

Manual del Residente de Medicina Familiar y Comunitaria.

Capítulo: La obtención de la información mediante entrevista.

Diseño y validación de cuestionarios.

Autores: J. Barranco Navarro, T. Martínez - Cañavete, O. Solas Gaspar

Introducción

La investigación de los fenómenos en el campo de la salud, y en general en todos los ámbitos, aunque aparentemente sencillos encierran una gran complejidad. Los comportamientos individuales y de los grupos relacionados con la salud y la enfermedad están influenciados por componentes sociales y culturales entre los que destacan hábitos, creencias, valores, actitudes... El conocimiento de estos en la comunidad constituye hoy día una prioridad para el desarrollo y evaluación de la atención primaria (programas de salud, utilización de servicios, satisfacción del usuario, estilos de vida, etc.). Para ello la entrevista en general, como técnica de investigación y el cuestionario como instrumento tienen una especial relevancia metodológica.

Concepto de entrevista, encuesta y cuestionario

Aunque tanto las entrevistas como los cuestionarios y encuestas basan su información en la validez de los informes verbales, existen diferencias entre ambos métodos. En un cuestionario la información que se obtiene se limita a las preguntas escritas que ya han sido preparadas y estructuradas. Mientras que en la entrevista se le permite al investigador una mayor flexibilidad, y ambos, entrevistador y entrevistado, están presentes al tiempo que las preguntas son formuladas y contestadas. A veces es difícil discriminar ambas técnicas, cuando se diseñan cuestionarios que son administrados por los propios investigadores. Los términos son usados con frecuencia en los textos de metodología de investigación como técnica, método o instrumento. Podemos considerar la encuesta como un tipo de diseño de estudio, la entrevista como una técnica de comunicación y el cuestionario como el instrumento, protocolo o formulario utilizado para la recogida de información.

Concepto de entrevista: Se han dado múltiples definiciones, Bingham y Moore la definen como: «Una conversación seria que se propone un fin determinado distinto del simple placer de la conversación.» Para Grawitz la entrevista es: «La concurrencia o conversación de dos o más personas en un lugar determinado para tratar un asunto.» Parece ser que existe un cierto acuerdo en las diferentes definiciones al aceptar los siguientes rasgos: comunicación verbal, guión estructurado, finalidad específica, proceso bidireccional y adopción de roles específicos por ambas partes, entrevistador-entrevistado.

Concepto de encuesta: la encuesta es definida por García Ferrando como «una investigación realizada sobre una muestra de sujetos, representativa de un colectivo más amplio, que se lleva a cabo en el contexto de la vida cotidiana, utilizando procedimientos estandarizados de interrogación, con el fin de obtener mediciones cuantitativas de características objetivas y subjetivas de la población». En definitiva, la encuesta hace referencia al procedimiento de investigación, y el cuestionario, al instrumento o formulario que se aplica a los sujetos.

Cuestionarios: concepto y tipos

Un cuestionario es un instrumento, soporte o esqueleto de la investigación por encuesta, específicamente diseñado para poder cuantificar y universalizar la información de los sujetos, así como controlar la estandarización del procedimiento de realización de la entrevista.

Presentan una serie de características comunes:

- a) No se observan los hechos directamente sino a través de las manifestaciones verbales o escritas de los sujetos investigados.
- b) Permite indagar aspectos objetivos y subjetivos de los fenómenos sociales.
- c) Permite reproducir la recogida de información de forma estandarizada.

El cuestionario es una técnica ampliamente utilizada en la investigación social y actualmente también en ciencias de la salud, por ello toda clasificación restringe la gran variabilidad de este instrumento.

Según la técnica de administración se clasifican en:

- a) Mediante entrevista personal: cuando es administrado por un entrevistador. Esta técnica obtiene una alta tasa de respuestas, pero su coste es elevado y exige un adiestramiento adecuado del entrevistador.
- b) Por correo: cuando es enviado al domicilio del sujeto. Es autoadministrado. Su principal dificultad es la alta tasa de «no respuestas».
- c) Mediante entrevista telefónica: ampliamente utilizada en otros países por su bajo coste y eficacia, en nuestro medio presenta aún dificultades en algunas zonas por la escasa cobertura telefónica. Las preguntas deben ser muy breves y sencillas.

Según sus contenidos pueden referirse a:

- a) Hechos: se indaga sobre acontecimientos ocurridos en los cuales el entrevistado es protagonista principal y se refieren características demográficas, ambientales y conductas manifiestas o aparentes.
- b) Opiniones: se recaba información sobre aspectos subjetivos del sujeto en relación con su estado de opinión sobre el tema.
- c) Actitudes, motivaciones y sentimientos: informaciones referidas a los aspectos que están en la base de la conducta del sujeto.

Ventajas y limitaciones del cuestionario

Entre las ventajas de la utilización del cuestionario respecto a la entrevista sin este instrumento se destacan:

1. El coste de aplicación es menor.
2. Requiere menos habilidades por el entrevistador en el manejo de la relación con el entrevistado.
3. En caso necesario, el anonimato permite garantizar la libertad de respuesta.
4. Puede aplicarse a un gran número de personas.
5. Cuando es enviado por correo, permite cubrir un área geográfica muy extensa.
6. El cuestionario autoadministrado permite al encuestado una mayor reflexión de las preguntas.

