

UN ESQUEMA CONCEPTUAL PARA ANALIZAR LA VALIDEZ EN LAS INVESTIGACIONES MEDIANTE ENCUESTA

Andrés González Gómez
José Luís Padilla García
Universidad de Granada

RESUMEN

Los controles habituales sobre la calidad de la información aportada por las encuestas son insuficientes. De hecho, sólo pretenden en la mayoría de los casos convencer de la idoneidad del muestreo realizado. Sin embargo, la utilidad de la información obtenida con la encuesta depende de más inferencias que la simple semejanza entre la muestra de personas encuestadas y la población objetivo. El trabajo propone un esquema conceptual para analizar de forma global la validez en las investigaciones mediante encuestas. La base de la propuesta es considerar cualquier utilización de la encuesta –teórica o aplicada– como una investigación cuasi-experimental. Esta consideración permite anticipar el impacto de los errores propios de las encuestas sobre la interpretación deseada de la información y analizarlos dentro de los esquemas de validación pensados para la investigación cuasi-experimental. La propuesta se concreta en una tabla donde se cruzan las inferencias propias de la investigación mediante encuesta con los errores más comunes. De esta forma el usuario de la encuesta puede identificar el contenido de las amenazas a la validez por las que se debe interesar y de las que dependerá la utilidad de la información obtenida mediante la encuesta.

Palabras clave: validez, encuesta, diseños de investigación, errores de investigación

Los investigadores y cualquier persona atenta a los fenómenos sociales han tenido en los últimos años la oportunidad de presenciar el "encubramiento" de la encuesta como estrategia de investigación social (Alós, 1996, 1997, 1998). Investigaciones o estudios aplicados con encuestas son realizados en una enorme variedad de ámbitos: barómetros de opinión, encuestas de intención de voto, estudios de mercado, evaluación de programas, etc. Todos los ámbitos de aplicación tienen en común el deseo de obtener información válida que apoye posteriores decisiones.

Desde la ocupación y la preocupación por la metodología, junto con la alegría por las oportunidades profesionales que el interés por las encuestas ofrece, al mismo tiempo no puede dejar de sorprender la escasa atención prestada a la calidad de la información obtenida con la encuesta. La práctica totalidad de las aplicaciones se limitan a informar sólo del error muestral, del tamaño de la muestra y del procedimiento de muestreo utilizado. Por ejemplo, Bosch y Torrente (1993) reflejan esta visión al plantear el problema de la calidad de los datos sólo en términos de la exactitud con la que los datos de la encuesta reflejan los que se obtendrían en la población; visión que parece suponer que las medidas poblacionales se podrían obtener sin error. Por supuesto, que esta información es relevante, pero limita el examen de las amenazas para la interpretación de los resultados a la correspondencia entre los datos muestrales y la población objetivo. Hay muchas otras amenazas. Llamar la atención sobre ellas y ofrecer un marco conceptual útil para su tratamiento es el objeto de este trabajo.

Aunque con una repercusión limitada, las dudas sobre la calidad de la información aportada por la encuesta han sido expuestas en diversos estudios. Belson (1986) enumeró las principales causas de error en la obtención de los datos mediante encuestas. Por parte de los encuestados: los errores en la comprensión de la intención prevista de las preguntas; la falta de esfuerzo o interés, los problemas de memoria o comprensión durante la situación "artificial" de la entrevista; y por parte del entrevistador: tendencia a cambiar la redacción de las preguntas, errores en los procedimientos de presentación, etc. Groves (1989) elaboró una exhaustiva clasificación de los errores en la encuesta dividiendo las fuentes de error en "errores de no-observación" y "errores de observación", dentro de cada una se mencionan desde los debidos a la no-respuesta, el muestreo, etc., hasta los achacables al entrevistador o el instrumento. Foddy (1993) comienza su monografía presentando una extensa relación de ejemplos que ilustran los problemas más frecuentes en la elaboración de las preguntas de la encuesta.

