

Improving social measurements: the experience of the ESS first rounds¹ in Spain

Anna Cuxart² (anna.cuxart@upf.edu)

Departament d'Economia i Empresa
Universitat Pompeu Fabra

Clara Riba (clara.riba@upf.edu)

Departament de Ciències Polítiques i Socials
Universitat Pompeu Fabra

Abstract

With the two aims of monitoring social change and improving social measurement, the European Social Survey is now closing its third round. This paper shows how the accumulated experience of the two first rounds has been used to validate the questionnaire, better adapt the sampling design to the country characteristics and efficiently commit fieldwork in Spain. For example, the dynamic character of the population nowadays makes necessary to estimate design effects at each round from the data of the previous round. The paper also demonstrates how, starting with a response rate of 52% at first round, a 66% response rate is achieved at the third round thanks to an extensive quality control conducted by the polling agency and the ESS national team and based on a detailed analysis of the non-response cases and the incidences reported by the interviewed in the contact form.

Keywords: European social survey, fieldwork, response rate, sampling design

Palabras clave: encuesta social europea, trabajo de campo, tasa de respuesta, diseño muestral

Journal of Economic Literature Classification: C42

¹ Paper supported by research grants SEC2002-03364 and SEJ2006-10974. The Spanish "Instituto Nacional de Estadística" (INE) provided the samples of the three rounds.

² The authors are members of the Spanish national team of the ESS.

Resumen

En esta comunicación se muestra cómo la experiencia acumulada de las primeras olas de una encuesta periódica permite afinar el diseño muestral de las siguientes y cómo el análisis del formulario de contacto facilita a su vez la mejora de los procesos de trabajo de campo en las olas sucesivas. El carácter dinámico de las poblaciones plantea la necesidad de actualización continua de las estimaciones de los efectos de diseño, usando para ello los datos de las olas precedentes. Por otra parte, el análisis detallado de las tasas de respuesta y de las incidencias consignadas por los encuestadores en el formulario de contacto, permite diseñar estrategias para la optimización de las actividades de trabajo de campo.

1. Introducción

El objetivo principal de la *Encuesta Social Europea* (ESE en sus siglas en castellano) es el estudio comparado y longitudinal de las actitudes, valores y comportamientos de los ciudadanos europeos en referencia a los ámbitos económico, social y político. La ESE, de carácter bianual y realizada en casi una treintena de países, nació con un triple objetivo. En primer lugar la ESE pretende crear una gran base de datos transnacional y longitudinal que incluya datos individuales y contextuales que permita mostrar y analizar los cambios en las actitudes, valores y comportamientos. En segundo lugar la ESE, con una clara orientación académica, persigue la difusión de estándares de rigor que garanticen la calidad de los datos obtenidos, estableciendo para ello unos estrictos requisitos que deben satisfacer tanto el diseño muestral como la ejecución y seguimiento del trabajo de campo en cada uno de los países participantes. Finalmente, la ESE se propone proporcionar las bases para la generación de indicadores sociales y actitudinales que no solamente sean reconocidos en los medios académicos sino que, al igual que ocurre con otros indicadores del ámbito de la economía, gocen de amplia aceptación en ámbitos políticos y sociales y puedan ser utilizados en el análisis comparado y en la toma de decisiones³.

La ESE está promovida por la European Science Foundation y cuenta con el apoyo de distintas instituciones públicas europeas. La principal institución financiadora del proyecto es la Comisión Europea, que a través de sus quinto y sexto Programas Marco financia la coordinación de todo el proyecto y, en cada país participante, los respectivos gobiernos aportan una financiación específica. Recientemente, acaba de ser firmado un *Memorandum of Understanding* por el que 18 de los países participantes han adquirido el compromiso de seguir financiado el proyecto a largo plazo. Hasta la fecha se han realizado tres olas de la encuesta (en 2002, 2004 y 2006), siendo uno de los objetivos específicos de la tercera la investigación de métodos de obtención de datos actitudinales.

³ La ESE es, sin duda, uno de los proyectos más ambiciosos de las Ciencias Sociales europeas (www.europeansocialsurvey.org). Cabe destacar que en el 2005 ha sido merecedor del *Premio Descartes de la Ciencia* que otorga la Comisión Europea, siendo ésta la primera vez que un proyecto en Ciencias Sociales recibe este prestigioso premio. Acceso universal y gratuito a los datos de la ESE ola 2002 y ola 2004 en <http://ess.nsd.uib.no>

A lo largo de las tres ediciones, han participado un total de 32 países, de los cuales, sin embargo, solo 17 países, entre ellos España, han participado en las tres (véase tabla 1). Casi todos los participantes forman parte de la Unión Europea de 27 países, excepto Islandia, Israel, Noruega, Rusia, Suiza, Turquía y Ucrania. Por otra parte, sólo dos miembros de la UE no han participado en ninguna de las olas: Letonia y Malta.

Tabla 1. Países participantes en las tres olas de la ESE

País	Ola 1	Ola 2	Ola 3	País	Ola 1	Ola 2	Ola 3
Alemania	•	•	•	Israel	•		
Austria	•	•	•	Italia	•	•	
Bélgica	•	•	•	Lituania			•
Bulgaria			•	Luxemburgo	•	•	
Chipre			•	Noruega	•	•	•
Dinamarca	•	•	•	Países Bajos	•	•	•
Eslovaquia		•	•	Polonia	•	•	•
Eslovenia	•	•	•	Portugal	•	•	•
España	•	•	•	Reino Unido	•	•	•
Estonia		•	•	República Checa	•	•	
Finlandia	•	•	•	Rumanía			•
Francia	•	•	•	Rusia			•
Grecia	•	•		Suecia	•	•	•
Hungría	•	•	•	Suiza	•	•	•
Irlanda	•	•	•	Turquía		•	•
Islandia		•		Ucrania		•	•

La coordinación de todo el proceso de diseño y ejecución de la ESE recae en un equipo de expertos (Central Coordinating Team, CCT) procedentes de centros de investigación europeos y de una serie de grupos que le asesoran en aspectos específicos como son la traducción del cuestionario o el proceso de muestreo. Además, en cada uno de los países participantes existe un Coordinador Nacional con su correspondiente equipo que se encarga de aplicar las directrices acordadas por el CCT. Finalmente, 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 la tabla 2 se detallan los requisitos técnicos de la ESE establecidos por el CCT para todas las olas de la encuesta. El cumplimiento de esos requisitos es una garantía de calidad de los datos (Riba y Cuxart, 2003).

