

LA CONSTRUCCIÓN DE INDICADORES VÁLIDOS EN LA INVESTIGACIÓN EN COMUNICACIÓN

Belinda de Frutos Torres
Universidad de Valladolid. Campus Público María Zambrano, Segovia
bfrutos@gmail.com

Resumen

El proceso de medición en la investigación cuantitativa aplicada a las ciencias sociales es una de las cuestiones claves en la metodología a la que no siempre se le dedica la atención debida. En el ámbito de la investigación aplicada a comunicación tendemos a confiar en escalas creadas para medir actitudes, en buscar indicadores de comportamiento respecto a usuarios, o en medir los efectos de los mensajes sin analizar el proceso de medida, es decir, el procedimiento por el que el indicador elegido o los indicadores elegidos están representando adecuadamente a la variable, concepto o constructo bajo estudio. Tomando como punto de partida la teoría de la medida planteada por Rossiter (2011) que critica el modelo psicométrico tradicional, el trabajo plantea la necesidad de garantizar la validez de constructo en la construcción y uso de indicadores aplicados a las ciencias sociales, en particular en el ámbito de la comunicación. Para ello se propone una técnica de valoración interjueces como un procedimiento que gracias a la independencia de las evaluaciones permita garantizar la validez de constructo de los indicadores. Así mismo se analiza la importancia de garantizar la validez de la escala de respuesta a fin de que se ajuste lo más adecuadamente posible al tipo de respuesta de los usuarios.

Palabras clave: Proceso de medida, teoría de medida en ciencias sociales, validez de contenido, validez de respuesta.

1. Introducción

Uno de los aspectos más importantes en la metodología de investigación cuantitativa es el proceso de medición. La característica principal de las técnicas cuantitativas es precisamente, y como su propio nombre indica, cuantificar o mediar las variables bajo estudio. El proceso mediante el que se realiza una asignación numérica al evaluar cada variable se conoce como cuantificación. La medición o asignación numérica en el proceso de investigación ha de seguir unas reglas explícitas, que indican la correspondencia entre el número asignado y alguna propiedad o característica del objeto de medición. Como indican Wimmer y Dominick (1996, pp. 51) “las reglas mediante las que se han asignado los números son la clave del sistema de medida, si estas fallan o resultan defectuosas todo el sistema fallaría”. En algunos casos la regla es bastante explícita y el proceso de medición evidente, si quiero conocer la relación entre la edad y el visionado de películas de acción, puedo medir la primera variable

Investigar la Comunicación hoy **Revisión de políticas científicas y aportaciones metodológicas**

registrando la edad cronológica expresada en años de las personas bajo estudio. Sin embargo, en comunicación muchas variables hacen referencia a procesos psicológicos donde no resulta tan evidente esta correspondencia. La credibilidad de un mensaje publicitario o la preferencia hacia los contenidos audiovisuales del género de acción son ejemplos de variables que requieren un proceso más complejo para su medición, puesto que implican procesos internos que no se evidencian de forma tangible. En este caso, la metodología de las ciencias sociales ha confiado en el desarrollo de escalas para medir constructos psicológicos que no podían ser observables externamente. El desarrollo de dichas medidas asume que para evaluar el proceso psicológico lo más apropiado es confiar en la información que proporcionan los propios individuos, puesto que no hay indicios externos del concepto bajo estudio. Para ello se formulan preguntas o afirmaciones, en la mayor parte de los casos, que permitan conocer sus respuestas y que a su vez se cuantifican mediante escalas. En el ámbito de las ciencias sociales el método de escalamiento que más se utiliza es la escala propuesta por Likert, en el que ante una afirmación las personas que responden declaran su grado de acuerdo con la afirmación presentada. La respuesta se gradúa con cinco niveles: muy de acuerdo, de acuerdo, neutral, desacuerdo o muy desacuerdo; de modo que a cada nivel se le asigna un número que suele ir del 1 al 5. Otra de las técnicas de escalamiento más utilizadas en la investigación es el diferencial semántico de Osgood, que a diferencia de la anterior tiene como objetivo medir el significado que para cada individuo tiene determinado asunto (Osgood, Suci y Tannenbaum, 1957). El aspecto más característico de esta forma de escalamiento es que la respuesta se recogen en una escala bipolar con un número impar de tramos 5 ó 7 en cuyos extremos se disponen adjetivos del tipo: agradable/desagradable, honesto/deshonesto, justo/injusto, etc. Thusthorne y Guttamn propusieron otras técnicas de escalamiento algo más complejas en su procedimiento, que no han sido tan utilizadas posteriormente por la literatura científica Wimmer y Dominick (1996).