El propio diseño de las preguntas del cuestionario plantea serios interrogantes sobre la calidad de la información. Las encuestas de intención de voto ofrecen un ejemplo oportuno para ilustrar la dificultad de que los encuestados interpreten las preguntas como pretende el investigador. Estas encuestas suelen incluir preguntas para averiguar si los encuestados recuerdan a qué partido votaron en las últimas elecciones (Encuestas Electorales del CIS, <http://www.cis.es>). Los análisis de las

respuestas a esta pregunta revelan diferencias importantes entre los porcentajes de voto registrado realmente en las elecciones y los obtenidos a partir del recuerdo de los encuestados. La cuestión planteada por este tipo de preguntas es interesante por diversos motivos. Basta considerar que 1) disponemos del verdadero valor de la variable —el voto obtenido por cada partido—; 2) el comportamiento registrado parece significativo y en principio fácil de recordar por los encuestados; y 3) ...sin embargo, no se observa un patrón sistemático de variación en las respuestas respecto, por ejemplo, al tiempo transcurrido desde las elecciones. Por tanto, ¿qué están midiendo realmente estas preguntas?

Dudas e interrogantes como los anteriores no parecen haber despertado el interés de los usuarios de las encuestas. Por una parte, los manuales dedicados a la encuesta o al cuestionario mantienen una concepción desfasada de la validez. La visión actual sobre la validez defiende su visión como un concepto único y donde el significado del constructo objeto de la medición juega un papel primordial (e. g., Messick, 1995). Basten algunas referencias como muestra. Litwin (1995) tras afirmar que un cuestionario es válido si mide lo que se ha establecido que mida, propone analizar su validez desde cuatro tipos de validez: aparente, contenido, criterio y constructo. Aiken (1997) aunque adopta una concepción más actual sobre el tema al exponer que la validez no es una característica general del cuestionario sino que varía cuando lo hace el objetivo para el que es utilizado, no puede desprenderse de inercias anteriores y de nuevo recomienda su estudio desde una visión "tripartita" de la validez: contenido, criterio y constructo. Y por otra, un acercamiento meramente cuantitativo a la ocupación de los investigadores por la cuestión de la validez ofrece resultados desalentadores. El análisis de las referencias que aparecen en la base PsycLit desde el año 1996 hasta la actualidad muestra que bajo los identificadores "validity" y "survey" aparecen sólo 15 registros, de los que 11 corresponden al uso restringido del término "survey" para referirse a instrumentos o cuestionarios específicos. El resto de referencias plantean un análisis de la validez limitado a la investigación realizada y coincidente al entender la validez exclusivamente como correlación entre distintas medidas. Por otro lado, las 27 referencias de la base Sociofile para el mismo intervalo temporal reflejan una situación similar: la mayoría presentan estudios de validación sobre una encuesta o cuestionario específico; aunque trabajos como el de Scherpenzeel y Saris (1997) apuntan hacia un acercamiento más global a la validez de la información obtenida mediante la encuesta.

Por último, el deseo de muchos investigadores de ir más allá de la mera descripción de variables, para intentar establecer relaciones de causalidad o al menos de variación conjunta entre las mismas, hace imprescindible abordar la validez de la investigación mediante encuestas desde una perspectiva global. Perspectiva que se ha de concretar en un esquema que oriente al investigador o responsable del estudio en el análisis de las amenazas a la validez de la interpretación que pretende hacer de los resultados de la encuesta.

La finalidad del presente trabajo es proponer un esquema conceptual para el análisis global de las investigaciones mediante encuesta. Abordar el estudio de la validez desde una perspectiva global aconseja agrupar las inferencias por las que se debe interesar al investigador en categorías que le permitan anticipar el efecto de los errores más comunes en la investigación mediante encuesta. Para ello, se desarrollan los siguientes objetivos: 1) completar, para el campo específico de la encuesta, las distintas amenazas a la validez propuestas por Cook y Campbell (1979) con las aportaciones de las concepciones más recientes sobre la validez (e. g., Cronbach, 1982; González y Padilla, 1998); y 2) señalar hacia qué tipo de inferencia pueden resultar una amenaza los errores más comunes en las encuestas dentro del esquema para analizar la validez de las investigaciones cuasi-experimentales.

