

Análisis de los elementos implícitos en la validación de contenido de un instrumento de investigación

Analysis of the elements implicit in the validation of the content of a research instrument

Luis Gibran JUÁREZ-HERNÁNDEZ ¹; Sergio TOBÓN ²

Recibido: 31/05/2018 • Aprobado: 15/07/2018 • Publicado: 11/11/2018

Contenido

1. Introducción
2. Metodología
3. Resultados
4. Conclusiones

Referencias bibliográficas

RESUMEN:

El proceso de determinación de validez de contenido de un instrumento de investigación articula elementos inherentes como son la connotación de expertos, su selección, y el tipo de evaluación que estos realizan al instrumento. Con base en los resultados se delimita que la selección de expertos debe basarse en métodos y criterios formales, la evaluación debe conjuntar un enfoque cuali-cuantitativo, y que el análisis de la evaluación cuantitativa se efectúe con coeficientes pertinentes.

Palabras clave: Instrumento, validez de contenido, experto, juicio de expertos, V de Aiken

ABSTRACT:

The process of determining the content validity of a research instrument articulates inherent elements such as the connotation of experts, the selection of them, and the type of evaluation they perform on the instrument. Based on the results, it was defined that the selection of experts should be well-founded on formal methods and criteria; the evaluation should combine a qualitative-quantitative approach; and the analysis of the quantitative evaluation should be carried out with appropriate coefficients.

Keywords: Instrument, content validity, expert, expert judgment, Aiken's V



1. Introducción

Clark-Carter (2002) refieren que la medición se conceptualiza como la asignación de una escala numérica a objetos y/o eventos a las unidades de análisis. Estas unidades de análisis requieren de la recolección de información, mediante la cual se pueden describir y/o predecir, siendo los instrumentos de evaluación el método de mayor proximidad. Los instrumentos de investigación son herramientas que permiten obtener evidencia, y cuando ésta se provee en términos de unidades de medida entonces el instrumento es catalogado como científico (Gamero et al., 2016). Específicamente, Herrera (1998) menciona que un "instrumento es la técnica o conjunto de técnicas que permiten la asignación numérica a las magnitudes de la propiedad o atributo, pudiendo ser por comparación con las unidades de medida y/o para provocar, cuantificar las manifestaciones del atributo cuando éste es medible sólo de manera indirecta (por las manifestaciones o consecuencias que se consideran producto del atributo)" (p. 16).

Mendoza-Mendoza y Garza (2009) refieren que las escalas son instrumentos de medición que agrupan un conjunto de ítems, los cuales permiten identificar niveles de las variables teóricas, las cuales no son directamente observables. Por lo tanto, cuando se efectúa la medición de manera indirecta, se necesita obtener evidencia robusta de la relación existente entre lo que realmente se está midiendo y el atributo que se supone se mide (cuanto más fuerte la evidencia, más precisa será la medición) (Kerlinger, & Lee, 2002). El instrumento deberá presentar una calidad óptima, ya que de esta manera se puede asegurar la obtención de evidencias que sean válidas y confiables (Zuñiga-Gonzalez & Cardenas-Aguilera, 2014). En este sentido, la vía usual para apreciar la calidad del instrumento es la consulta a expertos, específicamente referido como la evaluación de un instrumento mediante el procedimiento denominado como juicio de expertos (Sirecci, 1998a). En este orden, Ruiz (1998) indica que el juicio de expertos es representado por un número de personas con un alto grado de conocimientos sobre una temática, los cuales analizan y evalúan un instrumento. Como puede observarse, esta definición articula una serie de elementos, los cuales se requiere clarificar, de estos destacan los criterios para determinación e identificación de los expertos respecto al grado de conocimiento, número de ellos necesario, así como los métodos de evaluación que efectúan estos a un instrumento de investigación.

Sobre la base de lo antes expuesto surge el propósito del presente estudio, el cual consistió en efectuar un análisis del término validez de contenido, su proceso de determinación y los elementos inherentes de este proceso, en el afán de brindar pautas orientativas en la evaluación de contenido de un instrumento de investigación.