Tabla 2. Requisitos técnicos de la ESE

Población	<ul style="list-style-type: none">• Todos los residentes en domicilios particulares, de edad igual o superior a los quince años, sea cual sea su nacionalidad, ciudadanía o lengua.
Tipo de muestreo	<ul style="list-style-type: none">• Muestra totalmente probabilística en todas sus etapas de selección, con indicación de las probabilidades de selección de cada una de las unidades muestrales.• No se admite ni la aplicación de cuotas ni la sustitución de casos.
Tasa de respuesta	<ul style="list-style-type: none">• Objetivo: mínimo del 70%.
Tamaño muestral	<ul style="list-style-type: none">• Depende de cada diseño. En todo caso ha de ser suficiente para garantizar un mínimo de 2.000 entrevistas completadas y un tamaño muestral efectivo de 1.500⁴.
Trabajo de campo	<ul style="list-style-type: none">• Redacción de los cuestionarios en todas las lenguas habladas por un mínimo de un 5% de la población.• Entrevista cara a cara.• Período de realización entre uno y cuatro meses.• El primer contacto debe ser personal cara a cara.• Mínimo de cuatro intentos de contacto antes de abandonar una unidad de muestreo por improductiva, de los cuales uno debe ser en horas vespertinas y otro en fin de semana.• Número máximo de entrevistas asignadas a cada entrevistador igual a 48.• Formación específica de entrevistadores con participación del coordinador nacional.• Documentación por parte del entrevistador de todas las incidencias del trabajo de campo en un formulario de contacto.• Compensaciones materiales a los entrevistados.
Supervisión	<ul style="list-style-type: none">• Control por parte de la empresa de campo del trabajo de los encuestadores, con supervisión de un 5% de las entrevistas, un 10% de los rechazos y un 10% de los casos no contactados.• Realización de informes quincenales por parte de la empresa que realiza el trabajo de campo, con inclusión de las tasas de respuesta hasta el momento y desglose de negativas claras y negativas suaves.• Asignación de los casos correspondientes a negativas suaves a un segundo entrevistador que debe reiniciar el proceso.• Seguimiento continuo del trabajo de campo por parte del equipo nacional a través de la supervisión de los informes y de los ficheros de datos provisionales suministrados por la empresa encuestadora y envío de resultados al CCT.

La Encuesta Social Europea, además de establecer los ambiciosos estándares de calidad detallados en la tabla 2, introduce varias innovaciones destacables. La primera se refiere al cuestionario. En efecto, la encuesta dispone de un cuestionario principal y un cuestionario suplementario. El cuestionario principal, a su vez, está formado por un módulo central que se repite en cada ola y por dos o tres módulos rotativos destinados a tratar temas específicos. La temática de los módulos específicos se decide por concurso entre las propuestas de diferentes grupos de investigadores. Los temas de los módulos

⁴ El tamaño efectivo de una muestra es el tamaño de una muestra aleatoria simple con la cual se obtendría el mismo grado de precisión.

rotativos han sido hasta el momento: *Ciudadanos y Democracia y Inmigración* en la primera ola; *Familia, Trabajo y Bienestar, Opiniones sobre salud y cuidados y Moral y Economía en Europa: el mercado y la ciudadanía* en la segunda ola; *Bienestar personal y social: creando indicadores y La organización de nuestra vida* en la tercera ola. En la cuarta ola en preparación los módulos seleccionados son: *Experiencias y Expresiones ante el envejecimiento y Actitudes sobre el bienestar en una Europa en proceso de cambio*. A su vez, el cuestionario suplementario contiene la escala de valores de Schwartz (Schwartz et al., 2001) y una colección de preguntas como principal objetivo es evaluar la fiabilidad y validez de algunos ítems del cuestionario principal.

La segunda innovación se refiere al formulario de contacto. Se trata éste de un instrumento en el que los encuestadores recogen de manera exhaustiva todas las incidencias del trabajo de campo así como información sobre el entorno físico y la vivienda del individuo seleccionado, tanto si es finalmente entrevistado como si no lo es, así como una estimación del sexo y la edad de los individuos contactados y no entrevistados. La tercera innovación se refiere a la documentación que acompaña a la base de datos, documentación que incluye un fichero de “eventos” importantes de carácter político, social o económico acaecidos durante el período de realización del trabajo de campo en cada uno de los países y que puedan afectar a las respuestas de los entrevistados. Finalmente, otro elemento diferenciador consiste en el suministro por parte de la ESE de las probabilidades de selección de cada una de las unidades muestrales. Esta información es importante para corregir los posibles sesgos derivados de la existencia de distintas probabilidades de selección de los individuos dentro de un mismo país.

La presente comunicación consta de tres partes, que siguen a esta introducción y que se refieren a las tres etapas principales que componen el proceso de realización de una encuesta, a saber, la elaboración del cuestionario, el diseño muestral y el trabajo de campo. En cada una de ellas se expone cómo la experiencia obtenida en las olas anteriores permite mejorar la encuesta en las subsiguientes olas. La comunicación concluye con unas consideraciones finales.

2. El cuestionario

La redacción de las preguntas del cuestionario de la ESE es el fruto de varios años de trabajo metodológico de un Comité Científico Asesor, que es interdisciplinario e internacional en su composición. La participación de académicos de reconocido prestigio en ese comité posibilita la utilización por parte de la ESE de preguntas sólidamente contrastadas en investigaciones anteriores, lo que da garantías de la calidad del cuestionario. Por otra parte, la estrategia de abrir un concurso público para la selección de las temáticas de los módulos rotatorios y la supervisión del redactado de las preguntas propuestas para estos módulos, posibilitan la inclusión en el cuestionario de preguntas que están siendo utilizadas de manera experimental por grupos de investigadores que trabajan en temas emergentes. Así, estas nuevas líneas de investigación se benefician tanto de la *expertise* de la ESE como de las oportunidades de difusión que este proyecto les facilita.

La calidad inicial del cuestionario está sometida a un continuo proceso de mejora gracias a las posibilidades de contraste y validación de las preguntas que ofrece el proyecto. Así el cuestionario suplementario, que incluye preguntas sobre aspectos ya preguntados en el cuestionario principal pero usando otro formato, permite aplicar técnicas estadísticas avanzadas, como MTMM (Saris y Munnich, 1995) o Split-Ballot MTMM (Saris *et al*, 2004), para evaluar la calidad de los instrumentos de medida. Una consecuencia de la investigación metodológica realizada hasta el momento con los datos aportados por la ESE ha sido la creación de un programa semiautomático de predicción de la validez y fiabilidad de las preguntas de encuesta (Saris y Gallhofer, 2007). Se trata de una herramienta de gran utilidad que puede aumentar extraordinariamente la calidad de la investigación social puesto que, no sólo permite evaluar a priori la calidad de las preguntas de un cuestionario en su fase de diseño, sino que permite corregir a posteriori los efectos de los errores de medida.