La encuesta como técnica de investigación

La propuesta para una análisis global de la validez en las investigaciones mediante encuesta requiere dos aclaraciones: una respecto al significado atribuido al término "encuesta", y otra sobre su alcance metodológico. No es necesaria una revisión bibliográfica exhaustiva para detectar la confusión terminológica que arrastra el término "encuesta", podría bastar con un "examen de conciencia metodológica". Aiken (1997) tras advertir que pretende hablar en sentido estricto, enumera los contenidos ideales de un cuestionario como un conjunto de preguntas sobre los hechos, opiniones, actitudes u otras características estudiadas; observadas a través de un conjunto diverso de ítems objetivos o semi-objetivos, respuestas cortas, elección múltiple, tareas de ordenación, preguntas abiertas y cerradas, etc., a todo lo cual se podría denominar "inventario", "test", "escala", "encuesta", "estudio", etc.

La utilización que se hace en el trabajo del término encuesta es semejante a la realizada por Martínez (1995) bajo la etiqueta de "encuesta por muestreo" para referirse a la recogida sistemática de datos de poblaciones o de muestras poblacionales por medio de entrevistas personales u otros instrumentos de recogida. Por su parte, Cochran (1977) Abramson (1990) (citados en Martínez, 1995) y Gómez (1990) proponen una tipología de las encuestas dividiéndolas en "descriptivas" y "analíticas" que nos permite aclarar el alcance metodológico que asignamos a la encuesta.

Esta aclaración parte de la utilización del método científico como marco de generación y validación del conocimiento, tal como propone Gómez (1990). Este planteamiento lleva consigo la consideración de cualquier investigación como un todo, en el que el resultado final, las conclusiones que se establezcan sobre las hipótesis, dependen del conjunto de elementos que participan en la investigación (Bunge, 1976). Aún dentro de esa visión global del método científico, es posible distinguir lo que Arnau (1990) llama niveles de estructuración: teórico-conceptual, técnico-metodológico, y estadístico-analítico. El nivel técnico-metodológico es la parte "procedimental" aquella en que se realizan las acciones necesarias (i. e., el diseño) y se reco-

gen los datos pertinentes para la comprobación de las hipótesis planteadas en el nivel teórico como respuesta al problema que da origen a la investigación. La encuesta en sí, se identifica en gran medida con el nivel técnico. Por tanto, debe ser valorada como una técnica de recogida de los datos.

Desde esta perspectiva, el análisis de la "validez de la encuesta" es un análisis limitado; igual de limitado que abordar la validez de cualquier instrumento de medición, sea un test o todo un cuestionario de una encuesta. La expresión "validez de la encuesta" supondría analizar exclusivamente la validez en el nivel técnico-metodológico, aproximación errónea ya que la concepción general del método científico obliga a tener en cuenta los tres niveles al analizar la validez de una investigación.

Sin embargo, plantear el problema en términos de "la validez de la investigación" resulta poco útil por ser demasiado general. La propuesta de este trabajo acota el problema dos formas. Primero, adscribiendo las investigaciones con encuesta a la metodología cuasi-experimental. Podría parecer que esta adscripción incumple el criterio expuesto por autores como Fowler (1993) de que en la encuesta no haya manipulación ni intervención por parte del investigador. Sin embargo, los objetivos habituales de las encuestas analíticas (Martínez, 1995a): comparación entre subgrupos de población, predicción de comportamientos socialmente relevantes, análisis de la covariación entre variables, etc., obligan a la intervención de investigador al menos para la formación de los grupos, la elección de las variables de estratificación, etc. Se trata en definitiva de ir más allá de la "manipulación física" de las variables para referirse a la intervención del investigador. Segundo, analizando dentro de esta perspectiva, los errores que se suelen cometer al utilizar la encuesta como técnica de investigación. Por tanto, la expresión correcta del problema sería el de "la validez en las investigaciones mediante encuesta".

La validez como proceso de inferencia

Resulta necesario proponer una clasificación de los aspectos sobre la validez de la investigación de los que se debe ocupar el investigador. La clasificación que se presenta aquí parte de la propuesta por Cook y Campbell (1979). Se introducen dos nuevos aspectos: la validez de la asignación numérica y las consideraciones sobre las consecuencias sociales de la realización de la encuesta, que intentan responder a las visiones más recientes sobre la validez (e. g. Messick, 1995).