2. Fiabilidad y validez en el proceso de medida

La aplicación de una escala no garantiza el éxito en el proceso de medición. La teoría psicométrica ha sido dominante en la medición aplicada a las ciencias sociales y plantea la necesidad de garantizar la fiabilidad y validez en los instrumentos de medida. Antes de analizar cómo se plantean ambas propiedades debemos analizar dos supuestos de la teoría psicométrica. El primer supuesto de la teoría clásica de test, según fue planteada por Spearman (1904), asume que el resultado de la evaluación con un instrumento de medida (Observación) es el resultado de sumar la “puntuación verdadera” del individuo en la variable bajo estudio más un error aleatorio. Estos errores aleatorios ocurren durante el proceso de medición, se atribuyen a al evaluador y pueden deberse a circunstancias del momento de la evaluación (ej. Cansancio acumulado, tendencia a marcar el punto medio en las escalas bipolares, adivinar la respuesta deseada, omitir respuestas, por enumerar algunas de las fuentes de distorsión reconocidas). La teoría clásica de test reconoce la existencia de este tipo de errores si bien al tener un carácter aleatorio suman cero si tenemos un buen número de ítems y calculamos la puntuación media observada. De aquí se deriva un segundo supuesto, para poder inferir la puntuación verdadera es necesario obtener varias

Investigar la Comunicación hoy Revisión de políticas científicas y aportaciones metodológicas

puntuaciones observadas, ello implica desarrollar una escala formada por varios ítems dirigidos a medir el objeto y cuyo valor promedio reflejará la puntuación verdadera.

$$\text{Observación} = \text{Puntuación verdadera} + \text{error}$$

La fiabilidad es una propiedad que recoge la estabilidad y coherencia del instrumento de medición. El planteamiento es que si se mide, por ejemplo, la actitud hacia una marca, el resultado obtenido sea consistente en el tiempo, es decir, que si se vuelve a repetir la medición en un tiempo prudencial se obtendrán resultados semejantes. No obstante la teoría psicométrica admite que puede haber una variación o error de una ocasión de medición a otra. Por lo tanto, si consigo desarrollar un instrumento de medida que me aporta puntuaciones y éstas son estables en el tiempo, podré afirmar que el instrumento que he desarrollado es fiable, es decir, consistente en el tiempo. En la práctica la fiabilidad se mide con el índice de consistencia interna o coeficiente alpha desarrollado por Cronbach, y pocas ocasiones se informa de una prueba de test-rest. El índice alpha de Cronbach refleja el grado de relación entre los ítems desarrollados para medir la misma variable o constructo, es decir, su correlación. Por tanto, este indicador debería interpretarse como el grado de coherencia de los ítems entre sí. En la práctica, el índice alpha se utiliza para depurar los ítems que componen una escala, eliminando aquellos que tienen correlaciones bajas con el resto de los ítems. Además no puede aplicarse si sólo hay un indicador para medir una variable.

Continuando con las propiedades psicométricas de los instrumentos de medida, la fiabilidad de una herramienta por sí sola no garantiza su utilidad, puesto que para su uso es necesario garantizar también la segunda propiedad que es la validez. En una primera aproximación diremos que un instrumento es válido si evalúa adecuadamente su objetivo, tal propiedad que fue enunciada por Kelly (1927), plantea de una forma bastante sencilla que un instrumento será válido “si mide lo que pretende medir”. Es decir, se busca la correspondencia semántica entre la definición conceptual de un constructo y su medida. Wimmer y Dominick denominan a este tipo como *validez aparente*, planteando que se da “cuando examinando el instrumento de medición da la impresión a simple vista que mide lo que parece medir” (Wimmer y Dominick, 1996, pag. 61). La psicometría no obstante se ha ocupado de cuantificar empíricamente la validez a través del comportamiento de las mediciones con otras ya existentes, de modo que la validez se establece a través de la validez predictiva, la validez discriminante y/o la validez concurrente. Ninguna de ellas se dirige de forma directa a las propiedades de medición del instrumento respecto al constructo que pretende medir. Si examinamos, por ejemplo la validez predictiva, esta evalúa el grado de ajuste entre las puntuaciones obtenidas por el instrumento de medida y los resultados obtenidos en otra variable. Por ejemplo, podríamos ver la validez predictiva de la actitud hacia una marca y la compra posterior de un producto perteneciente a dicha marca. Como Wimmer y Dominick (1996) señalan la validez predictiva indica el grado de relación con otra variable, que en este caso es de tipo comportamental, pero no garantiza si mide adecuadamente el concepto que se desea medir. Lo mismo puede plantearse con la validez concurrente, esto es, la correlación entre el instrumento y otras medidas o elementos de comparación; o con la validez discriminante, que

Investigar la Comunicación hoy **Revisión de políticas científicas y aportaciones metodológicas**

verifica en qué medida el instrumento discrimina respecto a otras variables con las que teóricamente no guarda relación.

La importancia de la validez de contenido es crucial en todo proceso de medición empírica, a pesar de que un instrumento pueda alcanzar los estándares de fiabilidad o haya sido utilizado en anteriores trabajos de investigación publicados es necesario garantizar la validez de contenido. Con bastante frecuencia los investigadores tendemos a utilizar las mismas medidas que en otros trabajos de investigación publicados dando por hecho su validez para nuestros objetivos de investigación.

