urbana estaba integrada por las capitales de provincia y la categoría rural englobaba los restantes municipios.

La selección de la muestra dentro de cada uno de los 34 estratos resultantes constó de tres etapas:

1. Selección de secciones censales en cada estrato con probabilidad proporcional al número de hogares de cada sección. El número total de secciones seleccionadas fue de 346, 121 de las cuales correspondían a estratos urbanos y 225 a estratos rurales.

⁵ <http://www.upf.es/dcpis/grcp/ess/index.htm>

⁶ En la ola del 2002 fue Demoscopia

2. Selección de un número fijo de hogares dentro de cada sección seleccionada, 12 en cada sección urbana y 10 en cada sección rural.
3. Selección de un individuo en cada hogar. La selección la realizó cada entrevistador utilizando el método del último aniversario⁷.

No es necesario justificar el primer criterio de estratificación. La diversidad de características sociales, políticas y económicas existente entre comunidades autónomas así lo aconseja, siendo además la práctica habitual en la mayoría de los estudios sociales de nuestro país. Pero sí parece conveniente justificar la definición de los dos estratos según tamaño de hábitat. Según la opinión de las empresas encuestadoras, la tasa de respuesta en las grandes ciudades acostumbra a ser un 20% inferior a la del resto de municipios. La estratificación propuesta posibilitaba la selección de una muestra más amplia en los estratos urbanos de tal manera que dichos estratos quedasen adecuadamente representados a pesar de tener una tasa de respuesta inferior a la de los estratos rurales. Podría argumentarse que algunas capitales de provincia (Soria o Ávila, por ejemplo) tienen menos población que otros municipios que no son capitales de provincia (como Hospitalet de Llobregat o Badalona, por citar dos ciudades de Cataluña). Sin embargo, el Instituto Nacional de Estadística, INE, entidad que proporcionó la muestra de domicilios, defiende que una capital, aunque sea pequeña, sigue ejerciendo las funciones de capitalidad, motivo por el cuál sus habitantes son más “urbanos” que los de otras ciudades del mismo tamaño que no son capitales. De ahí que el INE utilice habitualmente este tipo de estratificación para la selección de domicilios. Las probabilidades de

⁷ Este método consiste en preguntar a la persona contactada cuál es el miembro del hogar que ha celebrado su aniversario más recientemente. Esta persona será la escogida para ser incluida en la muestra.

inclusión de las secciones y de los domicilios han sido proporcionadas por el mismo INE, mientras que las probabilidades de inclusión de los individuos han sido calculadas a partir de la información generada por el trabajo de campo.

Con el diseño aplicado, la probabilidad de inclusión de cada sección es directamente proporcional al número de hogares que la integran y, a su vez, cuantos más hogares hay en una sección menor es la probabilidad de cada uno de ellos de formar parte de la muestra. Por lo tanto, las secciones más pobladas tienen mayor probabilidad de ser seleccionadas que las poco pobladas y, a su vez, los domicilios de las secciones más pobladas tienen menor probabilidad de ser seleccionados que los domicilios de las secciones poco pobladas. Como consecuencia, el producto de estas dos probabilidades se mantiene constante, asegurando que los hogares son equiprobables⁸.

No obstante, no hay que olvidar que la ESS es una encuesta de individuos. La selección de los individuos es la última etapa del muestreo y, debido a la composición tan diversa de los hogares españoles, no es posible garantizar la equiprobabilidad en la selección de individuos. Así, en los hogares unipersonales la probabilidad de selección es 1, mientras que en un hogar con cuatro miembros de la población objeto de estudio la probabilidad es 0,25. El método utilizado para la selección del individuo garantiza que todos los integrantes de un mismo hogar tienen la misma probabilidad de ser escogidos, pero no garantiza la equiprobabilidad entre individuos de diferentes domicilios.

⁸ Las muestras obtenidas con este procedimiento se suelen denominar *muestras autoponderadas*.

Puesto que la probabilidad de inclusión no es la misma para todos los individuos, es necesario calcular esta probabilidad e introducirla en el fichero de datos. De esta manera el investigador podrá aplicar las correspondientes ponderaciones al realizar sus estimaciones.