Validez de asignación numérica

Este tipo de validez, se relaciona directamente con el proceso de escalamiento diseñado para la obtención de los valores numéricos con los que se cuantifican los constructos objeto de la medición. Veamos un ejemplo, que nos ayudará en la comprensión y que seguiremos utilizando más adelante.

Recientemente, la Federación Andaluza de Jugadores de Azar en Rehabilitación (FAJER), promovió un estudio sobre el nivel de adicción a los juegos de azar en Andalucía. Pero, ¿Cómo puede entenderse la variable "nivel de adicción"? ¿Se trata de un continuo unidimensional o, por el contrario, existen distintos tipos de adicción al juego? Estas preguntas tienen respuesta en el contexto de la teoría de la medida; contenidos a los que no puede permanecer ajena la investigación mediante encuesta. Sin entrar en detalles, una vez aceptada una definición específica de la variable, los posibles valores de esta conforman un sistema relacional empírico (Jáñez, 1989). El procedimiento de medida utilizado generará a su vez un conjunto de valores que conforman un sistema relacional numérico. La correspondencia entre ambos sistemas es el problema de la significación. Problema que, utilizando sus propias palabras, "Se plantea en el momento de realizar inferencias sobre un sistema relacional empírico, basándose en propiedades descubiertas en el correspondiente sistema relacional numérico", (Jáñez, 1989:414).

Volviendo a nuestro ejemplo, supongamos que conceptualizamos la variable "nivel de adicción" de forma tal que constituye teóricamente un sistema relacional empírico en que se pueden establecer las relaciones propias de una escala de intervalos. Es decir, en teoría es posible establecer diferencias significativas entre distintos niveles de adicción. Puede ocurrir que durante el proceso de asignación numérica, proceso que se concreta en la tarea planteada por los formatos de respuesta en el cuestionario (e. g. tarea de ordenación frente a una "tipo Likert"), se realice algún tipo de operación que distorsione los valores finalmente obtenidos, haciendo que en el sistema relacional numérico no se mantenga la significación de esas diferencias.

Validez de conclusión estadística

Una vez obtenidas las medidas y validada la asignación numérica, el investigador se embarcará en una serie de análisis de los datos para llegar a conclusiones como: "Las puntuaciones medias en adicción de la población disminuyen significativamente después de realizar campañas de concienciación". Para llegar a esta conclusión el investigador habrá usado algún test estadístico. Ninguno de estos tests es una herramienta infalible y de uso universal. El investigador habrá de preguntarse si las condiciones en las que realiza el análisis son apropiadas (e. g., tamaño de las personas realmente encuestadas, forma de la distribución de respuestas, etc.), y si permiten la interpretación que se quiere hacer de los resultados.

Validez interna

Suponiendo que la conclusión anterior estuviese justificada, el responsable de la investigación podría atribuir las diferencias encontradas en los niveles de adicción al programa de concienciación. Es decir, estableciendo una relación causal entre la

realización de campañas de concienciación y la disminución de la adicción a los juegos de azar. La justificación de la inferencia causal es terreno de la validez interna de la investigación. Cook y Campbell (1979) ofrecieron un listado con las posibles amenazas. Este tipo de interpretaciones, que en principio podrían parecer impropias de una investigación mediante encuestas, cobran relevancia en el contexto de las encuestas analíticas mencionadas antes.

Validez de constructo

Si aceptamos que la causa del cambio en los niveles de adicción es el programa de concienciación, cabría preguntarse qué significan realmente las mediciones de los niveles de adicción. En otras palabras, se debe analizar la validez de constructo de la medida realizada. No hay que confundir el problema de la validez de constructo con el de la validez de la asignación numérica. Es posible que la naturaleza teórica del constructo permita una interpretación cuantitativa (i. e., validez de la asignación numérica) pero que en la práctica, los resultados obtenidos no representen adecuadamente esa variable. Esto puede deberse a distintas causas señaladas por Cook y Campbell (1979) como amenazas a la validez de constructo, entre estas podríamos citar, por ejemplo, la falta de sinceridad en las respuestas de los entrevistados.