2. Metodología

Se efectuó un análisis documental respecto al término validez de contenido y los elementos relacionados con el proceso de validación de un instrumento de investigación. La información recabada procedió de bases de datos pertinentes (Sciendirect, Scielo, Latindex, Google Académico), seleccionando principalmente artículos de revistas indexadas y libros en español e inglés que incluyeran los términos validez de contenido, juez, juicio de expertos. Es relevante mencionar que el periodo de tiempo que se consideró para efectuar la búsqueda fue amplio, ya que se incluyeron trabajos clásicos de conceptualización así como trabajos de aplicación.

3. Resultados

Validez de contenido

La calidad y precisión de un instrumento de investigación se refiere como la validez de contenido, y esta se relaciona con la obtención de evidencias válidas. Cronológicamente la definición se ha mantenido estable (Tabla 1). De las contribuciones enlistadas, destaca la aportada por Haynes et al. (1995) quienes refieren que la validez de contenido es el grado en cual los elementos de un instrumento de evaluación son pertinentes y representativos de un constructo objetivo destinada a una finalidad específica de evaluación. De alta valía es la propuesta de estos autores, ya que además de brindar una definición, efectúan un análisis a detalle de cada elemento que integra la misma.

Tabla 1
Progresión temporal del término validez de contenido

Autor	Definición

Bohrnstedt (1976)	La validez del contenido es el grado en que la medición representa el concepto medido
Nunnally (1978)	La validez del contenido de un instrumento depende de la adecuación de un dominio específico del contenido muestreado
Magnusson (1982)	Se trata de determinar hasta dónde los ítems de un instrumento son representativos del dominio o universo de contenido de la propiedad que se desea medir
Bush (1985)	El grado en que el instrumento cubre el contenido que se supone debe de medir
Kerlinger (1986)	Es la representatividad del contenido
Mitchell (1986)	Se refiere a la relevancia y representación del atributo en los ítems de un instrumento de medición para un propósito particular
Haynes et al. (1995)	Es el grado en cual los elementos de un instrumento de evaluación son pertinentes y representativos de un constructo objetivo destinada a una finalidad específica de evaluación.
Ruiz (1998)	El grado en que un instrumento realmente mide la variable que pretende medir.
Wynd, Schmidt y Schafer (2003)	La evidencia necesaria para determinar el grado en que un instrumento muestre adecuadamente el dominio de investigación de interés
Sireci (2003)	Las principales evidencias de validez de contenido son la definición del dominio y su representación. La representatividad indica la adecuación con que el contenido del test representa todas las facetas del dominio definido, mientras que la relevancia alude al grado en que cada ítem del test mide el dominio definido.
Martín-Arribas (2004)	El grado en que un instrumento de medida mide aquello que realmente pretende medir o sirve para el propósito para el que ha sido construido.
Polit y Beck (2006)	la medida en que un instrumento de evaluación contiene una muestra adecuada de elementos para el constructo evaluado
Thomas y Nelson (2007)	el grado en que una prueba representa de forma adecuada lo que se ha realizado
Prieto y Delgado (2010)	Es la justificación de que los ítems para medir el criterio son una muestra representativa del contenido a evaluar.
Jaramillo y Osses (2012)	Es el grado en que la medición representa al concepto medido
Robles-Garrote y Rojas (2015)	Los ítems o aspectos elegidos para la elaboración del instrumento de medición son indicadores de lo que se pretende medir
Parrado-Lozano et al. (2016)	El grado en que el instrumento representa la totalidad del fenómeno que se pretende medir.
Koller, Levenson y Glück (2017)	La validez del contenido es determinada por la validez y representatividad en la definición del constructo, representatividad del grupo de ítems, aspectos gramaticales de los ítems y claridad de las instrucciones.