En nuestro país es una práctica habitual el suministrar unos pesos para aplicar ponderaciones en el caso de muestreos estratificados no proporcionales. Por ejemplo la documentación de las encuestas del CIS incluye esa información. Ésta es una práctica correcta y sería suficiente si se tratara de una muestra de hogares (como la Encuesta de Presupuestos Familiares) o se tratara de una muestra de individuos en la que todos los individuos de cada estrato tuvieran las mismas probabilidades de inclusión. Sin embargo, habitualmente no es así. Siempre que se seleccione una muestra de individuos y el método utilizado para ello sea polietápico con selección final de un individuo dentro de cada hogar (como sucede en gran parte de las encuestas de opinión), se debería corregir el sesgo originado por las diferentes probabilidades de inclusión de los individuos (Silva, 2000). Por lo tanto, la ESS al tener en cuenta las probabilidades de inclusión en todas y cada una de las etapas del muestreo aporta la información necesaria y suficiente para realizar inferencias válidas. En ese sentido, la ESS representa un avance significativo respecto de los estándares de calidad habituales.

Tamaño muestral

En el cálculo del tamaño de la muestra se tuvo en cuenta el efecto de diseño esperado y las exigencias del CCT en cuanto al tamaño mínimo efectivo y la tasa de respuesta.

Dos son las fuentes de efecto de diseño en el muestro aplicado en España: Las diferentes probabilidades de selección de los individuos y la existencia de conglomerados. El efecto de diseño total, EFDIS, será el producto de los efectos de diseño correspondientes a cada una de las dos fuentes, que llamaremos EFDIS_P y EFDIS_C, respectivamente.

El efecto de diseño debido a las diferentes probabilidades de inclusión individual se ha estimado, siguiendo la sugerencia de la empresa encuestadora, a partir de los datos de la encuesta Estudio General de Medios. Para esta primera ola, el valor estimado de este efecto ha sido:

$$\text{EFDIS}_P = 1,16$$

Como ya se ha comentado anteriormente, el grado de conglomeración induce una pérdida de eficiencia. En el cálculo del efecto de diseño correspondiente interviene el coeficiente de correlación intra grupos (ρ) que mide el grado de homogeneidad de los grupos. Dado que en el caso español no se disponía de información que permitiera calcularlo, se optó por seguir las recomendaciones del CCT al respecto, tomando $\rho = 0,02$. Así se obtuvo:

$$\text{EFDIS}_C = 1,20$$

El efecto de diseño total estimado, producto de los dos anteriores, resultó ser:

$$\text{EFDIS} = 1,16 \cdot 1,20 = 1,392$$

En consecuencia, para conseguir un *tamaño efectivo* de 1500 era necesario disponer de un *tamaño neto* (número de entrevistas completadas) de 2088, resultado del producto de 1500 por el efecto de diseño estimado más arriba. De esta manera también se satisfacía el

requisito de la CCT de obtener un número de entrevistas completadas igual o superior a 2000.

Tal como se ha dicho anteriormente, conseguir una tasa de respuesta del 70% era uno de los requisitos importantes que se había planteado la ESS. Se trataba de un reto difícil en el caso español. La experiencia de los miembros del equipo local, así como la de los técnicos de la empresa encuestadora, indicaba que era prácticamente imposible de alcanzar. A diferencia de la otros países participantes en la ESS, la cultura de los españoles tiene una tradición de recelo y de escasa participación es este tipo de consultas. El equipo local prefirió adoptar una postura prudente y calcular el tamaño de la muestra partiendo de una tasa de respuesta esperada del 56%, media ponderada de las tasas de respuesta en los estratos rural y urbano obtenida en las encuestas de mayor calidad realizadas en nuestro país. Suponiendo, pues, una tasa de respuesta global del 56%, el tamaño bruto de la muestra requerida era de unos 3700 casos.

Revisión de las hipótesis de partida en el diseño muestral

En el momento de realizar el diseño muestral se presumía que, siendo el INE el organismo que suministraba la muestra de domicilios, el número de direcciones no válidas sería despreciable⁹. Lamentablemente, esta presunción resultó ser errónea y fue la causante de que el número de cuestionarios realmente completados en España no alcanzara finalmente los 2000 exigidos. La explicación se halla en que la lista de domicilios usada por el INE para la extracción de la muestra procedía del censo de 1991.

⁹ Se consideran direcciones no válidas aquellas que no corresponden a domicilios, excluyendo por lo tanto de la muestra a los comercios, oficinas, sedes de instituciones y similares.