Validez externa

Aún admitiendo que la relación es causal y que se da entre los constructos supuestos, hay un hecho evidente, los resultados se han obtenido en una muestra de sujetos concreta y en un momento y lugar determinado. Este problema va más allá del grado de representación que la muestra tenga con respecto a la población de la que se extrajo y debe incluir aspectos como la estabilidad temporal que tengan los resultados. Siguiendo con el ejemplo, deberíamos preguntarnos si dentro de algún tiempo, los cambios en los niveles de adicción encontrados, y adjudicados al programa, se mantendrán o si, por el contrario, volverán a sus valores previos. Todos estos aspectos se encuadran dentro de la validez externa.

Consecuencias sociales de la realización de la encuesta

Podríamos decir que el proceso de inferencia desde un punto de vista técnico, en cuanto a la interpretación en sí de los resultados, termina con la validez externa. Sin embargo, y en la línea de otros autores (e. g., Cronbach, 1982; Messick, 1989) creemos necesaria una última reflexión. Ésta se refiere a la valoración de las consecuencias que la administración de la encuesta pueda tener sobre personas e instituciones y al análisis de los valores y contenidos ideológicos implicados en la encuesta. Incluso cabría plantear la responsabilidad del investigador mediante encuestas de contribuir a

la valoración de estas como técnica de investigación. Todas estas reflexiones cobrarán cada vez más importancia conforme aumente la relevancia de las decisiones tomadas a partir de las encuestas.

Los errores en la encuesta

Groves (1989), realiza un excelente análisis de las fuentes de error que pueden aparecer en una encuesta. Descompone el error total en dos apartados, varianza y sesgo en función de si este error es o no constante. Cada uno de estos elementos los divide a su vez en errores de no observación y de observación. Los primeros se descomponen en error de no cobertura, de no respuesta y de muestreo. Los segundos aluden al entrevistador, entrevistado, instrumento y modo de realización de la encuesta. En cualquier caso, nos permitimos hacer algunos comentarios sobre esta clasificación de los errores y que será utilizada para el posterior esquema conceptual:

1. Existe un matiz diferenciador entre lo que Groves llama errores de *no observación* y de *observación*. Mientras los primeros son realmente errores en sí, los segundos son fuentes de error que pueden provocar, entre otros, errores de no observación.
2. El modo en que se realiza la encuesta es una fuente indirecta de error. Lo es en la medida que va a determinar los errores de no observación por un lado, y la forma y/o conducta de entrevistadores, entrevistados e instrumento por otro.
3. Aunque Martínez (1995b) si lo incluye al citar su trabajo, Groves (1989) no trata directamente los errores de análisis y procesamiento. Es conveniente incluirlos y separarlos de los demás porque, al contrario que estos, se producen después de recogidos los datos.

Esquema de validación para el análisis de los errores

Hemos visto hasta ahora, de forma separada, la tipología de las inferencias propias de la encuesta y una clasificación de las distintas fuentes de error que pueden aparecer durante la realización de la misma. Sin embargo, como dijimos anteriormente, cuando realmente pueden resultar útiles ambas aproximaciones es cuando se tratan de forma conjunta. Saber, por ejemplo, que un entrevistador puede estar cometiendo algún tipo de error durante la encuesta, es un dato poco útil si no sabemos que efecto concreto puede estar ocurriendo. Es decir, a qué tipo de interpretación equivocada puede llevarnos su error. Por otro lado, de nada sirve reflexionar sobre la importancia de la validez si no somos capaces de encontrar los elementos específicos o las potenciales fuentes de error que pueden amenazar el proceso de inferencia. Veamos ahora como pueden conjugarse ambas perspectivas en un esquema conceptual conjunto.