3. El modelo COARSE

El presente trabajo afronta cómo desarrollar instrumentos de medida que aseguren la validez de contenido, para ello vamos a seguir el modelo COARSE formulado por el profesor Rossiter (2002, 2010). La propuesta de Rossiter plantea la construcción de instrumentos de medida basados en la argumentación racional o deducción lógica partiendo de la conceptualización teórica, a diferencia de la tradicional psicometría que se apoya únicamente en las pruebas empíricas. Esta teoría no se plantea necesariamente como una alternativa a los indicadores psicométricos de fiabilidad y validez, sino como un complemento a dichos indicadores, en este sentido, la teoría debe considerarse como un paso previo.

Para Rossiter (2002, 2010) sólo la validez de contenido es importante para el desarrollo de una escala, puesto que este tipo de validez asegura la relación entre el constructo y la medida que se ha desarrollado para su medición. Posteriormente, desde la psicometría se puede comprobar la relación entre la medida y las puntuaciones obtenidas. Para este autor, sólo podemos validar la correspondencia entre un constructo y su medida, en el sentido de que sea apropiada y representativa del constructo que pretende medir. Otros aspectos de la validez como la capacidad predictiva o discriminativa de una escala no tiene sentido su aplicación en tanto que se está utilizando un referente ajeno al propio constructo de partida. Tampoco la validez de constructo tiene sentido aplicarla en la medida en que este se desarrolla a partir de una teoría y debería ser desde ese marco donde se determinara si es apropiado o no (Rossiter, 2002).

El primer paso en la construcción de un instrumento de medida es la definición del constructo, que implica su definición en términos del objeto que va a ser evaluado, el atributo que va a ser evaluado y la entidad del evaluador. En ciencias sociales el constructo es una entidad creada por los investigadores, que se infiere su existencia a partir de sus manifestaciones o sus efectos, pero que no puede verse directamente. A pesar de que la definición de un constructo sea una abstracción su definición y sus componentes deben ser reales y concretos. Para la validez de contenido, no sólo debemos asegurar la correspondencia semántica entre el objeto de estudio y su definición, también se debe asegurar la correspondencia de la forma en la que recogemos la respuesta, de forma que esta sea relevante para el evaluador, de otro modo estaríamos cometiendo un error sistemático asociado al instrumento de medida.

Investigar la Comunicación hoy Revisión de políticas científicas y aportaciones metodológicas

Por lo tanto para que una medida sea válida debe cumplir con las siguientes propiedades:

“High item-content validity: This means that the semantic content of the question part of the ítem corresponds closely with a constituent or component of the object in the conceptual definition of the construct and with a component of the attribute in the conceptual definition unless the attribute is in the answer part.

High answer-content validity: This means that the semantic content of the answer part of the item allows the rater to see only the main answers alternatives that he o she has in mind, and to easily choose one answer that fits his or her true score. The answer part of the scale is always an attribute”. (Rossiter, 2010: 14).

Figura. 1. Representación gráfica de la teoría de Rossiter (2010)



Según plantea este autor en el proceso de medición están implicados, además del propio constructo los siguientes elementos:

El objeto que vamos a evaluar o conjunto de ítems que hemos desarrollado para ser evaluados.

El atributo en el que vamos a centrar la evaluación, es decir, la dimensión en la que queremos que se evalúe dicho objeto.

El evaluador, es decir, la persona que va a llevar a cabo la evaluación. Puede tratarse de un experto, un codificador, un directivo, un consumidor o un individuo.

El **tipo de ítems** y la **escala de respuesta** que se va a utilizar.

Investigar la Comunicación hoy **Revisión de políticas científicas y aportaciones metodológicas**

Las reglas de **asignación numérica**, es decir, cómo cuantificar las puntuaciones recogidas por las personas que responden.

Veamos con detenimiento el planteamiento de esta teoría, para después llevar a una aplicación práctica en el proceso de creación de una herramienta para medir valores publicitarios en jóvenes. Una primera recomendación que realiza Rossiter (2010) es distinguir el tipo de objeto bajo análisis, diferenciando entre objetos concretos o abstractos. Cuando el objeto a evaluar es un estímulo concreto este debe ser representado de forma concreta para su evaluación. Un ejemplo de estímulo concreto puede ser una película, una serie de televisión o un servicio. En el caso de productos concretos cuanto más directa es la evaluación más exacta será la medida, de modo que cuando el objeto tiene una representación icónica –por ejemplo un producto reconocible por el logotipo de la marca a la que pertenece- debe ser representado de la misma manera para asegurar la exactitud de la evaluación.