En otoño del 2002, el censo de 2001 no estaba suficientemente explotado y la correspondiente lista de domicilios por secciones no se hallaba todavía disponible. El resultado es que aproximadamente un 15% de las direcciones de la muestra resultaron ser no válidas. No obstante, se aprecian diferencias en la calidad de las listas según comunidades autónomas: desde un 5% en el País Vasco o un 7% en la Comunidad de Madrid hasta un 24% en Castilla-León o un 25% en Canarias (Tabla 1).

En el diseño se había previsto una tasa de respuesta media del 56%. Como se puede apreciar en la tabla, esta previsión resultó acertada, siendo muy cercana a la tasa de respuesta realmente obtenida, que ha sido de un 55%. Los datos parecen indicar que no existen diferencias tan importantes como las previstas entre la tasa de respuesta rural y urbana y que, en cambio, sí que existen algunas diferencias en la tasa de respuesta por comunidades autónomas no previstas inicialmente (Tabla 1).

Por que respecta a la estratificación según tamaño de hábitat se tomó como punto de partida que la tasa de respuesta en el estrato rural era un 20% más elevada que en el estrato urbano. En concreto, se había previsto un 50% en el estrato urbano y un 60% en el rural. Los resultados de la encuesta no contradicen esta hipótesis, pero parecen indicar que la diferencia entre los dos estratos es menor que la prevista. Por ejemplo, apenas se aprecian diferencias entre ambas tasas de respuesta en algunas comunidades como Madrid o Murcia. Por lo que respecta a las comunidades autónomas, la estratificación se justificaba por la variabilidad de características políticas sociales y económicas. Un primer indicio de la conveniencia de la estratificación por comunidades lo suministra la

tasa de respuesta diferenciada que se observa. Desde un 35% en el País Vasco hasta un 66% en Asturias (Tabla 1). Queda pendiente un análisis más detallado de estos resultados con la finalidad de ajustar mejor la estratificación más conveniente para utilizar en el futuro.

Tabla 1. Porcentajes de direcciones no válidas y distribución de las validas por comunidades autónomas. Resultados de la ESS 2002.

	No válidas	Válidas		
		No contacto	Rechazo	Respuesta
Andalucía	19	13	28	58
Aragón	14	16	36	48
Asturias	15	13	22	66
Baleares	12	13	22	64
Canarias	25	3	39	58
Cantabria	18	16	33	51
Castilla-La Mancha	22	16	29	56
Castilla-León	24	9	31	60
Cataluña	12	12	34	54
Valencia	8	7	30	62
Extremadura	17	18	27	55
Galicia	15	9	32	57
Madrid	7	10	38	51
Murcia	13	15	36	49
Navarra	11	14	28	58
País Vasco	5	15	49	35
Rioja	14	0	47	53
TOTAL	14	11	33	55

Fuente: Elaboración propia

REFLEXIONES Y PROPUESTAS DE FUTURO

La reflexión sobre los resultados de la implementación de la primera ola de la ESS en España permite discernir cuáles son aquellos aspectos susceptibles de mejora y proporciona elementos para ajustar mejor todos los apartados del diseño y su ejecución en olas posteriores. En esa dirección van algunas de las propuestas de futuro que se apuntan.

Sobre la preparación de la encuesta

Tal como se ha señalado anteriormente, en la determinación del tamaño de la muestra es imprescindible conocer de antemano el efecto derivado del diseño escogido. El desconocimiento del valor del coeficiente de correlación intra para nuestro país llevó al equipo local a la adopción del valor sugerido por el CCT ($\rho = 0,02$). El análisis de los datos españoles de la ESS del 2002 permitirá, sin embargo, estimar un coeficiente de correlación intra para España. El problema es que esta estimación no es simple ya que existe un coeficiente ρ para cada variable. Puesto que para el cálculo del efecto de diseño es imprescindible disponer de un único coeficiente para toda la muestra, debe diseñarse alguna estrategia para conseguirlo. Una de ellas podría ser el cálculo del promedio de los coeficientes de una selección de variables. Siendo ésta una propuesta sensata, deja en el aire cuáles pueden ser los criterios a utilizar en la selección de las variables.

Por otro lado, la existencia de diferencias en el número de miembros mayores de 15 años residentes en cada domicilio provoca que la probabilidad de selección de los individuos varíe de un domicilio a otro. El correspondiente efecto de diseño se ha estimado en la ola del 2002 a partir de la distribución de hogares proporcionada por el Estudio General de Medios. Se trata, por lo tanto, de una estimación obtenida a través de los datos de una encuesta. No obstante, para la próxima ola el INE se ha comprometido a suministrar la información poblacional sobre la composición de los hogares españoles.

