

## Dimensionando los hechos: la encuesta (II)

**Joan Guix**

*Unidad de Medicina Preventiva y Salud Pública. Universitat Rovira i Virgili.*

*Reus. Tarragona. España.*

*Agència de Salut Pública de Barcelona. Barcelona. España.*

Correspondencia: Dr. Joan Guix.

Agència de Salut Pública de Barcelona.

Pça. Lesseps, 1. 08023 Barcelona. España.

Correo electrónico: jguix@imsb.bcn.es

### Resumen

Los estudios cualitativos son imprescindibles para llegar a conocer los "qué", los "por qué" o los "cómo". No obstante esta información, esencial por otra parte, debe ser complementada mediante su cuantificación. Necesitamos conocer los "cuánto", "cuántas veces" o "con qué intensidad". Éste es el objetivo de la encuesta como herramienta cuantitativa. El cuestionario es, por lo tanto, un instrumento de medida de una realidad determinada.

En un artículo anterior<sup>1</sup> se revisaron los conceptos generales relacionados con la encuesta y las características, los aspectos positivos y negativos de los diversos tipos de realización de una encuesta (telefónica, presencial, postal).

En el presente artículo se presta una atención especial al diseño de encuesta, esto es: la selección de los ítems a tratar, la formulación de preguntas, los diversos formatos de respuesta (abiertas, Thurstone, Osgood, Likert, etc.), y se incide en la imprescindible validación de los cuestionarios, puesto que se trata de herramientas de medida, especialmente en cuanto a comprensión, fiabilidad, validez y buena adaptación cultural.

**Palabras clave:** Encuesta. Cuestionario. Validación. Fiabilidad. Validez.

En una primera parte de este artículo hemos planteado los principios generales de las encuestas y sus distintos tipos<sup>1</sup>. No obstante, no podemos olvidar que la encuesta es una herramienta de medida, por lo cual debemos estar completamente seguros de que estamos midiendo aquello que deseamos medir, y que los resultados obtenidos se corresponden con la realidad. Éstos son los aspectos que pretendemos tratar a continuación.

### Diseño de encuestas

#### Selección de ítems

Como ya se ha mencionado anteriormente, la encuesta es una herramienta de medida. Lógicamente, pues, al diseñar una encuesta debemos plantearnos ante todo qué queremos medir. Este aspecto nos lo debe proporcionar el mismo objetivo de la investigación. Ahora bien, la mayor parte de los fenómenos de tipo social presentan un cierto nivel de complejidad para ser abordados. Las realidades sociales son siempre multidimensionales. Para decidir qué preguntas formulare-

### Abstract

Qualitative studies are essential to determine "what?", "why?" and "how?" This essential information must be complemented by measures of its quantification. We also need to know "how much?", "with what frequency?" and "with what intensity?". The aim of this study was to evaluate the survey as a quantitative tool. The questionnaire is, therefore, an instrument of measurement of a certain reality.

In a previous study<sup>1</sup> we reviewed the general concepts of surveys, their characteristics, and the positive and negative features of distinct ways of performing them (telephone, face-to-face, postal, etc.).

The present study pays special attention to the questionnaire design, i.e., to the selection of items to be included, the formulation of questions, and the diverse answer formats (open, Thurstone, Osgood, Likert, etc.). An essential element of questionnaires is their validation, as a measurement tool, especially validation of comprehension, reliability, validity and cultural adaptation.

**Key words:** Survey. Questionnaire. Validation. Reliability. Validity.

mos es imprescindible identificar previamente las diversas dimensiones que componen el fenómeno a estudiar y, una vez identificadas, establecer cuáles son los ítems concretos que contienen. Para identificar las dimensiones básicas podemos utilizar diversas metodologías, desde la simple definición por expertos hasta el recurrir a la bibliografía. Cuando el fenómeno ha sido poco estudiado o el equipo de investigación considera que no conoce con la suficiente profundidad los diversos aspectos de la realidad que desea estudiar, solemos recurrir a alguna de las diversas metodologías cualitativas disponibles<sup>2</sup>, especialmente a los grupos focales, las entrevistas semiestructuradas<sup>3-6</sup> o al análisis de incidentes críticos<sup>7</sup>. Así, por ejemplo, para describir el fenómeno satisfacción con las organizaciones de servicios, Zeithaml et al<sup>6</sup> refieren que, de acuerdo con sus investigaciones, las dimensiones que componen este fenómeno son 5: tangibilidad (la impresión que nos transmite la apariencia de las estructuras y las personas), capacidad de respuesta, seguridad, fiabilidad y empatía (capacidad de ponerse en el sitio del cliente). Al mismo tiempo, cada una de estas dimensiones puede descomponerse en varios ítems o variables concretas. Siempre será importante que