Otra tipología de objetos a evaluar son los de tipo abstracto de tipo colectivo, en esta categoría encontramos a categorías de objetos que constituyen una abstracción colectiva, pero esta se ha formado a partir de realidades concretas. Un ejemplo de esta categoría de objetos la encontramos en la televisión, el medio es una entidad abstracta pero que se concreta en canales que son concretos. Si en una investigación queremos conocer la valoración de la televisión, nos referimos a la categoría abstracta, en estos casos, para asegurarnos que en la valoración sea lo más exacta posible debemos asegurar que en la valoración incluimos una muestra representativa de entidades que pertenecen a la categoría. Si no se cuenta con un número suficientemente amplio y variado de entidades pertenecientes a la categoría corremos el riesgo de llevar a cabo una evaluación estereotipada. La tercera y última tipología de objetos de evaluación se refiere a entidades abstractas creadas (*Abstract formed objects*), que son las más difíciles de evaluar en tanto en cuanto pertenecen a entidades abstractas que no se han categorizado a partir de entidades concretas como en el caso anterior, sino que son entidades creadas que tienen componentes que constituyen los principales significados del objeto. El caso es más claro de este tipo de entidades lo encontramos en la evaluación de los valores, el tipo de objeto que se pretende medir en este trabajo. Es evidente que los valores son entidades abstractas, que pueden definirse como las metas duraderas que una persona se plantea en la vida. En este caso, el investigador debe ser explícito sobre su significado y el principal significado debe ser incluido en la medida, como parte del objeto a evaluar. Este debe ser lo suficientemente explícito para que se comprenda adecuadamente por la persona que va a responder.

La segunda dimensión de la validez viene del formato de la respuesta. Al igual que en la construcción de una escala debemos asegurar su correspondencia semántica con el constructo o concepto que se pretende medir, debemos asegurarnos de que el formato de la respuesta se corresponde con el atributo que se desea medir. En este sentido, la primera recomendación es analizar el tipo de atributo que está siendo evaluado de modo que el formato de respuesta o el tipo de escala propuesto se ajuste lo máximo posible. Si por ejemplo, tratamos de evaluar la percepción de una serie de atributos en una marca, estamos ante una evaluación de creencias sobre la presencia o

Investigar la Comunicación hoy

Revisión de políticas científicas y aportaciones metodológicas

no de un atributo determinado, para Rossiter (2010) se trata de un objeto concreto (CONCOB) y un atributo perceptual (CONCPERC), por lo tanto, lo más adecuado es utilizar una medida binaria, es decir, que el evaluador asigne si el atributo está presente en el objeto o no. Lo mismo podría decirse si lo que estamos evaluando un rasgo de personalidad, el evaluador (encuestado) se sentirá identificado o no con ese rasgo de personalidad. Rossiter recomienda elegir un formato de respuesta que sea realmente discriminante, en este caso, lo verdaderamente discriminante es la dirección que toma la respuesta (aceptación o rechazo), de modo que lo más apropiado sería una respuesta dicotómica de Sí ó No, siendo este tipo de respuestas más precisas y estables. En investigación social normalmente acudimos a las escalas para posicionar la respuesta, asumiendo que graduar la respuesta llevará a una medición más precisa. El planteamiento es que simplemente añadir más posibilidades de respuesta no va incrementar la precisión de la medida, al contrario, puede convertirse en una fuente de distorsión. Las fuentes de sesgo en el escalamiento son bien conocidas, destacamos por ejemplo la tendencia a señalar el punto medio de la escala (si es impar), la aquiescencia o tendencia a estar de acuerdo con todas las afirmaciones o la tendencia a responder de forma extrema. La ventaja de las evaluaciones bipolares es que se eliminan las diferencias individuales a la hora de dimensionar la intensidad de la respuesta. Las escalas tienen sentido si se está midiendo por ejemplo un test de habilidad, y entonces, es interesante graduar las puntuaciones según el nivel de habilidad.

Otra tipología de ítems son los que se refieren a la evaluación de creencias, en este caso, el atributo que está siendo evaluado no es si un rasgo está presente o ausente, sino hacer una evaluación del mismo y tiene sentido que el evaluador cuantifique su intensidad, el caso más frecuente de este tipo de atributos son las actitudes. El componente de la actitud es una valoración positiva o negativa hacia el objeto o entidad que está siendo evaluada. El formato de respuesta en este caso tiene sentido que sea bipolar (negativa o positiva), y para ello se puede ayudar de las escalas de respuesta tipo Likert, cuyos extremos se corresponden con el calificativo de “Muy de acuerdo - Muy en desacuerdo”. Aún en estos casos donde se recomienda el uso de escalas bipolares la gradación de la respuesta no se traduce necesariamente en mayor fiabilidad ni validez. Según demostró Pebody (1962) la mayor variabilidad de las puntuaciones obtenidas en una escala bipolar proceden de la dirección que toma la evaluación (positiva o negativa), siendo muy poca la variabilidad que se debe a las diferencias en intensidad de las respuestas. Otro resultado bastante ilustrativo sobre la artificialidad de la medida lo aportan Swain, Waeathers y Niedrich (2008), demostrando que los ítems redactados en sentido inverso en una escala tipo Likert – que habitualmente son la mitad de los ítems- son tan confusos para los que responden, que generalmente se agrupan en un segundo factor del análisis factorial completamente artificial, un resultado que he podido comprobar en mi experiencia en el desarrollo y análisis de escalas de actitudes.