Entre las limitaciones del cuestionario se destacan:

1. La estandarización del contenido impide adaptar la formulación de las preguntas al encuestado.

2. No es flexible a la inclusión de nuevos temas aportados por el encuestado, ni permite al entrevistador acceder al tema por vías distintas de la prevista en el cuestionario.
3. Su utilización exige un proceso previo de validación del instrumento.

TIPOS DE PREGUNTAS

- Según su estructura básica:
 - a) Abiertas: permiten al encuestado expresar libremente con sus palabras la respuesta. Su principal dificultad radica en la codificación y análisis posterior de la información.
 - b) Cerradas: las respuestas posibles son especificadas y estandarizadas en el cuestionario. Exige un mayor esfuerzo en la búsqueda de las posibles categorías de respuesta.
- Según su función las preguntas pueden ser utilizadas con fines diversos:
 - a) Preguntas filtro: se utilizan para economizar preguntas y evitar que sea redundante la información. Ejemplo: Pregunta: ¿Ha utilizado el servicio de urgencias en el último año? Respuestas: SI/NO. En el caso de responder SI, ¿cuántas veces?, ¿cuáles fueron los motivos? En caso de responder NO se pasaría a las preguntas del siguiente tema.
 - b) Preguntas de consistencia: detectan la congruencia en las respuestas del sujeto. Para este fin se redactan varias preguntas que varían en la forma, pero no en el contenido de la información que obtienen.
 - c) Preguntas de control. intentan determinar la fiabilidad y veracidad de las respuestas incluyendo respuestas falsas o categorías de respuesta con trampa.
 - d) Preguntas de acceso: También llamadas «muelle» o «colchón», se introducen para abordar temas delicados.
 - e) Preguntas de aflojamiento: También llamadas «introdutorias», son útiles para comenzar el cuestionario o ganarse la confianza del encuestado o pasar de un tema a otro.

EXTENSION Y ORDENACION DE LAS PREGUNTAS

El número de preguntas incluidas en un cuestionario no está sujeto a normas rígidas, sin embargo, debe limitarse en lo posible; algunos autores no recomiendan superar las 25- 30 preguntas. El número es irrelevante si el cuestionario responde a los objetivos de la investigación y está bien construido. Respecto al tiempo necesario para responderlo, no debe superar los 30-40 minutos.

En relación con el orden de colocación de las preguntas en el cuestionario, se utilizarán distintos procedimientos dependiendo del tema que se estudie:

- a) El orden lógico comienza con las preguntas más sencillas para ir aumentando progresivamente su complejidad. El orden psicológico coloca las preguntas más sencillas al principio y al final, dejando la parte central del cuestionario para aquellas más complejas o delicadas.
- b) En batería: se agrupan formando unidades temáticas.

Algunas preguntas pueden tener ciertos efectos que influyen la respuesta en las siguientes (efecto contaminación) o sobre una batería de respuestas (efecto colocación). Para evitarlos puede utilizarse:

- 1) La técnica del embudo: comenzar con las preguntas generales para descender a las más específicas.
- 2) La técnica del embudo invertido: utiliza el sentido inverso.

LENGUAJE Y REDACCION DE LAS PREGUNTAS

El lenguaje y redacción de las preguntas dependerá de la población investigada, sus características socioculturales, edad, sexo, etcétera. No obstante Briones señala una serie de recomendaciones respecto al contenido:

1. Evitar el abuso de preguntas en las que el encuestado tenga poca experiencia o conocimientos. A veces las personas no reconocen su ignorancia y responden a todas las preguntas aunque no estén capacitadas.
2. Evitar preguntas muy generales cuando se pretende obtener información específica o concreta sobre un tema.
3. Cuando las preguntas investiguen actitudes, motivaciones y valores, necesitaremos hacer más de una pregunta para obtener la información desde las distintas perspectivas que pueda contemplar el mismo tema.
4. En los cuestionarios sobre actitudes se deberán contemplar los componentes cognoscitivos, afectivos y comportamentales al desarrollar las preguntas.

En lo que se refiere a la forma de redactar:

1. Las preguntas deben ser claras y sencillas, redactadas en un lenguaje comprensible y de fácil lectura.
2. Su extensión debe ser corta para evitar el aburrimiento y la dispersión.
3. Evitar la formulación de dos preguntas en una.
4. La forma de expresión debe ser «neutral». Evitar palabras cargadas de contenido emocional negativo o positivo, o aquellas que puedan inducir reacciones enojosas o actitudes defensivas. Con este fin es aconsejable hacer un planteamiento indirecto en temas comprometidos como sexualidad, drogas, política, etcétera.

Concepto y tipos de escalas

Las escalas son instrumentos que permiten medir o cuantificar las respuestas que se obtienen en una entrevista.