Juicio de expertos

Ruiz (2002) refiere que el procedimiento comúnmente empleado para determinar la validez de contenido es denominado "Juicio de expertos", el cual se ha referido como solicitar a un conjunto de personas la demanda de un juicio hacia un objeto, un instrumento, un material de enseñanza, o su opinión respecto a un aspecto concreto (Caballero & Llorente, 2013). Específicamente para instrumentos de investigación, la consulta a expertos es la vía más usual para apreciar la calidad del contenido de este (Sireci, 1998). Ruiz (1998) indica que el juicio de expertos es representado por personas con un alto grado de conocimientos sobre una temática, y quienes analizan un instrumento con el propósito de estudiar la exactitud con que puede hacerse medidas significativas y adecuadas con el mismo, y que mida el rasgo que se pretende medir. Concretamente Mendoza-Mendoza y Garza (2009) mencionan que se acude a un grupo de expertos con el objetivo de establecer los aspectos relevantes del dominio, de manera que en este proceso se consideren ítems no incluidos o se eliminen los considerados como no relevantes. Por su parte, Abad et al. (2011) refieren que el juicio de expertos se caracteriza por contar con un número de expertos, los cuales proponen los ítems o dimensiones que deben conformar el constructo de interés o evalúan los diferentes ítems en función de su relevancia y representatividad.

El juicio de expertos conceptualizada como estrategia de evaluación brinda la posibilidad de obtener una amplia y pormenorizada información sobre el objeto de estudio y la calidad de las respuestas por parte de los jueces (Cabero & Llorente, 2013). En concordancia con lo anterior se refiere que el juicio de expertos presenta una serie de ventajas, como lo son la teórica calidad de la respuesta que se obtiene del experto, el nivel de profundización de la valoración que se ofrece, su facilidad de puesta en acción, la no exigencia de grandes requisitos técnicos y humanos para su ejecución. (Barroso & Cabero, 2011; Cabero, 2001). Beck y Gable (2001) así como Mastaglia et al. (2003) precisan que el objetivo básico del juicio es la evaluación que indique en qué medida los elementos creados son representativos del constructo objetivo y el grado en que estos elementos representan la faceta del constructo para el que fueron desarrollados, es decir, su relevancia. Se trata de una técnica cuya realización adecuada constituye a veces el único indicador de validez de contenido del instrumento de investigación (Escobar-Pérez & Cuervo-Martínez, 2008), de ahí que resulte de gran utilidad en la valoración de aspectos de orden cualitativo (Robles-Garrote & Rojas, 2015).

Experto

De acuerdo con lo anterior, tres aspectos fundamentales destacan del juicio de expertos, siendo el concepto de experto, determinación del grado de conocimiento en el área o constructo y número de ellos necesarios para efectuar la evaluación del instrumento. Respecto a la definición de experto, de manera general se indica como aquella persona práctica o experimentada en algo o dicho de una persona especializada o con grandes conocimientos en una materia. (RAE, 2017). De mayor especificidad, Landeta (2002) refiere que el experto es un individuo cuya situación y recursos personales le posibilitan contribuir positivamente a la concesión del fin que ha motivado. Crespo (2007) refiere que se entiende por experto a una persona, grupo de ellas u organizaciones idóneas de ofrecer con un máximo de competencia, valoraciones conclusivas sobre un determinado problema. Por su parte, Escobar-Pérez y Cuervo (2008), define al experto como aquella persona que puede emitir una opinión o juicio que se basa en su propia historia en una línea de investigación. Mengual (2011) define al experto como el individuo o conjunto de personas que son capaces de proporcionar valoraciones fiables

sobre un problema en cuestión. Con el objetivo de brindar a los jueces, Abdolhammedi y Shanteau (1992) enlistan más 20 aspectos que caracterizan a un experto, siendo experiencia, conocimiento actual, metódico, creatividad, asunción de disponibilidad, como los aspectos de mayor relevancia, sin embargo, se incluyen otros aspectos (aparición física, cálido y amigable), que pudieran ser considerados aspectos relativos y controversiales. Posteriormente, Brill et al. (2006) así como García y Fernández (2008) refieren como características relevantes del experto su vinculación con el problema, experiencia profesional, cualidades personales para participar en las investigaciones y pericia profesional. Como puede observarse, el concepto de experto tiene múltiples connotaciones, siendo altamente polisémico y poco clarificador (Cabero-Almenara & Infante-Moro, 2014).