Otro elemento de calidad del cuestionario lo constituye el rigor usado en su traducción. En efecto, la ESE, al tratarse de una encuesta internacional con el objetivo explícito de facilitar el análisis comparado, puso ya desde el principio mucho énfasis en la elaboración de un protocolo de traducción. El cuestionario, inicialmente redactado en inglés, debe ser traducido a las lenguas oficiales de los países participantes y, también, a

todas las lenguas habladas como mínimo por un 5% de la población objeto de estudio. Los esfuerzos se dirigen no únicamente a asegurar la correcta traducción semántica de las preguntas sino también a asegurar la equivalencia funcional de las mismas. También en este aspecto, la experiencia adquirida ha permitido mejorar los redactados. Las diferencias observadas en la primera ola en las respuestas dadas a determinadas preguntas evidenciaron diferentes interpretaciones a las mismas preguntas en distintos ámbitos geográficos. Fruto de ello, el CCT recomendó a algunos países la revisión del redactado de esas preguntas en la segunda ola. Asimismo, también decidió eliminar de la encuesta otras preguntas que se prestaban a interpretaciones dudosas.

Otro tipo de mejoras introducidas en el cuestionario hacen referencia a las estrategias de obtención de la información necesaria para la investigación en ciencias sociales. Por ejemplo, es bien sabida la dificultad de obtener la clase social de un individuo. Ésta se derivaba en las primeras olas del nivel de ingresos del cabeza de familia o en su defecto del de su pareja. El análisis de los datos de esas dos olas mostró la imposibilidad de derivar la clase social de un buen número de encuestados debido a que ni el encuestado ni su pareja aportaban ingresos al hogar. Esta situación afectaba principalmente a muchas mujeres, hecho que imposibilitaba la utilización de la variable clase social en la mayoría de estudios. Pues bien, entre las mejoras que han surgido del análisis de los resultados de las primeras olas se encuentra una propuesta de reformulación de las preguntas que permiten estimar la clase social del entrevistado. Esta propuesta, realizada por el equipo español muy activo en ese tema, ya ha sido aceptada por el CCT y va a ser implementada en la cuarta ola. La opción que se ha tomado es la de preguntar en primer lugar quien es la persona que aporta más ingresos a la economía familiar y cuál es la ocupación del encuestado (o en su defecto de su pareja). En el caso que ni el encuestado ni su pareja sean quienes aportan el ingreso principal, entonces se propone investigar la clase social del encuestado a través de una batería de preguntas sobre la persona que más ingresos aporta (o aportaba) a la economía familiar. Esta estrategia, que combina información individual con información del hogar, permite obtener la clase social de todos los encuestados, completando con información del hogar aquellos casos en los que la información individual no es suficiente.

3. El diseño muestral en España

El segundo pilar de la calidad de una encuesta es el diseño muestral. Dada la heterogeneidad de contextos y disponibilidades de los diferentes países, la ESE optó por establecer unos estándares de calidad, recomendar a cada equipo nacional su adaptación a las características del país y supervisar las propuestas de diseños antes de su aplicación.

La tabla 3 presenta una comparación de las características del muestreo de la ESE en cada una de las tres olas en España.

Como se puede observar, aunque manteniendo el esquema inicial, el diseño ha ido evolucionando. En cuanto a la población, debe mencionarse el aumento de la cobertura a partir de la segunda ola con la incorporación de las ciudades autónomas de Ceuta y Melilla. La práctica habitual de las empresas de estudios de mercado en España excluía sistemáticamente dichos territorios del marco de muestreo con el argumento de que los elevados costos y el bajo porcentaje de población española que en ellos residía (un 0,3%) no justificaban tal esfuerzo. Además, predecían una tasa de respuesta bastante inferior a la esperada en la península. En la segunda ola, el equipo español forzó a la empresa encuestadora a incluir Ceuta y Melilla en el marco de muestreo, considerando al conjunto de las dos ciudades autónomas como un estrato adicional a añadir al de las 17 Comunidades Autónomas. El análisis de los datos de la segunda ola ha desmentido los supuestos iniciales sobre la dificultad del trabajo de campo en esas dos ciudades y sobre la tasa de respuesta. En efecto, la tasa de respuesta en Ceuta y Melilla fue del 71%, cifra superior que la del conjunto de la encuesta (55%) y la más elevada de todas las Comunidades Autónomas, revelando un interés por participar superior al que tienen el resto de los españoles.

Otro cambio importante entre la primera y las siguientes olas reside en la reducción del número de etapas. En efecto, en la primera ola se utilizó como marco de muestreo el registro de domicilios del INE, lo que conllevó un diseño muestral en tres etapas. El elevado número de direcciones no válidas (un 13%) llevó al equipo a reconsiderar la conveniencia de seguir trabajando con domicilios. El propio INE recomendó la utilización como marco de muestreo del registro de individuos que, aunque inicialmente procede del censo como el de viviendas, es actualizado anualmente a través de los datos

de los padrones municipales⁵. De esta manera, las tres etapas consistentes en selección de una sección, de un domicilio dentro de la sección y de un individuo dentro del domicilio, se reducían a dos. Además, la extracción de una muestra de individuos presentaba la ventaja adicional de la eliminación del sesgo de selección que podía ser introducido en la etapa final de selección aleatoria de un individuo dentro de cada domicilio.

Por lo que respecta a la estratificación según tamaño de municipio, también la experiencia ha aconsejado la modificación progresiva de los criterios aplicados. Así, en la primera ola se aplicó uno de los criterios más generalizados como es el de distinguir entre capitales de provincia y resto de municipios. Sin embargo, la existencia de ciudades en España más pobladas que muchas de las capitales de provincia, llevo a la inclusión de un nuevo estrato en la segunda ola, el correspondiente a poblaciones de más de 100.000 habitantes que no son capitales de provincia. Los datos mostraron que, realmente, es en esas grandes ciudades donde se da una menor tasa de respuesta⁶. Finalmente, en la tercera ola se optó por un refinamiento de la estratificación, con el establecimiento de cuatro tramos de tamaño de municipio. Por otra parte, como resultado de la baja tasa de respuesta alcanzada en el País Vasco en la primera ola (35%), se optó en la segunda por un incremento del 50% del tamaño de la muestra correspondiente a esta comunidad. En la tercera ola, sin embargo, dicha sobre-representación sólo se aplicó en los dos tramos superiores de tamaño de municipio, puesto que la participación en pequeñas ciudades y zonas rurales del País Vasco no difería significativamente de la de las demás comunidades autónomas.