A partir de estos datos se puede inferir la asistencia de «dimensiones»; estas dimensiones indican la presencia de una determinada característica o concepto. Por ejemplo, el concepto de menopausia podría estar formado por una serie de dimensiones como son la presencia de determinados síntomas, tales como: sudores, palpitaciones, temblores, creencias acerca de este período, emociones, etcétera.

Las escalas proporcionan a los entrevistadores un conjunto de «categorías» que representan una serie de posiciones de un individuo respecto a esta escala. Los investigadores se han esforzado por encontrar procedimientos que posibiliten la colocación de los individuos en la escala sin probabilidades de error. En este sentido se han elaborado cuestionarios estandarizados, donde se mide la respuesta de los sujetos en un solo formato, por ejemplo, tipo Likert.

Existe la posibilidad de medir estas respuestas en un cuestionario que contiene escalas con diferentes formatos, de manera que este cuestionario estará menos estandarizado y ofrecería la posibilidad de obtener información sobre una dimensión más a través de diferentes escalas.

Tipos de escalas

Vamos a comentar tres grandes grupos de escalas: escalas básicas, escalas comparativas y escalas no comparativas.

Escalas básicas Son escalas generales que hacen referencia al tipo de información que se recoge en las mismas. Hay cuatro tipos: las nominales, las ordinales, las de intervalos y las de razón.

Escalas nominales: estas escalas identifican diferentes categorías o alternativas de respuesta mutuamente excluyentes, es decir, no indican ningún orden u otro significado de respuesta. Ejemplo: ¿Cuál de estos síntomas ha experimentado? 1. Temblores. 2. Sudores. 3. Palpitaciones. 4. Dolor de cabeza.

Escalas ordinales: estas escalas, además de presentar las diferentes categorías, los números, indican la existencia de un orden. Ejemplo: ¿Cuántas veces a la semana ha tenido Ud. palpitaciones? 1. Menos de tres veces por semana. 2. De tres a cinco veces. 3. Seis o siete veces. 4. Más de siete veces.

Escalas de intervalos: estas escalas poseen las características de las escalas ordinales, pero la diferencia entre los valores de la escala poseen un significado. Las diferencias pueden compararse. Por ejemplo, la diferencia entre 1 y 2 es la misma que entre 2 y 3. En esta escala la ubicación del cero no es fija.

Ejemplo: ¿Qué probabilidad tiene a partir de ahora de experimentar palpitaciones por las noches? 0.41

1. Seguro que no tengo palpitaciones por la noche.
2. No es probable que tenga palpitaciones por la noche.
3. Es probable que tenga palpitaciones por la noche.
4. Seguro que tengo palpitaciones por la noche.

Escalas de razón: estas escalas poseen las características de las escalas descritas anteriormente, pero tienen un punto cero significativo. Ejemplo: ¿Cuánto dinero ha gastado en medicación la última semana? ptas.

Escalas comparativas Son escalas que facilitan al individuo un punto de referencia a la hora de elaborar un juicio. Se distinguen seis tipos de escalas: de comparaciones pareadas, de clasificación, de suma constante, de Guttman, de clases o similitudes y de protocolos verbales.

De comparaciones pareadas: en estas escalas se hacen tantas combinaciones de pares de elementos como sea posible. Por ejemplo, un profesional sanitario está interesado en conocer la presencia de un determinado síntoma en el problema de la menopausia. En concreto cuatro síntomas: A (temblores), B (sudores), C (palpitaciones), D (dolor de cabeza). Al utilizar esta escala se procede de la siguiente manera:

«Dígame de este par de síntomas A (temblores) y B (sudores) cuál tiene tendencia a padecer más.

» Después de elegir se le vuelve a presentar el síntoma A (temblores) para que lo compare con el C (palpitaciones), de forma que elija entre estos dos; posteriormente hará lo mismo con los síntomas A y D; B y C; B y D y C y D.

Escalas de Guttman: se basan en la posibilidad de ordenar un conjunto de ítems o incluso de individuos en base a una sola característica o atributo, de forma que se presentan los estímulos de más sencillo a más complejo, pudiendo contestarse a los mismos de forma dicotómica. Las respuestas que se obtienen implican que el valor alcanzado en la escala nos asegura que el entrevistado ha sido capaz de alcanzar todos los valores anteriores (menores) aunque no es capaz de contestar ninguno de los superiores al contestado.

En la menopausia, las mujeres no deberíamos estar preocupadas por los cambios que se producen biológica y psicológicamente:

De acuerdo
Desacuerdo

En la menopausia, las mujeres deberíamos preocuparnos sólo cuando se produzcan alteraciones importantes a nivel biológico en nuestro organismo:

De acuerdo
Desacuerdo

En la menopausia, las mujeres deberíamos preocuparnos sólo cuando se produzcan alteraciones importantes a nivel biológico y psicológico:

De acuerdo
Desacuerdo

En la menopausia, las mujeres deberíamos preocuparnos de cualquiera de los cambios biológicos o psicológicos que se produzcan, aunque no sean muy importantes:

De acuerdo
Desacuerdo

Escalas de clasificación: en estas escalas se le pide a los sujetos que clasifiquen u ordenen una serie de elementos en base a algún atributo o variable.