para el análisis de cada una de las dimensiones utilizemos diversos ítems, con el fin de dimensionarlas adecuadamente. Cuando queramos describir fenómenos, nos hará falta buscar una alta consistencia interna de los ítems, mientras que si queremos discriminar entre personas, la necesidad de esta consistencia interna será menor. Nos hará falta garantizar que un conjunto de ítems componentes de una misma dimensión en realidad midan aspectos distintos de una misma realidad, y no distintos aspectos de diversas realidades<sup>8</sup>. Cuando conozcamos las dimensiones y las variables que queremos medir, es decir, la información concreta que deseamos, entonces deberemos traducirlo en preguntas concretas.

### **Las preguntas del cuestionario**

Para Grawitz, citado por Sierra<sup>9</sup>, una pregunta bien formulada es aquella que no ejerce ninguna influencia en el sentido de la respuesta y que no incita a dar una respuesta inexacta, que no corresponde a la información buscada. Para lograr estas condiciones, debemos observar una serie de consideraciones a la hora de redactar las preguntas concretas y también es preciso plantearnos diversas cuestiones<sup>3,8,10-13</sup>: ¿Es necesaria la pregunta? ¿Tiene el encuestado la información necesaria para contestar la pregunta? ¿El objetivo está dentro de las vivencias del encuestado? ¿Puede recordar la información? ¿Responder la pregunta supone un gran esfuerzo por su parte? ¿Si no nos contesta la pregunta, es porque no quiere o bien porque es capaz de expresarse? ¿Para conseguir que cada pregunta contenga tan sólo un concepto, harán falta varias preguntas para lograr la información deseada en lugar de una sola?

Otro elemento previo a la redacción de las preguntas es decidir el tipo de respuesta que plantearemos. Básicamente hablamos de respuestas abiertas y de respuestas cerradas. En el primero caso dejamos un espacio en blanco para que el encuestado responda libremente y con sus propias palabras a la pregunta formulada. Esto lo podemos hacer mediante el encuestador, que reproducirá exactamente, literalmente, la respuesta del encuestado, o bien mediante la escritura directa del encuestado, es decir, de forma autoadministrada. Las ventajas y desventajas de cada una de las modalidades ya han sido examinadas anteriormente. Las preguntas con respuesta abierta nos son especialmente útiles para conocer aspectos del fenómeno que queremos estudiar y que nosotros podemos desconocer. Tiene un cierto parecido con la entrevista estructurada. Esta fórmula está indicada cuando las informaciones que queremos recoger son complejas y en estudios exploratorios. Las preguntas con respuesta abierta permiten añadir observaciones y comentarios, con lo cual la información desaprovechada es menor. No se inducen las respuestas, y éstas son más libres y abiertas. Al mismo tiempo, nos permite contrastar el grado real de información que el encuestado tiene sobre el tema preguntado. La redacción siempre es mucho más sencilla y son especialmente útiles para romper el fuego, al inicio de los cuestionarios, puesto que tienen una mayor capacidad por motivar al encuestado a responder. Los problemas esenciales de las preguntas abiertas son su manejo y explotación. Las respuestas abiertas son extremadamente difíciles de codi-

ficar y el manejo necesario para hacerlo se asemeja al de las metodologías cualitativas, con todas las potencialidades y los problemas que comportan, entre las cuales hay la subjetividad de los evaluadores. La codificación es compleja: se trata de convertir una herramienta cualitativa en cuantitativa. Es un proceso lento y caro que requiere más tiempo y más esfuerzo. No es útil en respuestas sobre temas considerados socialmente incorrectos. De alguna manera, este tipo de respuesta es más útil para ser aplicado en personas con un grado más alto de educación, ya que sus respuestas pueden traducirse en una mejor expresión que en una persona de nivel cultural inferior. La alternativa es la encuesta con respuestas cerradas. En este tipo de cuestionario se especifica previamente la totalidad de respuestas posibles y el encuestado se limita a escoger generalmente una sola.