4. Procedimiento

Investigar la Comunicación hoy **Revisión de políticas científicas y aportaciones metodológicas**

Siguiendo los aspectos enunciados en el planteamiento del trabajo a continuación se describe el procedimiento utilizado para desarrollar la herramienta dirigida a evaluar valores publicitarios en jóvenes (Sánchez y Frutos, 2011). Al tratarse de un tipo de objeto creado y de carácter abstracto se consideró oportuno dos cuestiones en el proceso, la primera es que la búsqueda de significados o tipologías de valores debían ser relevantes para las personas que actuarían como evaluadores, por ello se apostó por un proceso de abajo arriba. El segundo aspecto importante era asegurar que los indicadores creados fueran explícitos sobre su significado, para ello se recurrió al consenso inter jueces.

Rositer (2010) recomienda acudir a los propios sujetos con preguntas abiertas sobre sus metas, objetivos o propósito de la vida, para identificar los valores y posteriormente ser clasificado en las distintas tipologías. En la escala que se presenta en este trabajo se recurrió a un trabajo de investigación previo que de forma sistemática había recogido la tipología de valores presentes en la publicidad dirigida a jóvenes, que era el grupo al que se dirige nuestra investigación. La forma de identificación de los valores es relevante puesto que debe representar el universo de valores o metas finalistas de las personas que van a realizar la evaluación. El trabajo de Sánchez, Megías y Rodríguez (2004) se adecua a este proceso puesto que la extracción de afirmaciones sobre metas u objetivos en la vida fueron obtenidos a partir de una muestra de campañas publicitarias dirigidas a jóvenes procedentes de televisión, publicidad exterior y prensa. El conjunto de valores estaba formado por 30 afirmaciones sobre metas u objetivos finalistas que se habían agrupado en ocho categorías: valores ligados al pragmatismo y la competitividad, valores idealistas, valores vinculados con la transgresión y la aventura, valores ligados a la sociabilidad o de carácter gregario, valores altruistas, valores hedonistas y presentistas, valores de carácter normativo, y por último, valores centrados en la imagen personal. Cada afirmación describe de forma explícita el tipo de meta u objetivo en la vida, de modo que su interpretación fuera explícita. A pesar de que los valores habían sido utilizados en un trabajo previo, se realizó una validación de los mismos con una muestra de 34 campañas publicitarias elegidas en internet, la valoración se llevó a cabo con tres jueces expertos que formaban parte del equipo de investigación (Sánchez y Frutos, 2011). Una de las claves de este proceso de validación es que a diferencia del estudio previo de Sánchez, Megías y Rodríguez (2004), la validación interjueces se realizó con un criterio dicotómico, de modo que el tipo de respuesta en evaluación era 0 si el valor no estaba presente y 1 si este estaba presente. El hecho de utilizar una categoría dicotómica sin duda contribuyó a alcanzar un alto grado de consenso interjueces, garantizando así la validez de las afirmaciones cuyo consenso promedio fue del 88% (Frutos, Sánchez y Vázquez, 2012).

Con el objeto de garantizar la validez de la escala de respuesta y siguiendo con el modelo COARSE se analiza el tipo de atributo sobre el que se va a realizar la evaluación. Los valores se definen como una meta duradera en la vida, bien porque se busca en la vida (valores positivos) o bien porque se evita (valores negativos), por lo tanto, el tipo de atributo bajo análisis es la evaluación positiva o negativa hacia el objeto. Por otra parte, el evaluador debe pronunciarse sobre su propia percepción de si se identifica (aceptación) o no con el valor (rechazo). Para evitar la distorsión de la

Investigar la Comunicación hoy **Revisión de políticas científicas y aportaciones metodológicas**

respuesta con escalas que gradúan la intensidad de la respuesta se ha optado por la modalidad de respuesta de elección forzada. La presentación de las afirmaciones se realiza en triadas y el evaluador debe elegir la afirmación con la que más se identifica (aceptación) y la afirmación con la que menos se identifica (rechazo).