Ejemplo: Por favor, clasifique los siguientes síntomas, de mayor a menor, según la gravedad que Ud. los padece, asignando un 1 al más grave, 2 al siguiente y así sucesivamente:

Temblores 2.
Sudores 1.
Palpitaciones 4.
Dolor de cabeza 3.

Escalas de suma constante: en estas escalas se le pide al entrevistado que reparta una cantidad determinada (generalmente 100) entre un conjunto de elementos que le vienen dados.

Ejemplo: Por favor, reparta 100 puntos entre los síntomas que le presentamos, de forma que reflejen la gravedad que tienen para Ud:

Temblores 15.
Sudores 35.
Palpitaciones 40.
Dolor de cabeza 10.

Escalas de clases o similitudes: en estas escalas se les pide a los individuos que clasifiquen los estímulos en un número limitado de grupos (fijados por el investigador) atendiendo a la similitud de los grupos.

Ejemplo: De los siguientes síntomas que le vamos a mostrar, clasifique éstos en alguno de los siguientes grupos:

Menopausia Depresión Gripe

En este ejemplo, el investigador le muestra al entrevistado un grupo de X tarjetas con el nombre de diferentes síntomas, que ha de agrupar atendiendo a su similitud dentro de un grupo.

Protocolos verbales: en esta escala se le pide al entrevistado que seleccione aquella frase o conjunto de palabras que mejor representen su posición con respecto al estímulo que se le plantea. Lo más aceptado es plantear 5 ó 6 alternativas.

Ejemplo: ¿Qué opina de esta nueva etapa en relación con su etapa anterior?

- Es muchísimo mejor.
- Es mejor.
- Es más o menos igual.
- Es peor.
- Es muchísimo peor.

Escalas no comparativas

Son escalas donde las contestaciones a los distintos enunciados no se basan en acciones de comparación ante los estímulos presentados. Este tipo de escalas suelen usarse para medir las actitudes y valoraciones personales. Existen cinco tipos: de clasificación, continua, Likert, diferencial semántico, Stapel, Thurstone.

Vamos a comentar las escalas de clasificación continua. La escala Likert se comenta en otro apartado de este capítulo.

De clasificación continua: en estas escalas se presentan infinitas alternativas de respuesta. El individuo ha de marcar sobre una línea recta su posición respecto a la variable que es tema de investigación. Ejemplo:

Indique, por favor, marcando en esta línea cuál es su opinión respecto a la gravedad de los síntomas experimentados en la menopausia.

Nada graves

Muy graves 1 2 3 4 5 6 7 8 9 10

Cuestiones a tener en cuenta en la creación de las escalas

- ☞ Utilizar un formato de tipo gráfico que permita medir y cuantificar lo mejor posible las puntuaciones de los sujetos.
- ☞ Utilizar escalas comparativas o escalas no comparativas. Una escala comparativa ofrece una referencia a partir de la cual se ha de evaluar un objeto. El problema aparece cuando este punto de referencia no tiene significado para el sujeto. En la escala no comparativa esta referencia no se ofrece. Antes de utilizar una escala hemos de considerar las características del individuo del que se quiere obtener información.
- ☞ Utilizar respuestas forzadas o no forzadas, con el sentido de ofrecer al individuo la oportunidad de ocupar una posición neutral. Aunque no existe un acuerdo es conveniente dar a los sujetos la oportunidad de que ocupen una posición neutral. En estos casos también conviene tener en cuenta el tipo de individuos y la formación que se desea obtener.
- ☞ Utilizar categorías de respuesta equilibrada o no equilibrada, en relación al número de alternativas positivas o negativas. Cuando no se sabe a priori cuál es la actitud de un sujeto respecto al tema que se analiza, es mejor utilizar de respuesta equilibrada.
- ☞ Utilizar muchas o pocas alternativas de respuesta. No hay regla fija. Al utilizar muchas alternativas se puede conseguir una respuesta más precisa. En otras ocasiones utilizar escalas con muchas alternativas no tiene sentido, depende del tipo de variable y de las características del individuo. Por ejemplo, en las escalas tipo Likert se pueden obtener resultados más representativos cuando se emplean cinco posiciones en lugar de tres.

Escalas de actitudes. La escala de Likert

La actitud se define como una disposición mental que orienta nuestras acciones; es también una respuesta afectivo - emotiva y una predisposición a actuar en sentido positivo o negativo.

Está formada por tres componentes:

- a) Cognoscitivo: se refiere al grado de conocimiento, percepción, opinión y pensamiento que tiene el individuo hacia el tema que se investiga.
- b) Afectivo: es el sentimiento a favor o en contra del tema.
- c) Conductual: es la tendencia a reaccionar frente al tema de una determinada manera.