La mayor ventaja de las encuestas de respuestas cerradas es su facilidad para codificar y explotar los resultados, puesto que se obtiene una mejor precisión y uniformidad en la recogida de datos y esto facilita la comparación de respuestas. Es también un tipo de encuesta más rápido de contestar. Su principal utilidad es clasificar los individuos respecto de la pregunta realizada, puesto que no aporta más información que la que estrictamente pide el cuestionario. Es interesante utilizarlas para preguntar sobre cuestiones personales o delicadas. Hay pérdida de información porque no se aprovecha toda la de que dispone la fuente, es decir, el encuestado. No es del todo útil para información compleja. No es capaz de aportar matices por parte del encuestado más allá de las que el investigador le pueda ofrecer. Debemos conocer la totalidad de respuestas posibles para formular el mayor número de respuestas cerradas posibles. No nos aporta nuevos conocimientos, sino que cuantifica los que ya tenemos. Las respuestas planteadas tienen que ser exhaustivas y mutuamente excluyentes. El número de respuestas tiene que ser tan bajo como sea posible. La redacción del cuestionario es más compleja. Hace falta considerar el lenguaje utilizado y la posible ambigüedad de las palabras o el sentido de las frases. El orden de las respuestas puede inducir el sentido de las mismas, puesto que está contrastado el hecho que la respuesta formulada en primer lugar tiene más posibilidades de ser escogida<sup>7</sup>.

Es conveniente partir desde las preguntas más generales a las más concretas, utilizando la denominada técnica del embudo<sup>10</sup>. El orden del planteamiento de las respuestas tiene que ser lógico. Es relativamente frecuente utilizar en un mismo cuestionario preguntas abiertas y cerradas. En este caso, las preguntas abiertas se suelen situar ya sea al inicio o al final, y si se utilizan mezcladas, suele ser por precisar algún tipo de respuesta cerrada. En todos los casos, el lenguaje tiene que ser sencillo y claro, sin palabras o frases ambiguas. Las preguntas no deben ser indiscretas sin necesidad y, en la medida en que podamos, no deben provocar el afloramiento directo de prejuicios. Además, tienen que ser tan pocas como sea posible, y adecuadas al objetivo de la investigación (no debemos intentar aprovechar el cuestionario para introducir otros temas diferentes de los propios de la investigación estricta). Es recomendable formular las preguntas en positivo, puesto que las redactadas en negativo pueden comportar confusiones. Las preguntas han de estar formuladas en términos

neutros, sin adjetivos que puedan inducir o condicionar las respuestas, y debemos considerar si el orden de las preguntas puede condicionar las respuestas. Es imprescindible encuadrar siempre el cuestionario antes de iniciarlo y explicar el porqué de éste.

### Escalas de respuesta

Las preguntas se plantean con el fin de obtener información mediante respuestas. Cuando optamos por preguntas cerradas, la manera como planteamos las respuestas puede condicionar la calidad de la información recogida. Por eso es por lo que recurrimos a escalas de respuestas previamente validadas. Una escala es un instrumento que se utiliza para determinar diferencias de grado o de intensidad respecto de algún ítem o algún objeto actitudinal, entendiendo por actitud una disposición que condiciona el sujeto a percibir y reaccionar de una manera determinada. Existen múltiples formas de clasificación de las diversas escalas. Siguiendo al mencionado Del Rincón<sup>13</sup>, las podemos clasificar en escalas estimativas, escalas de actitud y escalas dicotómicas. Las escalas estimativas sitúan a la persona en un punto de una sucesión teóricamente continua o bien en una categoría de una serie ordenada, es decir, la persona encuestada se sitúa respecto de una frase (una opinión) concreta. Dentro de las escalas estimativas distinguimos las escalas gráficas, las conceptuales y las comparativas. En las escalas gráficas, el sujeto interrogado escoge un punto situado sobre una escalera construida previamente, que por ella misma sugiera un posicionamiento. Las escalas entre 1 y 10 son un ejemplo. Las escalas conceptuales presuponen el establecimiento de una relación de categorías definidas previamente y ordenadas de forma progresiva. Las escalas comparativas piden emitir un juicio relativo a un grupo de pertenencia. Las escalas dicotómicas (que sólo admiten por respuesta sí o no) proporcionan una información mucho más limitada cuando queremos analizar actitudes, en el sentido que no admiten posicionamientos intermedios. Las escalas de actitud se crearon para lograr estimaciones de actitudes más refinadas; estas escalas presentan al encuestado una sucesión de ítems o de frases en relación con cada una de las dimensiones que queremos estudiar. Lo que se busca es la puntuación total del sujeto y no tan sólo el mero posicionamiento del sujeto respecto del tema. Dentro de las escalas de actitud podemos mencionar como más conocidas las escalas diferenciales o de Thurstone, las aditivas o de Likert, las acumulativas o de Guttman y las diferenciales semánticas o de Osgood.

La actitud es definida como las “disposiciones permanentes de ánimo, formadas por convicciones y sentimientos, que hacen pensar, actuar y expresarse de acuerdo con ellas en los actos y opiniones en todos los campos de su relación social”<sup>9</sup>. Una actitud se compone de 3 factores:

1. Factor ideológico, que comporta ideas y convicciones.
2. Factor sentimental, que comporta sensaciones de simpatía o antipatía.
3. Factor reactivo, que impulsa las personas a actuar de acuerdo con sus convicciones y los sentimientos<sup>9</sup>.