Caracterización y selección del experto

La elección de expertos suele ser un tema de amplia discusión, ya que por su aspecto polisémico resulta compleja la estandarización de procesos de elección, así como es dependiente de los objetivos del investigador (Scapolo & Miles, 2006). Cruz-Ramírez y Martínez-Cepena (2012) así como Cabero-Almenara e Infante-Moro (2014) refieren que algunos criterios a considerar para su elección son el conocimiento y experiencia que tengan en la temática, experiencia profesional, voluntad de querer participar en el estudio, su disponibilidad de tiempo, comprometerse a la participación en todas las rondas que se establezcan, años de experiencia en la temática concreta, y su capacidad de comunicación efectiva. Además de estos criterios, Ténrière-Buchot (2001) denota la existencia de tres tipos de expertos denominados tácticos, conciliadores y comunicadores, siendo los primeros los que se seleccionan de acuerdo con el grado de experiencia y especialización en el tema de interés. En los segundos se apuesta por el equilibrio, imparcialidad y sentido común, y los terceros suelen estar implicados en el objeto de investigación y su percepción aporta criterios sobre viabilidad, contextualización, pertinencia, entre otros aspectos. Por su parte, Landeta (2002) refiere que existen dos tipos de expertos (especialistas y afectados), siendo los primeros los que poseen conocimiento científico y experiencia sobre la temática objeto de estudio, mientras que los segundos son los que se encuentran implicados de alguna forma en el área de estudio concreta. Además de su identificación basada en los criterios anteriormente descritos, Cabero-Almenara y Llorente-Cejudo (2013) refieren que el proceso de selección dependerá de una serie de aspectos, que van desde la rapidez con la que el evaluador desee llevar a cabo el estudio, la capacitación del evaluador, la eficacia y profundización de los resultados que desee alcanzar, el esfuerzo que el evaluador o investigador desee invertir en el proceso por mencionar algunos. Estos autores mencionan que los procedimientos de selección van desde aquellos que no involucran algún tipo de estructuración o criterio de selección (como puede ser la cercanía o afinidad al evaluador o investigador), hasta aquellos que poseen estructura e involucran la selección a través de la aplicación de diversos criterios.

Diversos procedimientos se han referido para efectuar la selección de expertos, destacando principalmente el biograma y el coeficiente de competencia experta (García & Fernández, 2008). Cabero-Almenara y Llorente-Cejudo (2013) especifican que el biograma es la elaboración de la biografía del experto, señalándose aspectos como lugar donde trabaja, años de experiencia, actividades desarrolladas, acciones formativas llevadas a cabo, experiencia en investigación, años de trabajo, lugares donde ha trabajado, con la finalidad de que se obtenga suficiente información que permita justificar la selección del experto, infiriendo la adecuación y pertinencia para la actividad solicitada. Respeto al coeficiente de competencia experta, se indica que se efectúa la identificación a partir de la autovaloración que el experto realiza en diferentes aspectos e indicadores, mediante los cuales se establece un valor que es utilizado por el investigador para seleccionar los expertos que pueden ser las más adecuadas para intervenir en la evaluación (García & Fernández, 2008; Mengual, 2011; Pedrosa et al., 2013; Zayas, 2011). Como unidad crítica se ha establecido, que aquellos expertos que obtengan una puntuación menor a 0,8, no serán contemplados para el estudio (Cabero-Almenara & Infante-Moro, 2014).