⁵ El INE dispone de una muestra marco de unas 3.500 secciones (seleccionadas entre las aproximadamente 33.000 secciones censales existentes) que es la que utiliza habitualmente como marco de muestro para la selección de las muestras de las encuestas que forman parte de la estadística oficial, como la Encuesta de Población Activa (EPA) o la Encuesta continua de Presupuestos Familiares (EPF). Con el fin de disponer de datos actualizados para esas encuestas, el INE actualiza anualmente el registro de domicilios de las secciones que forman parte de la muestra marco, usando para ello el padrón continuo de individuos. Esa muestra marco, sin embargo, no pudo ser usada para la extracción de la muestra de la ESE debido a que el INE consideró que a los 10 años de su extracción la muestra marco ya estaba agotada. En consecuencia realizó la extracción de la muestra de viviendas usando como marco el conjunto de todas las secciones censales, muchas de las cuales contenían información no actualizada desde 1991, con el consiguiente resultado de la presencia de un elevado número de direcciones no válidas en la muestra resultante.

⁶ En el diseño de la primera ola se estimó que la tasa de respuesta sería un 10% mayor en las zonas rurales, motivo por el cual el número de casos seleccionados en las secciones urbanas fue mayor que el las zonas rurales (12 y 10 respectivamente) con el objetivo de que el número de respuestas obtenidas fuera finalmente el mismo.

Tabla 3. Muestreo de las tres olas en España

	Primera ola (2002)	Segunda ola (2004)	Tercera ola (2006)
Cobertura de la población	Todas las personas de 15 o más años residentes en domicilios particulares, excluyendo Ceuta y Melilla.	Todas las personas de 15 o más años residentes en domicilios particulares.	Todas las personas de 15 o más años residentes en domicilios particulares.
Marco de muestreo	Registro de domicilios del INE, procedente de la actualización parcial del censo del 1991.	Padrón continuo de individuos del INE, actualizado en marzo de 2004.	Padrón continuo de individuos del INE, actualizado en abril de 2006.
Diseño muestral	<p>Muestra estratificada de tres etapas</p> <p>Los estratos se obtienen del cruce de CCAA (17) con dos tamaños de municipio que son:</p> <ul style="list-style-type: none"> • Capitales de provincia • Resto de municipios <p>En los 34 estratos resultantes, selección en tres etapas consistentes en:</p> <ul style="list-style-type: none"> • Primera etapa: selección de 346 UPMs (secciones censales) proporcionalmente a la población. • Segunda etapa: Selección de 10 o 12 domicilios en cada UPM seleccionada en la primera etapa (12 en el primer grupo de municipios y 10 en el segundo). • Selección aleatoria de un individuo en cada domicilio seleccionado en la segunda etapa. 	<p>Muestra estratificada de dos etapas</p> <p>Los estratos se obtienen del cruce de CCAA (18) con tres tamaños de municipio que son:</p> <ul style="list-style-type: none"> • Capitales de provincia y de CCAA • Ciudades de más de 100.000 habitantes no incluidas en el grupo anterior • Resto de municipios <p>En los 43 estratos resultantes (no vacíos), selección en dos etapas consistentes en:</p> <ul style="list-style-type: none"> • Primera etapa: selección de 503 UPMs (secciones censales) proporcionalmente a la población. • Segunda etapa: Selección aleatoria de 6 o 7 individuos en cada UPM seleccionada en la primera etapa (7 en los dos primeros grupos de municipios y 6 en el tercero) <p>Sobrerepresentación del 50% en el País Vasco.</p>	<p>Muestra estratificada de dos etapas</p> <p>Los estratos se obtienen del cruce de CCAA (18) con cuatro tamaños de municipio que son:</p> <ul style="list-style-type: none"> • Ciudades de más de 100.000 habitantes • Ciudades entre 50.001 y 100.000 habitantes y capitales de provincia con menos de 100.000 h. • Municipios entre 10.001 y 50.000 habitantes • Municipios de hasta 10.000 habitantes <p>En los 66 estratos (no vacíos):</p> <ul style="list-style-type: none"> • Primera etapa: selección de 504 UPMs (secciones censales) proporcionalmente a la población. • Segunda etapa: Selección aleatoria de 6 o 7 individuos en cada UPM seleccionada en la primera etapa (7 en los dos primeros grupos de municipios y 6 en el resto) <p>Sobrerepresentación del 50% en los dos primeros grupos de municipios del País Vasco.</p>
Efectos de diseño	$ED = (ED_c) \cdot (ED_p) = 1,20 \cdot 1,16 = 1,39$ ED_c no tomado en consideración $ED_c = 1 + (11-1) \cdot 0,02 = 1,20$ $ED_p = 1,16$	$ED = (ED_c) \cdot (ED_e) = 0,956 \cdot 1,284 = 1,225$ $ED_e = 0,956$ $ED_c = 1 + (4,225 - 1) \cdot 0,089 = 1,284$	$ED = (ED_c) \cdot (ED_e) = 1 \cdot 1,269 = 1,269$ $ED_e = 1$ $ED_c = 1 + (3,829 - 1) \cdot 0,095 = 1,269$
Tasa de respuesta esperada	56,5%	65%	62%
Previsión de no válidos	No se tomó en consideración	10%	5%
Tamaño muestral	3.702	3.221	3.290

Uno de los elementos importantes a tener en cuenta en el cálculo del tamaño de la muestra es el llamado *efecto de diseño*. Es decir, el incremento de la variabilidad de los datos obtenidos por el diseño muestral aplicado en comparación con la variabilidad que se habría obtenido en una muestra aleatoria simple del mismo tamaño. Es bien conocido que la estratificación redundante en una reducción de la varianza dentro de cada estrato, y por ello en una disminución de la variación total, mientras que el uso de conglomerados aumenta la variación de los datos de la muestra resultante (Kish, 1995; Silva, 2000). De ahí la conveniencia de estratificar cuando se usan conglomerados a fin de compensar ambos efectos.

En la muestra de la primera ola debía tenerse en cuenta, además de los ya citados, un tercer efecto de diseño derivado de la diversidad en la composición de los núcleos familiares. Así, por ejemplo, la probabilidad de que un individuo sea seleccionado cuando llaman a su puerta es más alta si el individuo vive sólo (1) que si tiene pareja y tres hijos mayores de 15 años (1/5). Esa desigualdad en las probabilidades de selección produce también un efecto de diseño, puesto que en una muestra aleatoria simple la probabilidad de selección de cada individuo es la misma. En una muestra polietápica, el efecto de diseño total (ED) es igual al producto de los distintos efectos de diseño presentes en la misma (el de estratificación ED_e , el de conglomeración ED_c , y el debido a las diferencias en las probabilidades de selección ED_p).