### Escala de Thurstone

Históricamente, la primera al surgir fue la escala diferencial o de Thurstone. En este tipo de cuestionario se plantean entre 20 y 22 preguntas, positivas, negativas o neutras, respecto de una actitud determinada, dispuestas de una manera determinada según un complejo acuerdo por parte de los investigadores, denominado “de jueces”<sup>13</sup>, en el cual se puntúa cada ítem entre 1 y 11 y considera el valor 1 como el máximo acuerdo y el valor 11 como el de máximo desacuerdo<sup>9</sup>. A partir de las conclusiones de los jueces se plantea la encuesta definitiva, a la cual el encuestado debe contestar con un “de acuerdo” o “desacuerdo” a cada una de las frases o ítems planteados. La suma de las puntuaciones obtenidas, en base a las ponderaciones de los jueces, resumirá la actitud del individuo frente al tema investigado<sup>3</sup>. Es preciso que las proposiciones cubran la continuidad de la actitud. Un ejemplo de este tipo de escala lo podemos encontrar en la tabla 1.

Esta escala es altamente difícil de elaborar y queda la duda de si realmente la opinión de los jueces es una base objetiva de ponderación.

### Escala de Osgood

La escala de Osgood, o diferencial semántica, hace que el encuestado se posicione entre 2 conceptos encontrados (agradable frente a desagradable) referidos a un ítem representativo de una actitud, con una gradación intermedia entre ambos de 7 niveles: se da un valor 1 a la respuesta más negativa y un valor 7 a la más positiva. La puntuación global es la media de las puntuaciones obtenidas<sup>13</sup>. Este tipo de escala permite el desarrollo de perfiles descriptivos que facilitan la comparación de artículos competitivos, por lo cual se utiliza en el estudio comparativo de marcas y/o productos<sup>14</sup>. Un ejemplo de esta escala puede verse en la tabla 2.

### Escala de Guttman

La escala de Guttman o escalograma pretende establecer una gradación jerárquica entre frases. Esta escalera tiene como limitación el hecho que tan sólo mide la actitud respecto de una sola característica, y no de diversas. Lo que hace es medir las diversas intensidades de la actitud. Las frases están

Tabla 1. Ejemplo de cuestionario tipo Thurstone

Sobre el tema aspiraciones de los padres respecto del logro educativo de sus hijos:

- Lograr el éxito es tan sólo una forma de que mis hijos compensen mis esfuerzos como padre  
De acuerdo ... Desacuerdo ...
- Ir a una buena escuela y obtener un buen trabajo son importantes, pero no esenciales, para la felicidad de mis hijos  
De acuerdo ... Desacuerdo ...
- La felicidad no tiene nada que ver con lograr hitos educativos o materiales  
De acuerdo ... Desacuerdo ...

Tabla 2. Ejemplo de cuestionario tipo Osgood

Cómo definiría las amistades de sus hijos

Estudiosos / \_\_\_ / \_\_\_ / \_\_\_ / \_\_\_ / \_\_\_ / No estudiosos

Puntuación: 7 6 5 4 3 2 1

Desobedientes / \_\_\_ / \_\_\_ / \_\_\_ / \_\_\_ / \_\_\_ / Obedientes

Puntuación: 1 2 3 4 5 6 7

Tímidos / \_\_\_ / \_\_\_ / \_\_\_ / \_\_\_ / \_\_\_ / Descarados

Puntuación: 7 6 5 4 3 2 1

ordenadas jerárquicamente, de mayor a menor intensidad, y la elección de una frase presupone la aceptación de las de intensidad inferior. El encuestado tiene que definirse con un sí o un no con respeto de la frase formulada en el enunciado. Es un tipo de escala que no discrimina lo suficiente entre individuos<sup>8</sup>. Un ejemplo de jerarquización de frases en el sentido planteado en la escala de Guttman se recoge en la tabla 3.

En este caso, el encuestado mide más de 1,85 metros, ya que al aceptar la primera frase, por ejemplo, acepta las dos siguientes, puesto que tiene un nivel jerárquico superior. Ésta es una escala difícil de construir y de plantear al encuestado.