Juárez-Hernández et al. (2017) proponen una aproximación basada en el biograma, en el cual se incluye variables cuantitativas referentes al número de años de experiencia en el área, así como en el diseño, revisión y validación de instrumentos de evaluación. Con la comparación de estas variables empleando el diagrama de Olmstead y Tukey (Sokal & Rohlf, 1981), es factible identificar por cuadrantes a los expertos que se consideran como óptimos para el juicio de expertos, denotando que los expertos con un mayor número de años de experiencia en el área, así como en la revisión de instrumentos se ubicaría en el cuadrante superior derecho. Además de esta propuesta, estos autores señalan que el empleo de variables de tipo cualitativas codificadas (i. e. nivel académico) en conjunción a las variables cuantitativas mencionadas, permitirían la identificación y caracterización de grupos de expertos mediante un análisis multivariado de clasificación (análisis de conglomerados).

Diversos autores han mencionado la importancia de la elección de los expertos, por lo tanto, su selección representa una fase fundamental. A este respecto Carlos (2009) refiere que la importancia de la selección de expertos radica en que, si se desean conseguir resultados útiles para los objetivos que se persiguen, se deberá de poner énfasis especial en la selección de los mismos, y para ello, la combinación del biograma y el índice de competencia experta se presenta como una opción adecuada. Rubio et al. (2003) destacan que se debe de efectuar una selección adecuada y representativa de los expertos para garantizar una evaluación pertinente y exhaustiva. Por su parte, Blasco et al. (2010) indican que la calidad de los resultados que se consigan en un estudio donde se ejecute el juicio de expertos, está completamente relacionada con los expertos seleccionados, por lo tanto, la utilización de un buen procedimiento de selección es uno de los aspectos al que se debe de prestar mayor atención.

Número de expertos

Otro aspecto fundamental en el juicio de expertos es el número de ellos necesario. En este orden, de acuerdo con Williams y Webb (1994) así como Powell (2003), mencionan que no existe un consenso respecto al número de expertos a emplear. En este sentido, diversos autores refieren que se recomienda un número mayor de diez expertos (García et al., 2016, Hyrkäs et al., 2003, Jiménez et al., 2013; Mills et al., 2012; Wiersma, 2001; Zhu, Ennis & Chen, 1998), ya que con este número mínimo se aporta una estimación aceptable de validez de contenido (Robles et al., 2016), facilitando la detección y exclusión de valores atípicos del evaluador (Carmines & Zeller, 1979; Lynn, 1986).

Una vez definidos los expertos y el número de estos que estarán involucrados en la evaluación de la validez de contenido, se determinara la forma o el método mediante el cual se efectuara el juicio de expertos. Cabrero-Almenara y Llorente-Cejudo (2013) refieren que se ha establecido cuatro propuestas para llevar a cabo el juicio de expertos (Agregación individual, método Delphi, técnica grupal nominal y método de consenso). De estas destacan los dos primeros métodos, ya que logísticamente posibilitan la participación de expertos de otras localidades y regiones.

Evaluación cualitativa

Referente al proceso de realización de la evaluación, es de vital importancia como efectuaran este proceso los jueces. Krippendorff (1990) indica que cualitativamente la validez de contenido se deriva de la estructuración exhaustiva del dominio, del grado de pertinencia que las categorías y las unidades definidas tienen en relación tanto al objeto de estudio como a los objetivos de la investigación. Por lo tanto, el análisis cualitativo del juicio de expertos se refiere al analizar a detalle la información y sugerencias proporcionadas por los expertos, dentro de estas se delimita la inclusión o exclusión de ítems, estructura de indicaciones, creación de nuevos ítems, creación de nuevas dimensiones. Wiersema (2001) menciona que la evaluación cualitativa permite ajustar el contenido y redacción de los ítems, tanto en el aspecto morfológico como sintáctico del enunciado de las preguntas y respuestas.