El problema con el que tuvo que enfrentarse el equipo español en el momento de concretar el tamaño de la muestra de la primera ola era el desconocimiento total acerca del valor aproximado de los efectos de diseño. En los estudios realizados en España rara vez se realiza la distinción entre *tamaño neto* de la muestra y *tamaño efectivo* de la misma. El primero, que incorpora el efecto de diseño, acostumbra a ser mayor que el segundo. Ante la falta de referencias en las que basar las estimaciones de los efectos de diseño, se optó por seguir las recomendaciones del CCT, tanto para el cálculo del efecto del uso de conglomerados (tomar el coeficiente de correlación intra-grupos ρ igual a 0,02) como para la fórmula de cálculo del efecto debido a las distintas probabilidades de selección individual, dejando sin estimar el efecto de la estratificación.

En la segunda ola, sin embargo, se estimaron los efectos de diseño a partir de los datos de la primera ola. Al utilizar una muestra de individuos, desapareció una de las fuentes

de incremento de la variación, la debida a la selección de un individuo dentro de cada domicilio. Por lo que se refiere al efecto de la conglomeración, se comprobó que había sido subestimado (coeficiente de correlación intra $\rho = 0,09$ en lugar del 0,02 utilizado en los cálculos iniciales). Se trató de compensar la elevación del valor de ρ con una disminución del número de individuos seleccionados por sección (6 o 7 en vez de 10 o 12) y con la toma en consideración del efecto de la estratificación. Con esta estrategia se consiguió que el efecto de diseño total esperado fuera menor que el de la ola anterior. Además, se estimó que con las mejoras introducidas en el trabajo de campo, se alcanzaría una tasa de respuesta cercana al 65% y que el número de casos no válidos se reduciría al 10%, con lo que se obtuvo una reducción considerable en el tamaño bruto de la muestra.

En la tercera ola, atendiendo a los resultados de las dos olas anteriores, donde la tasa de respuesta resultó inferior a la prevista, se optó por seguir las recomendaciones del CCT, y tomar como base para los cálculos una tasa de respuesta media del 62% que, sin renunciar al objetivo del 70%, era realista. Por otra parte, a pesar de que los datos de la segunda ola proporcionaron una estimación de ρ ligeramente mayor que en la ola anterior y que el efecto de estratificación era prácticamente igual a 1, la calidad de la muestra de individuos usada en la ola anterior sugirió la conveniencia de rebajar hasta el 5% la previsión de casos inválidos. Con ello el número medio de respuestas esperado por sección⁷ (3,8) resultó ligeramente inferior al previsto para la muestra del 2004 (4,2), redundando todo ello en un menor efecto de diseño debido a la conglomeración. En total, el efecto de diseño previsto y el tamaño bruto de la muestra fueron similares a los de la ola anterior.

Las oscilaciones que muestra la tabla 3 en las previsiones de las tasas de respuesta para las distintas olas se explican, también, por la falta de experiencia en muestreos totalmente probabilísticas en nuestro país. En efecto, la práctica habitual de las empresas de estudios de opinión es la realización de sustituciones automáticas de los casos correspondientes a individuos que se niegan a participar en la encuesta, con lo cual, no se dispone de datos fiables sobre la tasa de respuesta real entre los individuos seleccionados en la muestra inicial. Los datos de algunas encuestas de calidad tipo panel indicaban que difícilmente se podía alcanzar una tasa de respuesta superior al 50%.

⁷ El número esperado de entrevistas completadas por sección se calcula a partir del número medio de casos por sección (6,5) por la tasa de respuesta esperada (62%) y por el número de casos válidos (95%).

Considerando el intenso esfuerzo que la ESE dedica a la realización del trabajo de campo, se estimó que en los estratos rurales se podía llegar a una tasa de respuesta del 60% y en los urbanos del 50%, con lo que la tasa media alcanzaría el 56,5%. En el diseño de la segunda ola, los cálculos se hicieron considerando que la experiencia acumulada permitiría llegar a una tasa de respuesta cercana al objetivo del 70%. Sin embargo, los resultados obtenidos aconsejaron ser más prudentes en el diseño de la tercera ola. La tabla 4 recoge los cálculos realizados para la estimación del tamaño muestral de la tercera ola.

Tabla 4. Determinación del tamaño de la muestra

Concepto	Cálculo	Resultado
tamaño efectivo	fijado por el CCT	1.500
efecto de diseño total	efecto de conglomeración · efecto de estratificación	$1,269 \cdot 1 = 1,269$
tamaño neto	tamaño efectivo · efecto de diseño	$1.500 \cdot 1,269 = 1.904$
casos válidos	tamaño neto / tasa respuesta	$1.904/0,62 = 3.070$
tamaño bruto	casos válidos / proporción de casos válidos	$3.070/0,95 = 3.232$
tamaño bruto final	tamaño bruto + sobrerepresentación	$3.232 + 42 = 3.274$
tamaño bruto final afijado	redondeo del número de conglomerados por sección	3.290

4. El trabajo de campo en España

Es bien sabido que el seguimiento pormenorizado de las actividades del trabajo de campo es un requisito esencial para la garantía de representatividad de la muestra finalmente obtenida. La ESE tiene un protocolo detallado de trabajo de campo, cuya principal herramienta la constituye el formulario de contacto. Se trata éste de un documento que recoge para cada unidad muestral las incidencias habidas en el trabajo de campo, como por ejemplo: intentos de contacto, con sus respectivas fechas y horas; resultados de esos intentos; motivos por los cuales una unidad es declarada no válida; motivos y firmeza en caso de rechazo; código del entrevistador, y características del entorno del domicilio, entre otras. El formato de ese formulario, inicialmente poco ágil, ha sido mejorado por el CCT en la segunda y tercera olas tomando en consideración los informes de campo realizados por los equipos nacionales. En el caso de España, la utilización de CAPI en la segunda y tercera ola ha permitido una mejora notable en la labor de control ejercida por el equipo español (Torcal et al., 2006).

La información contenida en los formularios de contacto es de enorme utilidad, no sólo para la supervisión y el control de las actividades de campo sino también para la validación de las hipótesis utilizadas en el diseño muestral y, en su caso, la introducción de las oportunas adaptaciones en futuras ediciones.