### Escala de Likert

En un conjunto, todas estas escalas son complicadas de elaborar y de plantear, y por eso es por lo que la escala actitudinal más utilizada actualmente es la escala aditiva de Likert, por su sencilla elaboración y fácil respuesta, porque logra mayores coeficientes de fiabilidad con menos preguntas y porque, además de poder obtener una puntuación global, permite presentar porcentajes de grado de acuerdo o desacuerdo respecto de ítems concretos<sup>7</sup>. La escala de Likert expresa una continuidad bipolar entre una actitud muy favorable y una muy desfavorable, con una gradación en cada uno de estos sentidos y un punto de neutralidad intermedio<sup>3,14</sup>. Está basada en la escala de Thurstone y considerada más simple y segura que ésta. Es una escala aditiva con una puntuación final en torno a una actitud determinada. Busca conocer no tan sólo la dirección de la actitud, sino también su intensidad<sup>3,14,15</sup>. Una encuesta tipo Likert se elabora formulando, en primer lugar, toda una serie de frases en relación con la actitud que queremos estudiar, algunas en sentido positivo y otras en sentido negativo. Para cada una de estas frases la escala plantea un posicionamiento que pasa por “muy de acuerdo”, “de acuerdo”, “ni

Tabla 3. Ejemplo de cuestionario tipo Guttman

¿Mide usted más de 1,80 metros? Sí \_\_\_ No \_\_\_

¿Mide usted más de 1,70 metros? Sí \_\_\_ No \_\_\_

¿Mide usted más de 1,60 metros? Sí \_\_\_ No \_\_\_

de acuerdo ni en desacuerdo”, “en desacuerdo” y “muy en desacuerdo”. Las respuestas, que en el contexto de la frase impliquen una actitud favorable, puntúan 5 puntos la más favorable y un punto la más desfavorable. Podemos expresar una frase en sentido negativo con respecto a la actitud que se ha de estudiar y, en este caso, “un muy de acuerdo” en esta frase comportaría 1 punto, y un “muy en desacuerdo”, 5 puntos. La puntuación global de la actitud queda representada por la suma total de los puntos de los artículos referidos a aquella actitud concreta. Esta escala no necesita un acuerdo de jueces previo, como la escala de Thurston, sino que con una prueba sobre un grupo limitado de personas podremos discriminar cuáles de las frases formuladas pueden pasar a formar parte de la encuesta definitiva y cuáles son redundantes. Está demostrado que más de 5 frases o artículos por actitud no aportan más fiabilidad al estudio<sup>7</sup>. Un ejemplo para analizar actitudes con respecto del hábito de fumar se recoge en la tabla 4.

Si lo que queremos es medir la actitud contraria al tabaco, en el caso del ítem primero, como se puede ver, puntuaríamos el “muy de acuerdo” con un 5, la puntuación máxima, puesto que se expresa en contra del tabaco, mientras que en el segundo caso el “muy de acuerdo” se puntuaría con un 1, la puntuación mínima, puesto que expresa una actitud favorable al tabaco. Un problema de la escala de Likert está en el hecho que se pueden obtener puntuaciones similares o iguales con diferentes combinaciones de respuestas<sup>3</sup>. Ahora bien, la expresión porcentual del acuerdo o desacuerdo aclararía el resultado.

### Validación de la encuesta

Al inicio de este artículo nos hemos referido a la encuesta como un instrumento de medida. Es obvio, pues, que, si lo que pretendemos es medir, haya 2 extremos que nos preocupen: si el instrumento realmente mide aquello que queremos medir, y si el valor de la medida corresponde a la realidad del hecho medido. El proceso mediante el cual, una vez elaborada la encuesta, confirmamos su idoneidad para aquello para lo que la hemos diseñado lo denominamos validación. Las características básicas de una

Tabla 4. Ejemplo de cuestionario tipo Likert

Se debería prohibir fumar en todos los lugares públicos

Muy de acuerdo (5)

De acuerdo (4)

Ni de acuerdo ni en desacuerdo (3)

En desacuerdo (2)

Muy en desacuerdo (1)

El precio del tabaco debería bajar

Muy de acuerdo (1)

De acuerdo (2)

Ni de acuerdo ni en desacuerdo (3)

En desacuerdo (4)

Muy en desacuerdo (5)

buena encuesta son: sencillez, viabilidad, aceptabilidad, fiabilidad, validez, buena adaptación cultural, utilidad y sensibilidad<sup>8</sup>. De todas estas características, las básicas son la fiabilidad y la validez. La sensibilidad, es decir, la capacidad de captar leves variaciones en la realidad, es importante si se desea medir cambios o tendencias, pero no tanto en estudios sobre opciones o actitudes. La sencillez en cuanto al lenguaje, la presentación y la estructuración de las respuestas facilitará enormemente la respuesta a la encuesta. La viabilidad, es decir, si la herramienta será utilizable en la práctica, depende esencialmente de su fácil aplicación, pero también interviene en la viabilidad el hecho que sea económica en cuanto a su realización, que sea capaz de despertar el interés del encuestado, que sea breve y clara con respecto a las preguntas y respuestas y que sea fácil de codificar, registrar y corregir.