El recuento final de las entrevistas completadas arroja una cifra de 1729, inferior al objetivo del CCT que establecía un mínimo de 2000. Una de las causas, además de la tasa de respuesta, es la existencia no esperada de aproximadamente un 15% de direcciones no válidas. El censo del 2002, que se utilizará como marco de muestreo en la próxima ola, debería proporcionar una muestra actualizada, con un porcentaje substancialmente inferior de direcciones no válidas.

El CCT de la ESS indicaba en sus instrucciones la conveniencia de que el coordinador local participara personalmente en el proceso de formación del personal de campo, pero la empresa encuestadora se mostró reticente considerando que, por tratarse de un aspecto práctico, este tema recaía exclusivamente en su ámbito de actuación. Así, la transmisión de la filosofía del estudio, de los objetivos concretos que se habían marcado, de los estándares de calidad, así como de las instrucciones precisas sobre cómo llevar a cabo el trabajo de campo, quedó únicamente en manos de la empresa encuestadora. A la vista de cómo se ha desarrollado el trabajo de campo y de los problemas que han ido surgiendo,

es evidente que esta decisión deberá ser reconsiderada en la segunda ola. La impresión final que el equipo local ha obtenido del proceso del trabajo de campo es que las empresas encuestadoras no son conscientes de las repercusiones que para la representatividad de la muestra tienen algunas de sus prácticas habituales. Hay muchos aspectos probabilísticos y estadísticos de la teoría del muestreo que son infravalorados por los técnicos de estas empresas. Y si eso es así, las empresas no están en condiciones de formar adecuadamente por sí solas a los entrevistadores y transmitirles las razones de las especificaciones establecidas para el trabajo de campo.

Sobre la realización del trabajo de campo

La tasa de respuesta alcanzada ha sido de un 55%. A pesar de distar 15 puntos de la tasa que el CCT había establecido como objetivo (70%), el nivel de respuesta en la encuesta de la ESS representa una mejora sustancial respecto del de otras encuestas en España¹⁰. Lamentablemente, no se puede realizar la comparación de la tasa de respuesta de la ESS con la de otras encuestas de contenido político y social. El motivo es no se dispone de información para calcular dichas tasas al ser una práctica común la aplicación de sustituciones de los casos de no contacto y no respuesta sin que se documente la magnitud de tales sustituciones.

Es precisamente la documentación exhaustiva de todo el proceso exigida por el CCT la que proporciona garantías de la calidad de las prácticas realizadas y permite un control efectivo del trabajo de campo. Cabe señalar que gracias a esa documentación el equipo

local pudo supervisar el trabajo realizado por la empresa encuestadora. Por ejemplo, cuando la empresa encuestadora daba éste por terminado, con una tasa de respuesta hasta el momento del 48%, se procedió a la revisión de todos los casos de no contacto y no respuesta. Ello permitió agotar las posibilidades de respuesta adjudicando a un nuevo entrevistador todos los casos tipificados como “negativas suaves”. La estrategia seguida elevó la tasa de respuesta definitiva hasta el 55%.

Aunque inicialmente se había considerado la posibilidad de utilizar CAPI¹¹ para la anotación de las respuestas en la entrevista cara a cara, finalmente se optó por el uso de papel y lápiz debido a problemas logísticos. No obstante, para la ola del 2004 se ha puesto como exigencia a las empresas que opten a la adjudicación del trabajo de campo la utilización de esa modalidad. CAPI evitará errores en las anotaciones y permitirá al mismo tiempo un seguimiento actualizado de todo el proceso de recogida de datos, facilitando el control y la intervención de los supervisores.

Finalmente, se debe reconocer que no se ha conseguido completamente el objetivo inicial de crear un único fichero de datos contextuales de los países participantes en la ESS que incorporara el contexto económico, político y social así como los sucesos puntuales acaecidos durante el período de administración de la encuesta. Los equipos nacionales han recogido y estructurado la información, pero ésta no ha podido ser materializada en un único fichero de datos. La dificultad de convertir una información textual en un

¹⁰ Familitur, un panel de turismo del Ministerio de Economía consigue un porcentaje de respuesta del 47% y la encuesta financiera a las familias españolas realizada por el Banco de España alcanza un porcentaje de respuesta del 35%.