Las técnicas ipsativas o de elección forzada han sido utilizadas en la evaluación de constructos psicológicos duraderos, como por ejemplo los rasgos de personalidad, por la relevancia de identificar la prevalencia de unos rasgos sobre otros. Una revisión sobre este tipo de técnicas pone de manifiesto cierta controversia sobre su uso por su dificultad de aplicación de los parámetros estandarizados de las escalas psicométricas (Saville & Willson, 1991; Cornwell y Dunlap, 1994). El principal problema que ha presentado esta metodología es que no se pueden aplicar los estándares de fiabilidad y validez al no cumplir el supuesto de independencia de los ítems que componen la escala. Cornwell y Dunlap (1994) plantean que si las técnicas ipsativas no se basan en la teoría clásica de test (observación = puntuación verdadera + error), no tiene mucho sentido tratar de estimar su error. Si es cierto que no tienen sentido utilizar la consistencia interna de la escala, al ser las puntuaciones obtenidas en cada ítem co-dependientes, pero ello no implica que no se pueda estimar su estabilidad en el tiempo a través de una prueba de tes-retest. Incluso varios autores han informado reiteradamente de una mayor estabilidad en las escalas ipsativas precisamente por su formato de respuesta forzada (Saville & Willson, 1991; 2011; Bartram, 1996;). Además de la estabilidad de la respuesta, la verdadera ventaja de este tipo de técnicas está en el control sobre algunos de los sesgos que normalmente afectan a otras modalidades de escalamiento como es la deseabilidad social, la aquiescencia o la tendencia a puntuar en el punto medio de la escala (Saville and Willson, 1991). La elección forzada de una respuesta evita la tendencia a estar de acuerdo con todas las afirmaciones de la escala, aspecto que es especialmente relevante a la hora de evaluar valores con los que en alguna medida todas las personas se pueden sentir identificados. Al mismo tiempo, el formato de respuesta previene de la deseabilidad social puesto que el que responde debe elegir entre respuestas que en apariencia son igualmente favorables o desfavorables. Por último, al no tratarse de una escala de intensidad se evitan la tendencia hacia el punto medio de la escala o las diferencias individuales a la hora de interpretar la intensidad de las respuestas.

Por último, antes de comentar los resultados analizaremos como se ha procedido a asignar las puntuaciones de los encuestados. La elección de cada afirmación ha sido asignada la puntuación de un punto positivo, mientras que el rechazo de cada afirmación ha sido asignada la puntuación de un punto negativo. Las puntuaciones obtenidas en cada tipología de valores se han obtenido tras sumar las puntuaciones obtenidas en sus ítems. En la tabla 1 se aprecia que el rango de valores obtenidos se sitúa entre los cuatro puntos negativos y los cuatro puntos positivos. El valor numérico indica el número de ocasiones en las que los encuestados han rechazado o aceptado alguna afirmación perteneciente dicho valor. La escala de valores publicitarios se aplicó a una muestra formada por 320 jóvenes entre 15 y 20 años de edad, a los que se contactó a través de sus centros de estudios (centros de enseñanza media y universidad).

Investigar la Comunicación hoy
Revisión de políticas científicas y aportaciones metodológicas

5. Resultados

El análisis de las puntuaciones se centra en el comportamiento de la escala frente a los sesgos más frecuentes en ciencias sociales. En la figura 2 se muestran la distribución de las puntuaciones obtenidas. Si analizamos la simetría de la escala, es decir, el rango de las puntuaciones por encima y por debajo del punto medio, se puede apreciar que las categorías de valores se encuentran muy equilibradas, sólo en tres de las ocho categorías el índice de asimetría es mayor que 0,5. Este resultado implica que los participantes se han situado tanto en el lado positivo de la escala (aceptación) como en el lado negativo de la escala (rechazo). Las categorías de valores que han mostrado menos equilibrio han sido la normativa con una asimetría negativa, es decir, ha habido menor porcentaje de casos que han rechazado las afirmaciones pertenecientes a esta categoría y mayor porcentaje de casos que han aceptado sus afirmaciones. Algo similar ocurre en la escala idealista que también presenta asimetría negativa, es decir, mayor proporción de personas que se han identificado con las afirmaciones de esta categoría. La imagen personal también ha resultado tener asimetría, si bien en este caso es de signo positivo, de modo que ha habido mayor proporción de personas que han rechazado las afirmaciones pertenecientes a esta categoría de valores.

Tabla 1. Distribución de las puntuaciones obtenidas en la escala de valores publicitarios

Categorías de valores	Competitivo	Transgresor	Altruista	Hedonista	Gregario	Normativo	Idealista	Imagen personal
	% col.	% col.	% col.	% col.	% col.	% col.	% col.	% col.
-4,00	25,6	27,8	,0	,0	5,0	2,2	3,4	,0
-3,00	0	,3	27,3	16,9	0	,6	,0	44,2
-2,00	19,1	24,4	,3	,6	18,8	7,2	9,4	,6
-1,00	1,3	,3	23,5	25,4	,6	,0	,6	21,6
,00	17,5	15,6	,6	1,3	27,6	19,1	22,8	,3
1,00	,3	,9	19,1	35,7	22,6	,0	29,1	19,4
2,00	21,6	13,8	,6	,6	0	26,0	,3	,9
3,00	,0	,3	28,5	19,4	,0	1,3	,0	12,9
4,00	14,7	16,6	,0	,0	25,4	43,6	34,4	,0
N	320,0	320,0	319,0	319,0	319,0	319,0	319,0	319,0
Asimetría (estadístico F de Fisher)	,121	,360	,025	-,172	-,189	-,897	-,655	,581

Otra de las cuestiones analizadas es la tendencia a situarse en el punto medio de la escala. Según se puede apreciar los participantes no se agrupan en torno al punto medio de la escala, sólo tres categorías han agrupado un porcentaje de respuestas elevado en el promedio, se trata de los valores de tipo gregario, normativo e idealista.