La encuesta tiene que ser aceptable para la población diana. Debe estar redactada en términos correctos, no ofensivos e intentar evitar temas potencialmente conflictivos, ya sea abordándolos de manera suave o bien mediante preguntas indirectas.

Fiabilidad o seguridad implica que la herramienta produce resultados consistentes cuando se aplica en diferentes ocasiones<sup>9</sup>, es decir, el grado en que una escala ofrece los mismos resultados al ser aplicada más de una vez a un mismo grupo en situaciones similares<sup>13</sup>.

La fiabilidad comporta una serie de componentes:

1. Repetitividad: calidad de obtener resultados similares con más de una encuesta hecha a los mismos sujetos (test-retest). Implica que entre ambos pases no puede transcurrir demasiado tiempo y que los mencionados sujetos aceptan la repetición.

2. Fiabilidad interobservador: concordancia entre 2 o más observadores. Si es alta, también lo es la intraobservador. Si no lo es, no se puede asegurar si hay diferencias entre los observadores o en un observador. Para medirla, cuando la escala es cualitativa, utilizamos el test de kappa, y si es cuantitativa, el coeficiente de correlación intraclase.

3. Consistencia interna: evalúa la homogeneidad entre los diversos ítems que miden un mismo atributo. La medida de la consistencia interna implica pasar el cuestionario una sola vez. Se evalúa con el test de la  $\alpha$  de Cronbach, que mide el grado de acuerdo entre las diversas preguntas que exploran un mismo atributo e indica la profundidad con que los puntos del cuestionario están interrelacionados, a partir del análisis de la varianza de los puntos individuales y las covarianzas entre los puntos. Es el método más frecuentemente utilizado<sup>16</sup>. Se pondera de 0 a 1 y corresponde al valor 1 la máxima consistencia interna. También se puede hacer mediante la comparación entre las 2 mitades de la escala aplicando la correlación  $\rho$  con corrección de Spearman-Brown. Si el cuestionario contiene respuestas correctas e incorrectas, se utiliza la fórmula de Kuder-Richardson, que tiene en cuenta el número de individuos que aciertan las respuestas. El test de Tukey (test de no aditividad) informa de si los individuos han mantenido durante todo el test sus tendencias de respuesta en forma estable<sup>15</sup>.

La fiabilidad puede verse afectada por<sup>12</sup>.

1. Cambios de la característica estudiada en el tiempo.
2. Cambios debido a las condiciones de administración del cuestionario.
3. Variaciones del propio cuestionario.
4. Cambios causados por los encuestadores.
5. Errores en el manejo de los datos.

### Validez

Es el grado con que el cuestionario mide aquello que quiere medir.

La validez tiene una serie de componentes<sup>8</sup>.

1. Validez lógica: grado con que un cuestionario o una parte de él mide aquello que quiere medir. Su evaluación no tiene una medida objetiva y se hace a partir de la opinión de expertos.

2. Validez de contenido: si establece todos los aspectos relacionados con el concepto de estudio. Se evalúa a partir de la opinión de expertos, aun cuando podemos utilizar el análisis factorial de componentes<sup>17</sup>, que detectará los factores que se pretendía estudiar.

3. Validez de criterio: en muchas ocasiones no se puede medir, aun cuando, si es posible obtenerla, es la más importante. Presupone la comparación de los resultados obtenidos por el cuestionario con un patrón oro. Aun así, en muchas ocasiones no existe este patrón de referencia. Caso de disponer de él, mediríamos su concordancia. Como subcomponentes de la validez de criterio diferenciamos:

– Validez concurrente: a fin de evaluarla aplicamos la nueva medida (cuestionario) y la estándar conjuntamente. Utilizamos el análisis de la concordancia para dimensionarla.

– Validez predictiva: por analizarla se administra primero la nueva herramienta de medida (cuestionario) y, posteriormente, la estándar. Para medirla, en estudios cuantitativos, se utiliza la sensibilidad y la especificidad. En estudios cualitativos, el coeficiente de correlación intraclase.

– Validez de constructo o de concepto: cuando no disponemos de estándar (referencia) se utiliza esta medida. Se identifican grupos de opiniones extremas y se analiza la correlación. Si ello no es factible, se evalúa la validez convergente y la divergente, es decir, la correlación con ítems relacionados y no relacionados.

Varios elementos pueden influir sobre la validez<sup>8</sup>:

1. Orden de las preguntas: debemos situar siempre la más conflictiva al final del cuestionario.
2. Redacción de las preguntas: existe la posibilidad de obtener respuestas sesgadas.
3. Errores en la categorización de las respuestas: siempre ha de existir una opción de respuesta adecuada.
4. Sesgo de conveniencia social: no posicionarse en contra de la opinión supuestamente dominante en nuestra sociedad por miedo de ser mal considerado.