¹¹ Computer Assisted Personal Interview

fichero numérico plantea evidentes problemas de codificación y homogeneización que no pueden ser resueltos precipitadamente. El trabajo realizado en la primera ola permitirá disponer de un material con el que ensayar clasificaciones y discutir cómo confeccionar este fichero. Se espera poder disponer de un único fichero contextual integrado a partir de la segunda ola. Por el momento, y para los datos de la primera encuesta, el investigador deberá conformarse con consultar esa información en la página:

www.scp.nl/users/stoop/ess_events/events_overview.htm.

VALORACIONES FINALES

A la espera de la valoración final sobre la primera ola de la ESS que debe hacer el CCT, y sin conocer todavía los resultados del proceso en los demás países participantes, la impresión del equipo español es que, con algunas limitaciones, la encuesta se ha realizado satisfactoriamente.

Destacaríamos tres aspectos del trabajo en equipo que, si bien han representado una cierta dificultad inicial, al mismo tiempo han aportado un conocimiento y una experiencia que sus integrantes valoran positivamente. Se trata de las interacciones entre el equipo local, la empresa encuestadora y el CCT.

En la interacción entre universidad y empresa, en un inicio se produjo un cierto “choque de culturas profesionales” entre el mundo académico y el mundo de las empresas de

estudios de mercado. La desconfianza inicial se tradujo en una falta de comunicación entre los técnicos de uno y otro ámbito que produjo muchos errores y malentendidos. A través de reuniones, discusiones y mucha paciencia se consiguió generar poco a poco la confianza necesaria entre los dos colectivos para implementar el proyecto correctamente.

Dentro del mundo académico se produjeron, también inicialmente, algunas duplicidades e interferencias. Ha sido, pues, necesario un cierto rodaje hasta conseguir definir una manera eficiente de trabajar, en la que el coordinador delega responsabilidades, cada individuo ejerce una determinada tarea y la responsabilidad es asumida colectivamente.

En cuanto a la interacción entre el CCT y el equipo nacional, cabe decir que la facilidad de comunicación que suministra el correo electrónico y la existencia de tantas comisiones que conectaban directamente con los equipos locales, provocó una avalancha de mensajes con instrucciones que en ciertos momentos llegó a colapsar el trabajo del coordinador nacional. Entendemos que éstas son disfunciones propias de un primer rodaje y que es de esperar que desaparezcan en ediciones posteriores cuando gran parte de los protocolos estén ya definidos. Sin embargo, la experiencia ha sido ciertamente positiva. No tan sólo se ha adquirido conocimiento sobre aspectos metodológicos de suma importancia, debido a la colaboración con los expertos que forman el CCT y sus grupos de asesores, sino que el equipo español valora especialmente la forma de relación establecida con todos ellos. A pesar de ser el CCT quien ha establecido los criterios y ha decidido sobre las propuestas, se considera que el equipo local ha disfrutado de autonomía suficiente. En efecto, las mejoras introducidas a lo largo del proceso de implementación de la Encuesta

no han sido impuestas sino que han surgido como consecuencia de las “preguntas inteligentes” del CCT o de las “sugerencias” de sus expertos que obligaban al equipo español a replantearse algunos aspectos.

Finalmente, se debe insistir en que todo el esfuerzo realizado, económico y personal, habrá valido la pena si los datos obtenidos aportan conocimiento. Esto no se sabrá hasta que dichos datos sean utilizados por los investigadores sociales. Por el momento podemos avanzar que ha sido aprobado por parte del Ministerio de Ciencia y Tecnología un proyecto de investigación que lleva por título *Ciudadanos y Democracia en Europa: Análisis comparado de los datos de la primera ola de la Encuesta Social Europea* y que, bajo la dirección del coordinador nacional Mariano Torcal, agrupa una veintena de investigadores de centros de investigación de todo el país. Probablemente los participantes en este proyecto sean los primeros en empezar a trabajar con esos datos. Pero no deberían ser los únicos. Desde aquí se insta a toda la comunidad científica a hacer uso de la ESS dentro de sus respectivas áreas de trabajo y a participar en la definición de los contenidos variables de las sucesivas ediciones de la misma. A tal efecto, cabe recordar que el 16 de mayo de 2003 finalizó la convocatoria de propuestas de contenido para los módulos variables a incluir en la segunda ola, pero que se abrirán nuevas convocatorias en un futuro.