La Tabla 1 (ver Anexo I) presenta un cruce entre fuentes de error y los distintos tipos de validez. En el interior de la misma indicamos, de forma esquemática, cómo

puede concretarse cada tipo de error en el cruce de cada fila y columna. Determinadas casillas de la tabla aparecen sombreadas. Esto es así para indicar que esos errores, los de observación, no pueden en ningún caso afectar al tipo de validez que aparece en la fila. Tanto la validez de asignación numérica como la de constructo dependen exclusivamente del proceso de medida, y no pueden verse afectadas por la cuantía o tipo de datos no obtenidos. Por otro lado, el lector notará que no hay ninguna columna dedicada al modo en que se realiza la encuesta. Tal como expusimos con anterioridad, concebimos el modo como una fuente indirecta de error. Un cambio en la forma en que se realiza la encuesta va a afectar a todos los elementos que la componen. Por tanto, la forma adecuada de proceder consistiría en 1º) decidir el modo de realización de la encuesta; y 2º) en ese marco, analizar el efecto de las distintas fuentes de error sobre los distintos tipos de inferencia.

A continuación, comentamos los elementos que aparecen en el interior de la Tabla 1, haciendo constar que no pretendemos ser exhaustivos. Pueden existir muchos más factores que podrían identificarse con un tipo de error y validez concretos, pero su determinación depende de la encuesta específica que se intente llevar a cabo.

El interesado en analizar la validez de una encuesta en particular podría utilizar este esquema para ir planteándose preguntas como "¿Qué aspectos relacionados con *COLUMNA* (por ejemplo, el entrevistador) pueden estar afectando a la validez *FILA* (por ejemplo, de constructo)?" .

– *Varianza*. Afecta a la validez de conclusión estadística en la medida que determina la cantidad y la aleatoriedad de los datos. Factores a los que son muy sensibles la mayoría de tests estadísticos. No se debe confundir el significado del término varianza aquí utilizado con el del término estadístico. Nos referimos a un error no sistemático. Por ejemplo, durante el desarrollo particular de una de las muchas encuestas sobre la población activa, es posible que el número de potenciales encuestados finalmente encuestados resulte inusualmente reducido.

– *Sesgo*. Afecta a la validez interna ya que puede ocurrir que los sujetos "no-observados" posean características distintas con relación a las variables medidas. Por ejemplo, en encuestas telefónicas, las personas sin teléfono nunca podrán formar parte de la muestra seleccionada.

Ambos, varianza y sesgo, determinan qué poblaciones y en qué grado están representadas por la muestra (validez externa). Es, como dijimos al principio, el tipo de error o análisis de la validez de la encuesta tratado habitualmente.

– *Adecuación al proceso de escalamiento previsto*. El proceso de escalamiento previsto, va a depender de los actores encargados de llevarlo a cabo (entrevistador y entrevistado) y de la herramienta con la que cuentan (cuestionario). Si esta alteración se produce de forma sistemática, los números finalmente obtenidos no tienen porque tener la misma base que los que se obtendría si se procede de forma correcta a la asignación numérica, distorsionando por tanto este tipo de validez.

– *Fiabilidad en el proceso de medida* Presenta un matiz distinto al anterior, y puede identificarse más con los errores de medida aleatorios y con la fiabilidad en sentido clásico. En esta línea, afectaría en mayor medida a la validez de conclusión estadística.

– *Utilizar distintos entrevistadores o instrumentos.* Este hecho representa una clara amenaza a la validez interna. Retomando el ejemplo del juego, contar con entrevistadores o instrumentos distintos antes y después de la realización de programa, podría ser la causa oculta de las diferencias encontradas.

– *Historia y maduración.* Los encuestados no son personas que viven en burbujas sometidos únicamente a los efectos de las variables que se investigan. Siguiendo con el mismo ejemplo, es posible que las personas cambien (maduración) o que sucedan hechos simultáneamente a la implementación del programa de concienciación que puedan explicar los cambios encontrados.

– *Expectativas del entrevistador.* Especialmente cuando debe codificar respuestas o indagar el sentido de algunas respuestas. Los valores obtenidos pueden estar representando más la imagen que el entrevistador tiene del encuestado que el verdadero valor de éste en la variable.

– *Deseabilidad social.* En relación también con las expectativas pero desde el punto de vista del entrevistado, cuando no responde en función de sus verdaderas creencias sino en función de lo que cree que se espera de él.

– *Perspectiva psicométrica.* Con esta expresión queremos resumir los aspectos clásicos de la literatura psicométrica. Entre ellos, por ejemplo, la definición del constructo, el muestreo del dominio, etc.