5. La existencia frecuente de una tendencia sistemática a dar siempre la misma respuesta (favorable o contraria).

6. Las características personales del encuestador, que pueden afectar el sentido de la respuesta.

7. La tendencia a la elección de la puntuación o categoría intermedia para no comprometerse.

Si se aumenta el número de ítems, se incrementan la validez y la seguridad de una escala.

### ***Buena adaptación cultural***

Es ideal el poder utilizar una encuesta previamente validada. Ahora bien, en caso de que esta encuesta validada haya sido realizada en un entorno cultural diferente del nuestro, debemos garantizar su buena adaptación cultural. Para lograrlo, hará falta que la encuesta original, en su idioma original, sea traducida a nuestro idioma por un traductor bilingüe, capaz de captar el significado no tan sólo de las palabras, sino también de las expresiones. Una vez realizada la primera traducción, un segundo traductor bilingüe, diferente del primero, traducirá el cuestionario obtenido en nuestro idioma, procedente de la primera traducción nuevamente al idioma de origen. Así obtendremos una segunda versión en el idioma original. Compararemos la original verdadera con la segunda traducción al idioma original y se analizarán las posibles diferencias. De esta manera garantizamos una versión correctamente adaptada a nuestro entorno cultural y lingüístico<sup>18-20</sup>.

### **Presentación**

La presentación del cuestionario es otro elemento muy importante, especialmente en el caso de encuestas autoadministradas, hasta el punto de poder facilitar un mayor nivel de respuesta<sup>12</sup>. Pretendemos transmitir al encuestado la convicción de que responder al cuestionario es importante. Ello implica que el soporte esté a la altura. Sin caer en lujos excesivos, el papel debe ser de buena calidad y, caso de estar compuesto de más de una hoja, ayudará una composición en cuadernillo, descartando el grapado. Un tamaño tipo A4 o similar será adecuado. Formatos mayores o más pequeños dificultarán su manejo. Las páginas estarán numeradas y el tipo de letra será claro y con un tamaño no demasiado pequeño. Un tipo Arial, Courier o Times, con un tamaño 10 o 12, darán a la redacción un aspecto limpio y fácil de leer. El diseño deber ser "airoso", con espacios suficientes entre las preguntas como para no dar una impresión de densidad y pesadez. Las escalas de respuesta deben ser fáciles de comprender y su disposición no debe llevar a errores. Un cuadro para ser marcado al lado de cada una de las posibles respuestas será más comprensible en los cuestionarios de respuestas cerradas.

Un caso aparte lo constituyen las llamadas preguntas (o respuestas) filtro, es decir, aquellas que, al ser condicionales, comportan el saltar a otras preguntas que no son las consecutivas (p. ej.: Pregunta 28 "caso de no haber asistido, pase directamente a la pregunta 35"). Este tipo de preguntas puede ser útil de cara a facilitar el posterior manejo y explotación del

cuestionario, pero pueden ser complejas e incluso tediosas para el encuestado. Si no podemos evitarlas, deben de ser las menos posibles y los saltos han de estar formulados de forma muy clara, o incluso mediante flechas o líneas de ayuda.

Por último, recordemos que estamos solicitando para nosotros un favor importante al encuestado. Es elemental iniciar el cuestionario con una demanda cortés, exponiendo brevemente el objetivo y el interés de la investigación que estamos realizando, y lo valioso de su colaboración en él. Al finalizar, un agradecimiento para el encuestado por su ayuda.

### **Instrucciones**

En el caso de las encuestas con encuestador, tanto las presenciales como las telefónicas, es necesario el entrenamiento del encuestador y, para ello, es imprescindible la existencia de unas instrucciones escritas.

Estas instrucciones, dirigidas a ayudar al encuestador en el trabajo de campo, incluirán una fórmula de presentación estándar, consejos sobre la adecuación del aspecto personal, en consonancia con el ámbito en el que se deba desarrollar la encuesta, así como indicaciones claras sobre su actitud y la manera de plantear las preguntas de tal forma que no induzcan unas determinadas respuestas. El encuestador deberá comprender perfectamente el contenido y el sentido del cuestionario, la terminología utilizada, y los formatos de respuesta cuando se trate de encuestas con respuestas cerradas. También deberá recibir instrucciones claras acerca de cómo transcribir las respuestas abiertas.