Las escalas son los instrumentos metodológicos más utilizados para estudiar actitudes sociales, de opinión, estilos de vida, etcétera. La más conocida y aplicada en nuestro medio es la Escala de Likert. Parte del supuesto teórico de que una actitud ha de ser definida como disposiciones de los sujetos hacia una acción manifiesta. Las declaraciones de opinión son un método indirecto para medir una actitud. La elaboración de preparación de los ítems consiste en recoger los enunciados que se refieren a la actitud que se pretende medir y que hagan referencia a comportamientos deseados. Se enuncian de forma categórica, con un lenguaje claro, conciso y formulados de manera que la respuesta modal se sitúe en el centro de la escala, evitando asimetrías. Las preguntas se presentan en un cuestionario de modo que cada una de ellas tenga adscrita una escala gradual con varias alternativas posibles de respuesta. Ejemplo: absolutamente de acuerdo, de acuerdo, indiferente, en desacuerdo, absolutamente en desacuerdo.

Concepto y tipos de sesgos

Entendemos por sesgos las «fuentes potenciales de error» en el diseño y aplicación de la técnica de entrevista, que dificultan la validez y fiabilidad de las respuestas. Se clasifican según sean debidos a la influencia de] entrevistador, del entrevistado o del instrumento - cuestionario.

Influencia del entrevistador

- a) Condicionamiento verbal como la entonación, el volumen, etcétera, de la voz, y no verbal como gestos, miradas, actitud física, postura, etcétera, pueden expresar al encuestado actitudes positivas o negativas del entrevistador respecto a la pregunta que se formula.
- b) Personalidad del entrevistador, aspecto físico, edad, sexo, etcétera, pueden modificar las respuestas del encuestado.
- c) Fenómeno de sugestión: el sujeto responde lo que piensa que el entrevistador desea obtener: «sesgo de expectativas».

Los sesgos debidos al entrevistador se evitarán con una adecuada formación y adiestramiento de los entrevistadores.

Influencia del entrevistado

- a) Juicios basados en apariencias: los entrevistadores pueden enjuiciar rasgos psíquicos o actitudes del entrevistado.
- b) Efecto de halo: producido por la impresión general que el entrevistado produce sobre el entrevistador.

- c) Efecto prestigio: debido a que las preguntas no están bien formuladas y el encuestado no responde a la verdad sino a lo que considera que el investigador espera como respuesta. Es equivalente al apartado c) anterior.
- d) Efecto de aprendizaje: se produce como consecuencia del conocimiento mutuo que van adquiriendo encuestador y encuestado previamente o en el transcurso de la entrevista. Para evitar este sesgo y el anterior, se recomienda que la encuesta sea administrada por personas desconocidas y ajenas al tema de investigación, por ejemplo, evitar que sea su médico o enfermera.

Influencia del instrumento-cuestionario

- a) El orden de colocación de las preguntas y respuestas: puede generar un efecto de falta de coherencia, desorden o desconcierto e influir negativamente en el recuerdo de los hechos, en el interés y en la sinceridad de las respuestas.
- b) Efecto de cadencia: el cuestionario internamente debe seguir un ritmo que facilite la comunicación fluida entre el entrevistador y el entrevistado.
- c) Efecto «las no respuestas»: para cada pregunta obtendremos una proporción de respuestas y no respuestas. Debemos investigar las posibles causas de esas «no respuestas» (ambientales, de confidencialidad, falta de comprensión de la pregunta, desacuerdos con el enunciado, etcétera). Los acontecimientos que se produzcan en el entorno social inmediato (presencia del acompañante, cuidador, padres, etcétera) del entrevistado en relación con el tema que se aborda deben ser cuidados por el entrevistador para evitar que mediaten sus respuestas.

Validez y fiabilidad. Concepto. Tipos. Técnicas de control

Por validez entendemos el grado en que un instrumento mide realmente aquello que pretende medir. La fiabilidad comprueba que el instrumento está midiendo algo de manera estable y reproducible. Para comprender la diferencia imaginemos que hemos diseñado un nuevo cuestionario para conocer la satisfacción de los pacientes con su médico de familia. Encontramos que el grado de satisfacción es similar cuando se aplica el cuestionario en dos ocasiones diferentes y que se han alcanzado resultados similares cuando el cuestionario ha sido aplicado al mismo paciente por dos entrevistadores diferentes. Podemos afirmar que el cuestionario es fiable. Sin embargo, aún no podemos asegurar que las diferencias encontradas respondan a diferentes grados de «satisfacción con el médico» y no a la repercusión del modelo de organización del centro de salud, por ejemplo el horario de consultas o la organización de la cita previa, etcétera. Para ello estudiaremos la validez del cuestionario para medir realmente la «satisfacción con el médico». De forma genérica se habla de «validar un cuestionario» cuando procedemos al estudio de su validez y fiabilidad.

Validez

En ciencias de la salud se utilizan con frecuencia cuestionarios creados por nosotros o por otros investigadores sin comprobar el proceso de validación seguido por éstos. Antes de utilizar cualquier instrumento es necesario comprobar quién realizó la validación, cómo y cuáles fueron los resultados obtenidos. Una serie de preguntas nos surgen al plantear este tema: ¿Por qué es necesario validar un cuestionario? ¿Cuándo es necesario validar un cuestionario? ¿Cómo se realiza el proceso de validación?