Las fuentes de error de observación pueden afectar de forma indirecta a la validez externa en la medida en que alteren los errores de no observación. Por ejemplo, entrevistadores que dejan de realizar un subconjunto específico de encuestas.

Con respecto a las consecuencias de la realización de la encuesta, se puede ver afectada, por ejemplo, por los errores de no observación en la medida en que no se refleja el *derecho a opinar de los sujetos no observados* y por los errores de observación si en el proceso de medida se atenta de alguna manera contra la intimidad o dignidad de las personas.

Por último, un *procesamiento y análisis de datos inadecuado* puede invalidar una investigación bien desarrollada. Esta fase es crucial puesto que, en definitiva, es la que traduce los datos brutos recogidos en la investigación en los resultados que serán la base de toda interpretación.

Conclusiones

La encuesta ha evolucionado y su uso se ha hecho cotidiano en la sociedad actual. Sin embargo, existe cierto desfase entre teoría y práctica. Esto puede ser porque la encuesta ha ampliado su ámbito de actuación abarcando cada vez objetivos más

amplios de investigación. De ahí, la necesidad de un esquema conceptual que permita al investigador o usuario de la encuesta juzgar la calidad de la información aportada por la misma desde una perspectiva global.

Por otro lado, su sencillez conceptual, unida a su facilidad de aplicación en múltiples campos, hace de la encuesta una herramienta muy atractiva en múltiples disciplinas. Los investigadores de distintas áreas tienen por tanto un acceso sencillo a lo que podríamos denominar "parte procedimental" en la realización de una encuesta. Con esto queremos indicar la secuencia de acciones más o menos automáticas que pueden dar a una investigación la apariencia de una encuesta. Por ejemplo, la selección de una muestra, la redacción en apariencia sencilla de las preguntas, la aplicación de un cuestionario y el análisis descriptivo de los datos. No ocurre lo mismo con el "soporte teórico". Con esta expresión no nos referimos al contenido específico de cada área o investigación, en el que cabe esperar que el investigador tenga un conocimiento suficiente. Nos referimos a aspectos como la teoría de la medida, teorías de tests, teoría de la validez, etc. Estos aspectos, de pertinencia indudable en cualquier encuesta, quedan fuera del interés específico de muchas disciplinas.

Esta actitud puede apreciarse claramente en el trabajo de Groves (1989) ampliamente citado en este artículo. En la página 27, tras un breve repaso a distintas definiciones de validez dice:

*"The more one reads of errors in psychological measurement, the more types of validity that one encounters. Each author seems to feel both free to ignore phrases used in prior research and compelled to invent his/her own."*¹ (p. 27)

No podemos dejar de apreciar en este texto una cierta infravaloración de la importancia que el tema de la validez tiene en una encuesta. En nuestra opinión, en cambio, la validez debería ser el centro de interés, tanto de la investigación mediante encuestas como de cualquier otro tipo de investigación. No olvidemos que el objetivo final de cualquier estudio es realizar algún tipo de inferencia. Y que la validez, en su definición más sencilla, no es otra cosa que la justificación empírica con que cuentan esas inferencias. La motivación fundamental de este trabajo ha sido mostrar cómo pueden aunarse distintas perspectivas dentro de un esquema único para alcanzar un resultado más válido.

¹ Cuanto más lee uno sobre errores en la medición psicológica, más tipos de validez encuentra. Cada autor parece sentirse, tanto libre de ignorar expresiones usadas en investigaciones previas como obligado a inventar las suyas propias.