En lo que se refiere a los cuestionarios autoadministrados, las instrucciones van dirigidas a los encuestados, por lo que debemos emplear un lenguaje sencillo y sin tecnicismos, evitando, en la medida de lo posible, la utilización de siglas o abreviaturas. Debemos intentar evitar que el encuestado deba codificar alguna de sus respuestas, ya que es tedioso y puede dar lugar a confusiones. Las instrucciones deberán cumplir los mismos requerimientos de formato y presentación que el cuestionario, aunque deben estar claramente diferenciadas, por ejemplo utilizando otro color de papel. Se deberá insistir en la definición de los conceptos utilizados y explicar la forma de responder a las preguntas mediante algún ejemplo. Siempre que sea posible, será bien recibido facilitar instrumentos para la respuesta, tales como sobre franqueado, lápiz o bolígrafo, y debe explicarse con absoluta claridad el mecanismo de envío de respuestas al equipo investigador.

Es importante inducir al encuestado a leer las instrucciones antes de responder al cuestionario. No obstante, el colocar las instrucciones delante del cuestionario podría desanimar al encuestado, transmitiendo una sensación de complejidad y desinteresándolo con respecto al contenido de la encuesta. Una buena opción es colocar al inicio del cuestionario una nota destacada recomendando, para facilitar la contestación del cuestionario, la lectura previa de las instrucciones.

Estamos ya en disposición de realizar el trabajo de campo.

## Bibliografía

1. Guix J. Dimensionando los hechos: La encuesta (I). *Rev Calidad Asistencial*. 2004;19:402-6.
2. Guix J.  $p < 0,05$ , sí pero ... la aplicación de la metodología cualitativa en la investigación sanitaria. *Rev Calidad Asistencial*. 2003;18:55-8.
3. Cea MA. *Metodología cuantitativa*. Madrid: Síntesis; 1996.
4. Carbonell JM, Guix J, Fernández-Ballart J, Xifró A. Opinión de los médicos de atención primaria acerca de su hospital de referencia. *Todo Hospital*. 1997;140:7-11.
5. McKinley R, Manku-Scott T, Hastings AM, French DP, Baker R. Reliability and validity of a new measure of patient satisfaction with out of hours primary medical care in the United Kingdom: development of a patient questionnaire. *BMJ*. 1997; 314:193.
6. Zeithaml V, Parasuraman A, Berry L. *Calidad total en la gestión de servicios*. Madrid: Díaz De Santos; 1993.
7. Hayes B. *Cómo medir la satisfacción del cliente. Desarrollo y utilización de cuestionarios*. Barcelona: Ediciones Gestión 2000;1995.
8. Argimón Payas JM, Jiménez Villa J. *Métodos de investigación clínica y epidemiológica*. Barcelona: Harcourt; 1999.
9. Sierra Bravo R. *Técnicas de investigación social. Teoría y ejercicios*. Madrid: Paraninfo; 1995.
10. García Ferrando M. La encuesta en el análisis de la realidad social. En: García Ferrando M, Ibáñez J, Alvira F, editores. *Métodos y técnicas de investigación*. Madrid: Alianza Universidad Textos; 1996.
11. Doménech Massons JM. *Técnicas de muestreo y división aleatoria*. Documentos del Laboratorio de Psicología Matemática. Bellaterra: Universidad Autónoma de Barcelona; 1995.
12. Boyd H, Westfall R, Stasch S. *Investigación de Mercados*. 5.ª ed. México: Grupo Noriega Editores; 1990.
13. Del Rincón I. *Técnicas de investigación en ciencias sociales*. Madrid: Kimsan; 1995.
14. Audes-Egg E. *Técnicas de investigación social*. Madrid: Lumen; 1995.
15. Mira JJ, Orozco D. Estudios de validación y aplicación de cuestionarios (aplicados a diabetología, salud mental o calidad de vida). En: *Tratado de calidad en atención primaria*. Madrid: Du Pont Pharma; 1995.
16. AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines. The Agree project. *Qual saf Health Care*. 2003;12:18-23.
17. Wert JI. La encuesta telefónica en el análisis de la realidad social. En: García Ferrando M, Ibáñez J, Alvira F, editores. *Métodos y técnicas de investigación*. Madrid: Alianza Universidad Textos; 1996.
18. Badia X. Sobre la adaptación transcultural de medidas de la calidad de vida relacionada con la salud para su uso en España. *Med Clin (Barc)*. 1995;105:56-8.
19. Hyrkas K, Appelqvist-Scmidlechner K, Paunonen-Ilmonen M. Translating and validating the Finnish version of the Manchester Clinical Supervision Scale. *Scand J Caring Sci*. 2003;17: 358-64.
20. Sabbah I, Drouby N, Sabbah S, Retel-Rude N, Mercier M. Quality of life un rural and urban population in Lebanon using SF-36 Health Survey. *Health Qual Life Outcomes*. 2003;1:30.