¿Por qué es necesario validar un cuestionario?

Algunas variables pueden ser fácilmente objetivables y medibles: es el caso de la temperatura, el pH o la presión arterial. La correcta utilización de los instrumentos de medida, junto a la comprobación de su fiabilidad (estabilidad), sensibilidad y especificidad, nos garantizan su validez. No ocurre lo mismo cuando la variable que

queremos medir es subjetiva, como es el caso de la «satisfacción del usuario con su médico» o la «calidad de vida». En este caso la validez dependerá de la definición que elaboremos sobre este concepto, de la capacidad para operacionalizar (hacer medibles) los elementos que la componen y de la relación que exista entre dichos elementos y con el resultado final obtenido.

¿Cuándo es necesario validar un cuestionario?

La respuesta inmediata es: «Siempre». ¿Cómo podremos si no afirmar que el 80% de los pacientes están satisfechos con su médico, si en realidad no podemos asegurar que las preguntas realizadas estén midiendo realmente «satisfacción con el médico» y no otras variables de la organización?

La importancia del estudio, su trascendencia (inferencias) y la repercusión de los resultados en la toma de decisiones nos permitirán valorar la necesidad de utilizar un instrumento validado.

Como principio básico buscaremos instrumentos validados ya por otros investigadores. Pero además es necesario comprobar que se haya realizado en el mismo idioma y sobre un grupo de población de características similares. La traducción es una de las mayores fuentes de problemas. Los giros lingüísticos y patrones culturales modifican en muchas ocasiones el sentido de las preguntas, haciendo perder la validez al instrumento.

Pueden darse varias circunstancias:

1. Cuestionario validado en castellano sobre una población de similares características. En este caso puede ser utilizado sin problema, sin olvidar solicitar autorización a los autores.
2. Cuestionario validado en otro país o sobre una población de características diferentes. Por ejemplo, ancianos - adolescentes, nivel sociocultural bajo-profesionales sanitarios, etcétera. Será necesario realizar una traducción bidireccional y proceder a su validación posterior.
3. Utilización «parcial» de un cuestionario validado. Consultando el proceso de validación, comprobaremos si el conjunto de preguntas o ítems utilizado corresponde a una dimensión teórica y real, en cuyo caso puede ser utilizado. En caso de seleccionar ítems aislados dentro de otro cuestionario se procederá de nuevo a su validación. Se trata de un nuevo instrumento.

¿Cómo se valida un cuestionario?

La validación es un proceso mediante el cual realizamos una serie de pruebas que nos informaran si las cualidades de un cuestionario permitirán la inferencia de sus resultados sobre la población y con qué niveles de confianza. Estudiaremos tres aspectos básicos de la validez, sin olvidar que se trata de un proceso continuado: validez de contenido, validez de constructo y validez de criterio.

Validez de contenido

También llamada validez lógica o «de sentido común». Es una comprobación cualitativa, no estadística, sobre la capacidad que tiene el conjunto de variables para definir una variable compleja. En un estudio sobre «satisfacción del usuario con los servicios de salud», se investigarán aspectos de la atención recibida, de las preferencias y de las expectativas del usuario a través de las siguientes dimensiones: competencia profesional, relación médico-paciente, calidad técnica, accesibilidad, comodidad, continuidad, imagen externa del servicio. Este conjunto de dimensiones

pretenden abarcar el mayor número de factores que puedan explicar teóricamente el concepto de «satisfacción».

Para seleccionar estas dimensiones partiremos de los objetivos del estudio y revisión cuidadosa de la bibliografía en la búsqueda de posibles instrumentos ya validados por otros autores. En caso de no encontrar ninguno que responda a nuestras necesidades se recomienda consultar grupos de expertos utilizando técnicas cualitativas de discusión y consenso. Estadísticamente, el Análisis Factorial de Componentes principales es la técnica que permite estudiar el grado de correlación entre las variables de cada dimensión teórica.

Validez de constructo

Para observar las variables de interés es necesario hacerlas «operativas» o «medibles». Frecuentemente la variable que pretendemos medir no es posible observarla y operativizarla (medirla) tan fácilmente como el peso, la talla o la presión arterial. Se trata de variables complejas generalmente sobre actitudes, conocimientos o aspectos psicológicos que requieren para su completa definición la construcción de una «teoría» que incorpora diferentes aspectos relacionados. Siguiendo con el ejemplo anterior, la «satisfacción» no es posible medirla físicamente, el «constructo» de este concepto estará formado por el conjunto de aspectos, relacionados entre sí, que podrían explicar este concepto (competencia, relación médico -paciente, accesibilidad, etcétera).

Existen dos aspectos a estudiar en la validez de constructo:

Validez convergente y validez discriminante

Consistiría en ver «cuánto» de correlacionadas están entre sí las variables que componen el constructo. En nuestro ejemplo los resultados de la variable «competencia profesional» debería correlacionarse estrechamente con la «satisfacción», lo que demostraría su validez convergente. Sin embargo, esta correlación tampoco debe ser muy elevada, porque en ese caso nos haría pensar que están midiendo la misma cosa, es decir, que la «competencia» es equiparable a la «satisfacción», algo que teóricamente no aceptamos. Pero además estas variables no deberían correlacionarse con aspectos teóricamente no relacionados como la «ubicación del archivo de historias clínicas», por ejemplo. En esta ocasión hablamos de validez discriminante.