Investigar la Comunicación hoy
Revisión de políticas científicas y aportaciones metodológicas

Tabla 2. Distribución de las puntuaciones obtenidas en la escala de valores publicitarios

	Rechazo	Neutral	Aceptación
	% de fila	% de fila	% de fila
Competitivo	9,3	73,1	17,6
Transgresor	27,9	55,4	16,7
Idealista	3,4	62,2	34,4
Normativo	2,8	52,9	44,3
Gregario	5,0	70,0	25,1
Hedonista	16,7	64,1	19,2
Altruista	26,9	44,9	28,2
Imagen personal	43,7	43,7	12,7

Por último, una referencia a las puntuaciones globales de las escalas permite describir cómo se identifican los participantes con los valores publicitarios. En la tabla 2 se muestran los resultados de los participantes según su identificación en las categorías de valores. Se han clasificado en el grupo de rechazo a los que han obtenido entre 3 y 4 puntos negativos; los que han obtenido una puntuación entre 3 y 4 puntos positivos se han asignado en la categoría de aceptación. Podríamos decir que los valores que tienen mayor aceptación son los normativos (44,3%), seguidos de los idealistas (34,4%) y los gregarios (25,1%). En contraste, hay mayor proporción de participantes que rechazan los valores centrados en la imagen personal (43,7%), la transgresión (27,9%) y el altruismo (26,9%). La interpretación de los resultados debe centrarse en el perfil de los participantes, en detrimento de las puntuaciones medias u otras medidas de tendencia central, debido precisamente a la codependencia de las elecciones.

6. Conclusiones

El trabajo aborda de una forma explícita el proceso de elaboración de una escala para medir valores publicitarios que aseguren la validez de contenido siguiendo el modelo COARSE formulado por el profesor Rossiter. El primer aspecto de la validez se centra en la construcción de los indicadores relevantes para el objeto de estudio que se pretende medir, en este sentido se ha apostado por un procedimiento explícito partiendo de los mensajes publicitarios que ha sido consensuado a través de la valoración interjueces con un alto grado de consenso.

Para garantizar la validez de la respuesta se ha optado por un formato de respuesta de elección forzada, que según los resultados obtenidos, ha tenido un buen rendimiento frente a los sesgos más frecuentes en las técnicas de escalamiento. Así se ha comparado que se ha evitado la tendencia de puntuar en el punto medio de la escala, hecho que facilita la clasificación de los participantes en sus preferencias hacia los valores. Así mismo, se ha mostrado un alto grado de simetría en las puntuaciones en la mayor parte de las escalas, evitando así la aquiescencia en las escalas que miden valores. Por último, la escala permite clasificar a los participantes según su aceptación

Investigar la Comunicación hoy
Revisión de políticas científicas y aportaciones metodológicas

o rechazo en las ocho categorías de valores publicitarios, indicativo de que este instrumento ha sido discriminante en el proceso de medida.

Por último, el proceso de construcción se vería completado si se evalúa la consistencia de la escala en el tiempo a través de un prueba re-test, que permite valorar la fiabilidad de la medida. Este aspecto se completará en la siguiente fase de la investigación.

7. Bibliografía y referencias metodológicas

BARTRAM, D. (2011) "The relationship between ipsatized and normative measures of personality." *Journal of Occupational and Organizational Psychology* 69.1: 25-39.

CORNWELL, J. M., and DUNLAP, W. P. (1994). "On the questionable soundness of factoring ipsative data: A response to Saville & Willson (1991)." *Journal of Occupational and Organizational Psychology* 67.2 (1994): 89-100.

FRUTOS TORRES, B., SÁNCHEZ-VALLE, M. & BARRIO, T. (2012). "Empirical analysis of values on interactive advertisement aimed at a teenage audience". *Comunicación y Sociedad*, 25(2), 17x-202.

KELLEY, T. L. (1927). *Interpretation of educational measurements*. Mcmillian: New York.
Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*. 140: 1-55.

OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.

PEABODY, D. (1962). Two components in bipolar scales: Direction and extremeness. *Psychological Review*, 69, 65.

ROSSITER, J.R. "The C-OAR-SE procedure for scale development in marketing." *International Journal of Research in Marketing* 19.4 (2002): 305-335.