La tabla 5 recoge las principales características del trabajo de campo de la ESE en España en las tres olas realizadas hasta el momento, permitiendo compararlas y comprobar su evolución. Los cambios más importantes se observan entre la primera y segunda olas, mientras que los cambios introducidos en la tercera consisten en una consolidación y profundización de las mejoras introducidas en la segunda. Las mejoras se centran en la formación inicial de entrevistadores y coordinadores, en los materiales de apoyo que se les subministraron, en el establecimiento de incentivos económicos para los entrevistadores y en el aumento del número mínimo de intentos de contacto antes de abandonar una unidad por improductiva.

Por lo que respecta al personal de campo, las principales mejoras introducidas en la segunda ola y que se mantuvieron en la siguiente son: la formación extensiva a todos los entrevistadores (y no sólo a los coordinadores), el aumento del número de sesiones formativas, el suministro de material de apoyo y la participación activa del equipo nacional de la ESE en todo el proceso.

En referencia a los incentivos a los entrevistadores, en la segunda ola se establecieron bonificaciones con el objetivo de aumentar la tasa de respuesta, mientras que en la tercera, se incrementaron los incentivos económicos orientándolos esta vez a la mejora de la fiabilidad de los registros, independientemente de que dieran como resultado o no una entrevista completada.

Finalmente, es de destacar el mayor esfuerzo realizado a partir de la segunda ola en los intentos de localización y contacto de las unidades muestrales, consistente en el establecimiento de un mínimo de siete intentos de contacto (en lugar de los cuatro requeridos por el CCT), de los cuales dos debían ser en horario vespertino y dos en fin de semana (en lugar de uno y uno establecidos por el CCT respectivamente).

Tabla 5. Características del trabajo de campo

Aspecto del trabajo de campo	1ª edición (2002-2003)	2ª edición (2004-2005)	3ª edición (2006-07)
Administración del cuestionario principal	Papel y lápiz (PAPI)	Asistido por ordenador (CAPI)	Asistido por ordenador (CAPI)
Administración del cuestionario complementario	Auto cumplimentado	Integrado en CAPI	Integrado en CAPI
Piloto	100 entrevistas	50 entrevistas	50 entrevistas
Formación de entrevistadores	<ul style="list-style-type: none"> Sólo por Demoscopia. Sólo 1 sesión. Sólo coordinadores de zona. Sin control de contenidos por equipo de ESE. 	<ul style="list-style-type: none"> Realizados por TNS-Demoscopia junto equipo ESE. 7 sesiones. Asistencia de 47 entrevistadores y coordinadores. Mejoras sustanciales en los materiales para los entrevistadores. Entrenamiento específico para maximizar tasa de respuesta. 	<ul style="list-style-type: none"> Realizados por Metrosocopia junto equipo ESE. 5 sesiones de 7 horas. En Madrid y Barcelona. Asistencia de 56 entrevistadores y coordinadores. Actualización de los materiales para los entrevistadores. Entrenamiento específico para maximizar tasa de respuesta incluyendo estrategias de conversión de rechazos
Envío de cartas de presentación del estudio	1 carta, 1 semana antes del comienzo del trabajo de campo.	<ul style="list-style-type: none"> 1ª carta: 2 meses antes del trabajo de campo. 2ª carta: 2 semanas antes del trabajo de campo, con cuartilla informativa. 3ª, específicamente dirigida a “no localizados” o a “rechazos. 	<ul style="list-style-type: none"> 1ª carta: 2 meses antes del trabajo de campo. 2ª carta: 2 semanas antes del trabajo de campo, con cuartilla informativa. 3ª, específicamente dirigida a “no localizados” o a “rechazos.
Otros materiales	Ninguno	Entrega a los entrevistados de un “cuadernillo” con resultados de la 1ª edición en el primer momento de interacción.	Entrega a los entrevistados de un “cuadernillo” con resultados de las olas anteriores en el primer momento de interacción.
Incentivos a los entrevistadores	Pago 10% superior al habitual.	<ul style="list-style-type: none"> Pago 30% superior al habitual Bonificación del 5% por entrevista a quienes alcancen 90% de tasa de respuesta 	<ul style="list-style-type: none"> Pago 30% superior al habitual Bonificación del 10% por entrevista realizada en cada grupo de 24 unidades correctamente completadas Bonificación del 15-20% por entrevista completada a quienes alcancen el 70% de tasa de respuesta Bonificación del 15-20% por entrevista completada a quienes alcancen 90% de tasa de respuesta
Incentivos a los ciudadanos	Regalo simbólico.	Tarjeta-regalo de El Corte Inglés o regalo simbólico.	Bono de compra en grandes almacenes por valor de 12 € en el primer contacto personal
Organización del trabajo de campo	<ul style="list-style-type: none"> Mínimo de 4 intentos 1 en fin de semana + 1 en horario de tarde/noche 	<ul style="list-style-type: none"> Mínimo de 7 intentos 2 en fin de semana + 2 en horario de tarde/noche 	<ul style="list-style-type: none"> Mínimo de 7 intentos 2 en fin de semana + 2 en horario de tarde/noche
Supervisión del trabajo de campo	<ul style="list-style-type: none"> Informes quincenales Proceso muy lento 	<ul style="list-style-type: none"> Informes semanales Proceso “casi-instantáneo” por uso de CAPI. 	<ul style="list-style-type: none"> Informes semanales Proceso “casi-instantáneo” por uso de CAPI.

A pesar del compromiso de realizar un mayor esfuerzo de localización de los individuos seleccionados, el análisis posterior de los datos mostró que buena parte de los entrevistadores de la segunda ola no lo respetaron. No obstante, en la tercera, con una mayor implicación de la empresa encuestadora, los resultados han sido bastante más satisfactorios.

La tabla 6 presenta el resultado final del trabajo de campo en las tres olas de la ESE, detallando el motivo por el cual algunos casos no han finalizado en entrevista. Comparando los resultados de las tres olas, puede apreciarse un aumento sostenido del porcentaje de entrevistas realizadas y una disminución de los rechazos. Aunque en la ola más reciente se mantiene el porcentaje de rechazos, ha habido una disminución importante de los casos no localizados, compensado en parte por un aumento de los casos inelegibles⁸, siendo la mayor parte de estos últimos debidos a cambios de domicilio, ya sean al extranjero o dentro del mismo país. La estricta supervisión del trabajo de campo realizado en la tercera ola, tuvo como consecuencia la anulación de un puñado de entrevistas (contabilizadas en el apartado de “otros”) por haberse detectado errores en la administración del cuestionario. Es posible que si el mismo rigor se hubiera aplicado en la segunda ola, también se habrían detectado anomalías en algunas entrevistas y, por consiguiente, la categoría “otros” no estaría en blanco.