Método de los grupos extremos

Consiste en pasar el cuestionario a dos grupos. Uno que presenta la característica y otro que no la presenta. Si en el primer grupo se obtienen puntuaciones significativamente más altas en la escala que en el segundo grupo, entendemos que el cuestionario es capaz de detectar esa diferencia. La principal dificultad del método consiste en la necesidad de conocer la presencia / ausencia de la característica a priori; para seleccionar los grupos, lo que implica que contamos con otro instrumento válido de observación, que no siempre existe.

Validez de criterio

Se basa en el hallazgo de una correlación positiva entre los resultados obtenidos con el nuevo cuestionario y otro instrumento utilizado como referencia al que llamamos «patrón oro» y que es aceptado como válido. Llamaremos a esta comprobación validez concurrente. Como ejemplo podemos medir la frecuencia y motivos de utilización de la consulta por los usuarios a través de su respuesta a un cuestionario y consultando la historia clínica y registros de cita previa («patrón oro»). En ocasiones, no contamos con esta medida de referencia válida y fiable. En este caso, la medida óptima a tomar como referencia es aquella que tiene mayor validez lógica, o que ha sido probada previamente, demostrando tener una alta validez de criterio. Cuando no existe esta posibilidad, nuestro cuestionario sería considerado el más válido. En otras ocasiones, el criterio de comparación se obtendrá paralelamente o incluso *a posteriori*. Un ejemplo serían los test de screening, donde el criterio de comparación utilizado sería el resultado de los diagnósticos definitivos. A esta la llamaremos validez predictiva.

El cálculo de la correlación se realizará mediante procedimientos estadísticos.

Fiabilidad

Su objetivo es comprobar que el instrumento está midiendo algo de forma precisa, estable y reproducible, expresada por la capacidad de una medida para dar resultados similares al ser aplicado en dos o más ocasiones a un mismo grupo de sujetos (reproductibilidad) a una misma persona en momentos diferentes (estabilidad), o a una sola persona frente en momentos diferentes (estabilidad), o a una sola persona frente a la situación real observada (precisión).

Cuando nos referimos a un instrumento de laboratorio, fiabilidad es sinónimo de «precisión» y mide la variabilidad de una medición individual expresada como la desviación estándar (DE) de ese valor respecto a la media. En general estudia el grado de acuerdo o desacuerdo entre observadores respecto a la presencia o ausencia de una característica es fácil en variables dicotómicas, sin embargo conforme aumentan el número de categorías, disminuye la probabilidad de acuerdo. Por esta razón se utilizará como medida el «grado de asociación entre observadores» en cada una de las posiciones de la escala.

Existen varios métodos para el estudio de la fiabilidad:

1. Método con observadores. Caso de cuestionarios administrados por entrevistadores:
 - a) Coeficiente de fiabilidad inter-observador. Mide las variaciones entre los resultados obtenidos por varios entrevistadores sobre el mismo sujeto. Puede ocurrir que las diferencias encontradas se deban a diferencias entre los propios entrevistadores y no al instrumento de medida.
 - b) Coeficiente de fiabilidad intra-observador: mide las diferencias entre varias medidas realizadas por un mismo entrevistador sobre el mismo sujeto.
2. Método para cuestionarios autoadministrados:

Test-retest: el cuestionario se aplica en dos ocasiones, con un intervalo de tiempo entre ambos, a los mismos sujetos, y se comparan los resultados obtenidos. La correlación entre ambas expresará el grado de fiabilidad del instrumento. Estadísticamente, se utilizarán diferentes técnicas de medida según el tipo de variable:

 - a) Para variables continuas se aplicará el Análisis de la Varianza (ANOVA), que estima la razón de la varianza entre sujetos sobre el total de la varianza (suma de la varianza entre sujetos más el error de la varianza). Su resultado, llamado Coeficiente de fiabilidad, varía de 0 (no reproductibilidad) a 1 (máxima reproductibilidad).

- b) Para variables categóricas ordinales se utilizará el «Coeficiente Kappa (K)», cuyo valor mínimo deseable se sitúa en torno a 0,8.
- c) Para variables nominales se aplica- rá una variación del anterior, el «Coeficiente Cohens Kappa».