Referencias

- Abramson, J. H. (1990) *Survey methods in community medicine*. London: Churchill Livingstone.
- Aiken, L. R. (1997). Questionnaires and inventories: Surveying opinions and assessing personality. New York: John Wiley and Sons, Inc.
- Alós, J. S. (1996) Industria de los Estudios de Mercado en España 1995. *Investigación y Marketing*, vol. 56, pp. 73-74
- Alós, J.S. (1997). Industria de los Estudios de Mercado en España 1996. *Investigación y Marketing*, vol. 56, pp. 73-74.
- Alós, J.S. (1998). Industria de los Estudios de Mercado en España 1997. *Investigación y Marketing*, vol. 61, pp. 61-62.
- Arnau, J. (1990) Metodología Experimental. En J. Arnau, M. T. Anguera y J. Gómez (1990) *Metodología de la investigación en ciencias del Comportamiento*. Murcia: Universidad de Murcia.
- Belson, W. A. (1986). *The design and understanding of survey question*. Aldershot: Gower.
- Boch, J. L. y Torrente, D. (1993). Encuestas telefónicas y por correo. Madrid: CIS.
- Bunge, M. (1976) *La ciencia y su método*. Buenos Aires: Siglo Veinte.
- Cochran, W. G. (1977) *Sampling techniques*. New York, John Wiley.
- Cook, T. D. y Campbell, D. T. (1979). *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Chicago, IL: Rand McNally.
- Cronbach, L. J. (1982). *Designing evaluations of educational and social programs*. San Francisco, CA: Jossey-Bass.
- Foddy, W. (1993). *Constructing questions for interviews and questionnaires*. Cambridge University Press.
- Fowler, W. (1993). *Constructing questions for interviews and questionnaires*. Cambridge: Cambridge University Press.
- Gómez, J. (1990) Metodología de la Encuesta por muestreo. En J. Arnau, M. T. Anguera y J. Gómez (1990) *Metodología de la investigación en ciencias del Comportamiento*. Murcia: Universidad de Murcia.
- González, A. y Padilla, J. L. (1998) La validez en la Investigación mediante encuestas. En V. Manzano y M. Sánchez (comps.) *Investigación del Comportamiento. Innovaciones metodológicas y estrategias de docencia*. Sevilla: Instituto Psicosociológico Andaluz de Investigaciones. 85-99.
- Groves, R. M. (1989). *Survey errors and survey cost*. New York: John Wiley & Sons.
- Jáñez, L. (1989) *Fundamentos de Psicología Matemática*. Madrid: Pirámide.
- Litwin, M. S. (1995). How to measure survey reliability and validity. California: Sage Publications, Inc.
- Martínez, R. M. (1995a). El método de encuestas por muestreo: conceptos básicos. En M. T. Anguera, J., Arnau, M. Ato, R. Martínez, J. Pascual y G. Vallejo (Eds.).

- Métodos y técnicas de investigación en Psicología* (pp. 485-510). Madrid: Síntesis.
- Martínez, R. M. (1995b). Las decisiones posteriores al muestreo. En M. T. Anguera, J., Arnau, M. Ato, R. Martínez, J. Pascual y G. Vallejo (Eds.). *Métodos y técnicas de investigación en Psicología* (pp. 485-510). Madrid: Síntesis.
- Messick, S. (1989). Meaning and values in test validation: the science and the ethics of assessment. *Educational Research*, 18, 5-11.
- Messick, S. (1995). Validity of psychological assessment. *American Psychologist*, 50, 741-749.
- Padilla, J. L. y González, A. (1998). Elaboración del cuestionario. En A. J. Rojas, J. A. Fernández y Pérez, C. (Eds.). *Investigar mediante encuestas*. Madrid: Síntesis.
- Scherpenzeel, A.C. y Saris, W.E.. The Validity and Reliability of Survey Questions: A Meta-Analysis of MTMM Studies. *Sociological Methods and Research*. 25, 341-383.

Anexo I

Tabla 1: Esquema para el análisis de los errores en la encuesta

		FUENTES DE ERROR						
		Cobertura	No respuesta	Muestreo	Entrevistador	Entrevistado	Instrumento	Procesamiento
TIPOS DE VALIDEZ	Asignación numérica	—	—	—	Adecuación al proceso de escalamiento previsto			Alteración de datos y adecuación del análisis
	Conclusión estadística	Varianza			Fiabilidad en el proceso de medida			
	Interna	Sesgo			Distintos	Historia y maduración	Distintos	
	Constructo	—			Expectativas	Deseabilidad social	Perspectiva psicométrica	
	Externa	Sesgo y varianza			Efecto sobre los errores de no observación			
	Ética	Derecho a opinar de los sujetos <i>no observados</i>			Violación del derecho a la intimidad, preguntas ofensivas, engaños...			