ANEXO: EFECTO DE DISEÑO, TAMAÑO EFECTIVO Y MARGEN DE ERROR

El muestreo por etapas (ME) es, en general, menos eficiente que el muestreo aleatorio simple (MAS). Sea π un parámetro proporción y p su estimador. El efecto de diseño, que mide la reducción en la precisión del estimador, se calcula como el cociente entre la varianza del estimador obtenido en la muestra por etapas y la del estimador obtenido en la muestra aleatoria simple de igual tamaño:

$$\text{EFDIS} = \frac{\text{Var}_{ME}(p)}{\text{Var}_{MAS}(p)}$$

El valor estimado para la ESS 2002 en España fue 1,39. Tomando la raíz cuadrada del efecto de la expresión anterior se obtiene:

$$\frac{EE_{ME}(p)}{EE_{MAS}(p)} = \sqrt{1,39} = 1,18$$

De ahí que el error estándar del estimador derivado del muestreo por etapas sea igual al correspondiente error estándar del muestreo aleatorio simple multiplicado por 1,18.

$$EE_{ME}(p) = 1,18 * EE_{MAS}(p)$$

Cálculo del margen de error para la muestra de la ESS en España

Sabiendo que el margen de error máximo en la estimación de una proporción π para un nivel de confianza del 95% en el caso de un muestreo aleatorio simple de 1729 entrevistas es aproximadamente:

$$z \cdot \sqrt{\frac{\pi(1-\pi)}{n}} \cong \sqrt{\frac{1}{n}} = \frac{1}{\sqrt{1729}} = 0,0240 \quad (1)$$

Podemos obtener el margen de error de la estimación correspondiente al diseño aplicado por dos vías alternativas:

a) Multiplicando la expresión (1) por 1,18:

$$z \cdot EE_{ME}(p) = z \cdot EE_{MAS}(p) \cdot 1,18 = 0,0240 \cdot 1,18 = 0,0284$$

b) Utilizando el tamaño efectivo, $nef = 1729/\text{EFDIS} = 1729/1,392 = 1242$, en vez del número de entrevistas:

$$z \cdot \sqrt{\frac{\pi(1-\pi)}{nef}} \cong \sqrt{\frac{1}{nef}} = \frac{1}{\sqrt{1242}} = 0,0284$$

Por lo tanto, el margen de error máximo para la inferencia sobre proporciones es 2,84%.

Referencias

- Dustmann, C. y Preston, I. 2000. "Attitudes to ethnic minorities, ethnic context and location decisions", *Economic Journal*, 111.
- Jones, K., Johnston, R.J. y Pattie, C.J. 1992. "People, Places and Regions: Exploring the Use of Multi-Level Modeling in the Analysis of Electoral Data", *British Journal of Political Science*, 22: 343-380.
- King, G.; Murray, C.J.L.; Salomon, J.A. y Tandom, A. 2003. *Enhancing the Validity and Cross-cultural Comparability of Survey Research*. Documento de trabajo publicado en la web: <http://gking.harvard.edu/>
- Kish, L. 1965. *Survey Sampling*. New York: Wiley.
- Litwin, M.S. 1995. *How to measure survey reliability and validity*. Thousand Oaks: Sage.
- Rubin, D.B. 1987. *Multiple Imputation For Nonresponse In Surveys*. New York: Wiley.
- Saris, W.E. y Andrews, F.M. 1991. "Evaluation of measurement instruments using a structural modeling approach", en Biemer, P. P., Groves, R. M., Lyberg, L. E., Mathowetz, N. y Sudman S., eds. *Measurement errors in surveys*. New York: John Wiley & Sons, 575-599.
- Saris, W.E. y Van Meurs A. eds. 1989. *Evaluation of measurement instruments by meta-analysis of multitrait multimethod studies*. Amsterdam: North-Holland.
- Schwartz, S., Melek, G., Lehmann, A., Burgess, S., Harris, M. y Owens, V. 2001. "Extending the cross-cultural validity on the theory of basic human values with a different method of measurement", *Journal of Cross-cultural Psychology*, 32: 519-542.
- Silva, L.C. 2000. *Diseño razonado de muestras y captación de datos para la investigación sanitaria*. Madrid: Díaz de Santos.
- Steenberger, M.R. y Jones, B.S. 2002. "Modeling Multilevel Data Structures", *American Journal of Political Science*, 46: 218-237.
- Van Deth, J. 2000. "Interesting but irrelevant: Social capital and the saliency of politics in Western Europe", *European Journal of Political Research*, 37: 115-147.