Fases en la construcción de un cuestionario

El cuestionario ha de estar enmarcado en las distintas fases de la metodología de investigación científica:

1. Definición del objeto de estudio.
2. Revisión y análisis de las diversas aportaciones teóricas al tema.
3. Objetivos específicos: Concretamos el objetivo general en cuestiones más particularizadas.
4. Definir la población y muestra: definimos el universo poblacional y seleccionamos una muestra de esta población, indicando las características demográficas (edad, sexo, nivel educativo... y tamaño de la muestra) y la dimensión temporal M estudio.
5. Formulación de hipótesis y operacionalización de las variables: formulamos la hipótesis de trabajo que será objeto de verificación o refutación y operacional izamos las variables, es decir, convertimos las abstracciones para que puedan* ser medidas. Cuan- do el estudio es descriptivo no se formulan hipótesis a contrastar.
6. Elección de la técnica de recogida de información y elaboración del instrumento: la elección de cuestionario o encuesta está sujeta a una serie de criterios ya mencionados: a) El tiempo de instrumento a utilizar dependerá: 1. De nuestros fines científicos: descriptivos o explicativos. 2. De la forma de administración: entrevista personal, por correo o encuesta telefónica. 3. Del contenido: puede referirse a hechos, opiniones, actitudes, motivaciones y sentimientos, y la dimensión temporal o aplicación en el tiempo. b) Hemos de elegir la forma y tipos de preguntas. Las clasificaciones más frecuentes de preguntas son: 1. Preguntas abiertas o cerradas. 2. Por su naturaleza, puede referirse a hechos, acción, información, intención, opinión. 3. Por su función son las preguntas filtros, de consistencia, de control, de acceso y afloja- miento. c) Hemos de tener en cuenta las reglas y recomendaciones que se deben seguir para la formulación de las preguntas. d) Extensión y ordenación del cuestionario.
7. Realización del pre - test o estudio previo y presentación del cuestionario: una vez delimitada la información, el tipo de preguntas, la formulación de las mismas y la ordenación, que dependerá de nuestro objetivo de investigación, se procede a la «prueba piloto» para comprobar el correcto funcionamiento del cuestionario. Esta prueba se administra a un grupo de personas (30 a 50) que no tienen representatividad, pero que se «aproxima» a las características de la muestra definitiva. La prueba nos indica: los tipos de preguntas más idóneos, la comprensión de las preguntas, su sencillez y extensión, el adecuado funcionamiento de las categorizaciones y abanicos de respuesta, las resistencias psicológicas, la fatiga del entrevistado y el ordenamiento interno de las preguntas.
8. Elaboración del cuestionario definitivo: cuando se ha comprobado el buen funcionamiento del cuestionario se procede a su confección definitiva. Este cuestionario se pasará a la muestra representativa de la población objeto de estudio. Una adecuada presentación del cuestionario requiere: a) La correcta identificación del organismo, empresa o institución que lleva la investigación. b) El título del estudio. c) Garantizar el anonimato. d) Registrarse la fecha del período de la investigación. e) Indicar la fecha en que ha sido cumplimentando el cuestionario.
9. Formación de los encuestadores: tiene por objeto disminuir en lo posible los sesgos debidos al entrevistador que ya fueron comentados. Pelecha- no sugiere tres condiciones que han de cumplir estos programas de formación: a) Ofrecer información sobre las técnicas de recogida de datos. b) Enseñar las técnicas de

- entrevista a utilizar en cada caso. c) Realizar demostraciones prácticas de las entrevistas.
10. Codificación y tabulación: la codificación consiste en un procedimiento técnico en el que los datos del cuestionario son transformados en símbolos ordinariamente numéricos. Para ello se recurre a la adscripción de valores numéricos marginales a cada categoría de respuesta. La tabulación forma parte del proceso estadístico del análisis de datos. Puede ser manual o mecánica.
 11. Análisis de resultados: se aplican las técnicas estadísticas adecuadas. Se comparan los resultados con la hipótesis de trabajo y se establecen conclusiones relacionadas con la teoría y los resultados de otras investigaciones.

Bibliografía

1. Grawittz M. Métodos y técnicas en las ciencias sociales. Hispano Europea, Barcelona, 1975.
2. Ander Egg E. Introducción a las técnicas de investigación social. Humanitas, Buenos Aires, 1971.
3. Argimón Palias JM, Jiménez Villa J. Métodos de investigación aplicados a la atención primaria. Doyma, Barcelona, 1991: 135-148.
4. Selitiz C et al. Métodos de investigación en las relaciones sociales. Rialp, Madrid, 1980: 399-452.
5. S. Goode Wj, Hatt PK. Métodos de investigación social. Trillas, México, 1970.
6. Briones G. Métodos y técnicas de investigación en las ciencias sociales. Trillas, México, 1982.
7. Ortega E. Manual de investigación comercial. Pirámide, Madrid, 1981.
8. Morales Domínguez, JE Metodología y Teoría de la Psicología. UNED, Madrid, 1985.
9. García Ferrando M et al. En Análisis de la realidad social. Métodos y Técnicas de Investigación. Alianza Universidad, Madrid, 1986.
10. Summers GE Medición de actitudes. Trifias, México, 1976.
11. Festinger L y Katz D. Los métodos de investigación en las ciencias sociales. Paidós, Buenos Aires, 1972.
12. Padua J. Técnicas de investigación aplicadas a las ciencias sociales, FCE, México, 1979.
13. Streiner DL, Norman GR. Heath measurements scales. I-N, Oxford University Press, 1989.
14. Grau G. Metodología para la validación de cuestionarios. MEDIFAM, 1995; 5(6):351-359.