ROSSITER, J. R. (2010). *Measurement for the social sciences: The C-OAR-SE method and why it must replace psychometrics*. Springer.

SÁNCHEZ, L., MEGÍAS, I., & RODRÍGUEZ, E. (2004). Jóvenes y publicidad. Valores en la comunicación publicitaria para jóvenes. *Madrid: Fundación de Ayuda contra la Drogadicción. Instituto de la Juventud*.

SÁNCHEZ-VALLE, M., & FRUTOS, B. (2011). La convergencia de metodologías para el estudio de los efectos de la publicidad en medios interactivos. In III Congreso Internacional Latina de Comunicación Social.

Investigar la Comunicación hoy
Revisión de políticas científicas y aportaciones metodológicas

SAVILLE, P. & WILLSON, E. "The reliability and validity of normative and ipsative approaches in the measurement of personality." *Journal of Occupational Psychology* 64.3 (2011): 219-238.

SAVILLE, P. & WILLSON, E. "The reliability and validity of normative and ipsative approaches in the measurement of personality." *Journal of Occupational Psychology* 64.3 (1991): 219-238.

SPEARMAN, C. (1904). " General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology*, 15(2), 201-292.

SWAIN S. D., WEATHERS, D. & NIEDRICH R. W. (2008). Assessing three sources of misresponse to reversed Likert items. *Journal of Marketing Research*, 45 1: 116-131.

WIMMER, R. D., & DOMINICK J. R. (1996). *La investigación científica de los medios de comunicación :Una introducción a sus métodos*. Bosch comunicación. Vol. 18. Barcelona: Bosch.

Investigar la Comunicación hoy
Revisión de políticas científicas y aportaciones metodológicas

Anexo 1. Afirmaciones utilizadas para expresar los valores publicitarios desarrolladas a partir de Sánchez, Megías y Rodríguez (2004)

TIPOLOGÍA	AFIRMACIONES
Altruismo	Colaborar con otras personas como tu sobre todo en situaciones difíciles
	Ayudar desinteresadamente a personas en apuros o que pasan por dificultades de cualquier tipo
	Apoyar y luchar por los más necesitados
Normativo	El esfuerzo y la superación personal para superar las adversidades y los obstáculos
	Prepararse para el futuro aunque cueste
	Cuidar la salud y mantenerse en buena forma física
	Vivir en armonía con las personas que te rodean y mostrar tu afecto a tus seres queridos
Gregario	Conectar con personas que piensan como tú
	Sentirse parte de un grupo con el que compartes experiencias, aficiones e intereses
	Ganarse la confianza y el afecto desinteresado de las personas
	Saber ser tu mismo y no dejarte llevar por las presiones de tus padres
Hedonismo	Dejarse llevar por los placeres de la vida
	Vivir al día y disfrutar de cada momento porque no sabes que te depara el futuro
	Divertirse y saber disfrutar del tiempo libre y del ocio
Idealismo	Tener libertad para actuar y poder elegir lo que quieres
	Poder elegir en las opciones que te ofrece la vida
	Ser uno mismo y expresar tu propia personalidad, tu manera de ser
	Ser independiente, tomar tus propias decisiones al margen de los demás
Imagen personal	Cuidar la imagen ante los demás
	Sentirse físicamente atractivo
	Saber seducir, tener gancho
Competitivo	Tener éxito profesional, ser reconocido por los demás
	Ir a por todas y conseguir ser el mejor
	Luchar para ganar y estar entre los mejores
	Poder comprar o disfrutar de las cosas buenas de la vida
Transgresor	Romper los moldes y resistir a las presiones del entorno
	Arriesgarse y no dejarse llevar por lo establecido
	Experimentar nuevas sensaciones y emociones
	No tener que ajustarse a lo establecido y a lo que esperan de ti

VALID MEASUREMENT FOR COMMUNICATION RESEARCH

Belinda de Frutos Torres
Universidad de Valladolid. Campus Público María Zambrano, Segovia
bfrutos@gmail.com

Abstract

The measurement process in quantitative research applied to social sciences is a key issue in methodology which not always devotes due attention. In communication research often we tend to trust on scales previously created to measure attitudes, to trust at user's performance indicators or measure the effects of the messages without analyzing the measurement process. It's quite frequent choose an indicator without questioning if it does represent the variable, concept or construct under study. This paper discusses the need to ensure the construct validity on the use of indicators applied social sciences, particularly in the field of communication. Taking the COARSE theory, proposed by Rossiter (2011), which criticizes the traditional psychometric model, we analyses the keys on the measurement process and we propose an inter-rater evaluation process with independent judges that would ensure the construct validity of the indicators. It also discusses the importance of ensuring the validity of the response scale to fit as best as possible the type of user response.

Keywords: measurement process, measurement theory in social science, content validity and answer validity.

Investigar la Comunicación hoy
Revisión de políticas científicas y aportaciones metodológicas