Tabla 6. Resultado final del trabajo de campo en las tres olas

Resultado final	1a ola		2a ola		3a ola	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Entrevista	1.736	47	1.663	52	1.876	58
Rechazo	1.041	28	553	17	559	17
Entrevistado no localizado ^a	419	11	360	11	225	7
Entrevistado no disponible	161	4	368	11	167	5
Unidad inelegible	286	8	180	6	430	13
Cita pendiente	0	0	89	3	0	0
Otros	59	2	0	0	33	0
Total	3702	100	3213	100	3290	100

^a Se trata de las unidades para cuya última visita el resultado es una no localización del individuo, lo cual no quiere decir que nunca se haya establecido contacto con el hogar.

⁸ Los casos inelegibles los constituyen aquellos casos que a pesar de estar incluidos en la muestra no han podido ser entrevistados por error en el dirección del individuo, por cambio de domicilio, por demolición de la vivienda que consta como domicilio o por defunción del interesado.

Tabla 7. Grado de cumplimiento de objetivos

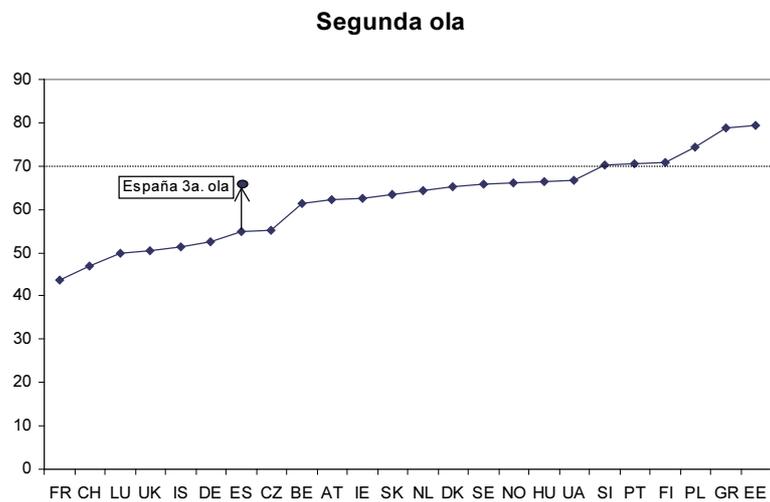
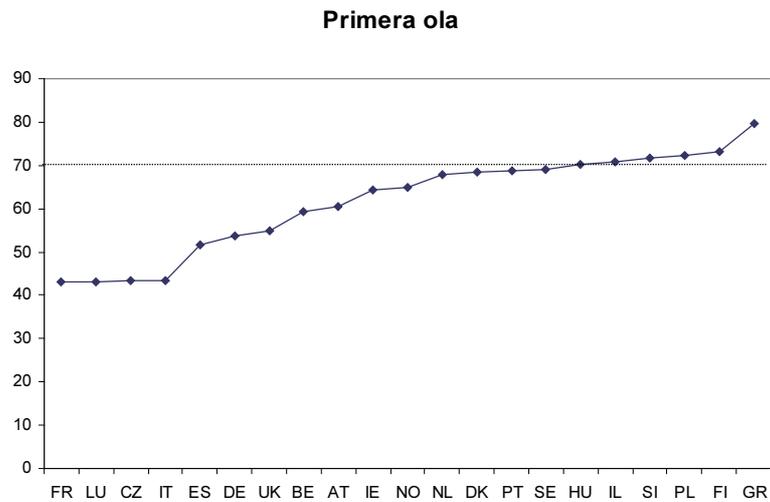
	1a ola	2a ola	3a ola
Número de entrevistas necesarias para un tamaño efectivo de 1.500	2.088	1.829	1904
Número de entrevistas completadas	1.736	1.663	1.876
Porcentaje de respuesta ^a	51,7	54,8	65,6
Entrevistas completadas / requisito de 2.000	0,87	0,83	0,94
Entrevistas completadas / necesarias para tamaño efectivo de 1.500	0,83	0,91	0,98

^a La tasa de respuesta es el cociente entre el número de entrevistas completadas y el número de casos elegibles. En la tercera ola dicha tasa es igual a $1876/(3290-430)$.

La Tabla 7 muestra el grado de cumplimiento en cada ola de los objetivos fijados por el CCT, consistentes en obtener como mínimo 2.000 entrevistas, un tamaño efectivo de 1.500 y una tasa de respuesta del 70%. Aunque en la tercera ola todavía no se han alcanzado estos valores, las mejoras introducidas en el trabajo de campo y su mayor cumplimiento por parte de la empresa encuestadora han permitido acercarse a ellos (98% de las entrevistas necesarias para el tamaño efectivo de 1.500 y tasa de respuesta global del 66%). Merece destacarse que en los dos grupos de municipios de menor tamaño (los de menos de 10.000 habitantes y los de 10.001 a 50.000 habitantes) se ha superado el objetivo del 70%.

No se dispone todavía de los resultados finales de la tercera ola para el conjunto de países participantes. El gráfico 1 permite calibrar la dificultad de alcanzar un 70% en la tasa de respuesta ya que son tan sólo 6 los países que lo han conseguido en cada ola, es decir un 27% y un 25%, respectivamente, del total de países participantes. Grecia, Finlandia, Polonia y Eslovenia son los únicos países que han superado la barrera del 70% en los dos casos, siendo Grecia quien ha obtenido siempre los mejores resultados (cerca del 80% en las dos olas). En el otro extremo, se encuentran Francia, Suiza y Luxemburgo, aunque estos dos últimos países mejoraron sustancialmente su tasa de respuesta en la segunda ola. En el caso de España, para el que sí disponemos de información, la mejora de la segunda a la tercera ola ha sido notable, tal como evidencia el gráfico 1. Es de esperar que la experiencia acumulada de olas anteriores permita también a los demás países participantes en la tercera un incremento significativo de sus tasas de respuesta.

Gráfico 1. Tasas de respuesta de la ESE por países⁹



5. Consideraciones finales

La Encuesta Social Europea se planteó desde sus inicios unos objetivos ciertamente ambiciosos en tres aspectos importantes: la difusión de buenas prácticas y de estándares de rigor en la investigación por encuesta, la calidad y comparabilidad de los datos que proporciona y la generación de indicadores sociales de amplia aceptación en la comunidad académica y con utilidad social. Estando en la fase final de la tercera ola,

⁹ Alemania (DE), Austria (AT), Bélgica (BE), Dinamarca (DK), Eslovaquia (SK), Eslovenia (SL), España (ES), Estonia (EE), Finlandia (FI), Francia (FR), Grecia (GR), Hungría (HU), Irlanda (IE), Islandia (IS), Israel (IL), Italia (IT), Luxemburgo (LU), Noruega (NO), Países Bajos (NL), Polonia (PL), Portugal (PT), Reino Unido (UK), República Checa (CZ), Suecia (SE), Suiza (CH).

quizás es pertinente realizar un cierto balance del camino recorrido y de los resultados obtenidos en esos tres aspectos, por lo que se refiere a la participación de España en el proyecto de la ESE. Para un análisis comparado de la calidad de los datos de las dos primeras olas en el conjunto de países participantes, véase Billet y Pleysier (2007).

Una de las primeras dificultades que tuvo que afrontar el equipo español fue el de superar la desconfianza inicial de la empresa encuestadora respecto lo que ésta consideraba una intromisión del mundo académico en la práctica profesional. La percepción de esta empresa era que ella tenía experiencia consolidada en la realización de encuestas, con una metodología que consideraba suficientemente contrastada y que el papel del equipo nacional de la ESE, compuesto exclusivamente por universitarios, debía reducirse a la preparación del cuestionario. En ese sentido, para los técnicos de dicha empresa fue una novedad que se les suministraran instrucciones precisas para el diseño muestral (totalmente probabilístico), que no se admitieran sustituciones, o que se hablara en términos de tamaño muestral efectivo. Las resistencias a aplicar unas normas que diferían sustancialmente de las habituales y a aceptar el liderazgo del equipo de la ESE en España, crearon ciertos malestares iniciales que se superaron con el tiempo y largas sesiones de trabajo. Esta primera colaboración entre universidad y empresa propició un acercamiento entre dos culturas profesionales distintas, que posibilitó un funcionamiento mucho más fluido en la segunda ola.

En la tercera ola, la experiencia adquirida por unos y otros ha tenido claras consecuencias positivas. Por una parte, el equipo investigador ha definido con mayor precisión el pliego de condiciones asociado al contrato con la empresa de campo y ha realizado una supervisión de sus actividades mucho más estricta. Por otra parte, la empresa encuestadora ha tenido mayor conocimiento de las exigencias de la encuesta de la ESE, ha dispuesto de los medios necesarios para responder a ellas y se ha implicado plenamente en su desarrollo, haciendo suyas las buenas prácticas exigidas como consecuencia del orgullo profesional de protagonizar una encuesta de esta envergadura. La experiencia española muestra que, con respecto a la primera, en la segunda y tercera olas los entrevistadores iniciaron las tareas de trabajo de campo con mayor información sobre los objetivos de la encuesta, conscientes de la necesidad de conseguir la colaboración de los encuestados, con mayor formación sobre estrategias de conversión, y aplicando en todo momento mayor rigor en la documentación de su actividad.

Esta documentación, obtenida a partir de los formularios de contacto, no sólo es importante para asegurar la calidad de los datos e incrementar la tasa de respuesta, sino que permite investigar las características de los encuestados que muestran disponibilidad a colaborar y las de aquellos que son reacios a participar. Ese conocimiento permite elaborar de manera más precisa estrategias de conversión de rechazos y formar a los encuestadores con ese objetivo. Todo ello redundará en una mayor tasa de respuesta y en una mayor representatividad de los datos finalmente obtenidos.

Por lo que respecta al diseño muestral, parece que ha sido acertada la decisión del CCT de dar libertad a cada país para proponer su propio diseño pero, al mismo tiempo, supervisar de cerca el mismo, exigiendo a los participantes la justificación de las opciones que toman y sugiriendo alternativas cuando éstas son mejorables. En España, el diseño muestral ha seguido un proceso de adaptación a la realidad territorial, adaptación que, en cada ola, se ha ido alimentado del análisis de los datos de las olas anteriores. El resultado de este proceso es la aplicación en la tercera ola de una estratificación suficientemente fina, justificada por una variación en la tasa de respuesta según tamaño de municipio y la diversidad de opiniones entre comunidades autónomas, una reducción en el número de etapas de muestreo y una mayor precisión en la estimación de los efectos de diseño. Uno de los parámetros al que inicialmente no se dio demasiada importancia y que ha resultado difícil de estimar ha sido el porcentaje de casos ilegibles presentes en la muestra. Según se desprende del trabajo de campo de la tercera ola, la movilidad de la población española ha incrementado en los últimos años, hecho que deberá ser tenido en cuenta en los diseños de las futuras olas.

De todo lo anterior se deduce que la ESE avanza sin problemas en los objetivos de difusión de buenas prácticas y generación de datos de calidad. Sentadas las bases, ahora es cuando se está en condiciones de avanzar en la creación de indicadores actitudinales y de opinión sólidamente contrastados y de uso generalizado, tanto en el ámbito de la investigación, como en el de la gestión. Sería deseable que, de la misma manera que nadie pone en duda la utilización de determinados indicadores económicos para la toma de decisiones, también se consiguiera el mismo reconocimiento para los indicadores sociales. Éste es el reto que se plantea para el futuro.

Referencias bibliográficas

Billiet, Jaak y Stefaan Pleysier (2007) “Response Based Quality Assessment in the ESS-Round 2. An update for 26 countries” disponible en <http://ess.nsd.uib.no> (survey documentation).

Kish, Leslie (1995) *Survey Sampling*, New Jersey, John Wiley and Sons.

Riba, Clara y Anna Cuxart (2003) “Construyendo las bases para una comparación fiable: la Encuesta Social Europea 2002 en España” en *Revista Española de Ciencia Política*, nº 8, pp.165-185.

Saris, Willem E. y Gallhofer, Irmtraud (2007) *Design Evaluation and Analysis of Questionnaires for Survey Research*, New Jersey, John Wiley and Sons.

Saris, Willem E. y AKos Munnich (1995) *The Multitrait-Multimethod Approach to Evaluate Measurement Instruments*, Budapest, Eötvös University Press.

Saris, Willem E., Albert Satorra y Germà Coenders (2004) “A new approach to evaluating the quality of measurement instruments: the split-ballot MTMM design” en *Sociological Methodology*, vol. 34, pp. 311-347.

Schwartz, S., G. Melek, A. Lehmann, S. Burgess, M. Harris, and V. Owens. (2001) “Extending the cross-cultural validity on the theory of basic human values with a different method of measurement” en *Journal of Cross-cultural Psychology*, vol. 32, pp. 519-542.

Silva, Luis Carlos (2000) *Diseño razonado de muestras y captación de datos para la investigación sanitaria*, Madrid, Ediciones Díaz de Santos.

Torcal, Mariano, Laura Morales y Clara Riba (2006) “Supervisión y control de calidad del trabajo de campo de la Encuesta Social Europea en España: Evaluación y resultados” en *Metodología de Encuestas*, vol. 7, nº 2, pp.75-97.