Evaluación cuantitativa

Sireci (1998b) establece que, a la par del análisis cualitativo de los expertos, es vital que estos aporten una valoración cuantitativa a los ítems, ya que el hecho de que contar con expertos que informen sobre la falta o exceso de ítems representativos del constructo o que determinen a qué dimensión corresponde cada elemento, no aporta de por sí información relevante para el proceso de validación. Por lo tanto, Haynes et al. (1995) así como Kimberly y Winterstein (2008) refieren que el estudio de la validez de contenido debe ser un proceso multimétodo, involucrando la evaluación cualitativa y cuantitativa y el análisis de los resultados de estas evaluaciones.

Diversos autores coinciden que cuantitativamente la validez de contenido es determinada por el grado en que la medida obtenida representa el concepto en su conjunto y de cada uno de sus componentes o dimensiones (Abad et al., 2011; Furr & Bacharach, 2008; García-Sedeño & García-Tejero, 2014; Martínez-Arias, 1996; Sireci, 1998a). Ante este hecho, Kerlinger y Lee (2002), refieren que la validez de contenido es cuantificable a través de índices de relevancia entre las evaluaciones de los jueces por lo cual es necesario emplear una escala de evaluación numérica (escala Likert). Esta escala puede contemplar una amplitud de uno a cinco (Mussio & Smith, 1973), de uno a tres (Hambleton, 1980), o bien de uno a cuatro (Davis, 1992), y es aplicable para calificar aspectos como relevancia, utilidad, importancia (Drauden & Peterson, 1974), representatividad, comprensión, ambigüedad, claridad (Davis, 1992), relevancia, representatividad, especificidad, claridad (Haynes et al., 1995) de los ítems que integran el instrumento.

Análisis cuantitativo

Existen diversas propuestas para efectuar el análisis cuantitativo de validez de contenido, los cuales van desde métodos básicos (i. e. promedio o

media), hasta de mayor complejidad (i. e Rango interpercentil ajustado la simetría (Fitch, et al., 2001; Índice promediado de la desviación media (Claeys et al., 2012), Índice de validez factorial (Rubio et al., 2003)). En el presente estudio, se realiza énfasis en coeficientes de mayor difusión y que de manera general no suponen un esfuerzo para su cálculo e interpretación. De los procedimientos de mayor sencillez es el referente al cálculo de la media de las evaluaciones de los jueces para cada ítem. A este respecto, Nunnally y Bernstein, (1994) concluyen que la media y desviación estándar pueden guiar juicios sobre la validez del contenido, incluso sin criterios formalizados para su interpretación. Basados en este hecho, Escobar-Pérez y Cuervo-Martínez (2008) proponen que los ítems que posean un mayor promedio serán los que se deben de conservar, precisando que aquellos ítems que se acerquen a la mitad de la media serán sujetos a revisión y mejoramiento, coincidiendo con lo propuesto por Zambrano et al. (2015). Por su parte, Nieves et al. (2013) señalan que además del análisis de los promedios se debe de incorporar el cálculo y análisis de la desviación estándar, concluyendo que entre mayor sea esta medida de dispersión, revelará mayor diferencia de opinión entre los expertos, denotando la revisión de estos ítems.

Es relevante mencionar que el método anteriormente descrito resulta de fácil aplicación, pero no contempla variables propias del juicio de expertos como lo son el número de jueces, así como la longitud de la escala empleada. Ante este hecho, se mencionan una serie de coeficientes, los cuales incorporan las variables señaladas. Aiken (1980; 1985) propone el coeficiente de validez de contenido denominado V, el cual cuantifica la relevancia de cada ítem respecto de un dominio de contenido formulado por N jueces, oscilando desde cero hasta uno, siendo el valor de uno indicativo de un perfecto acuerdo entre los jueces respecto a la mayor puntuación de validez de los contenidos evaluados. Este coeficiente se ha mencionado como uno de los principales medios para cuantificar y evaluar el contenido o relevancia de cada elemento en relación con el contenido dominio de N juicios (Merino-Soto & Livia-Segovia, 2009). Aunado al cálculo del coeficiente, se debe de determinar la significancia estadística, para lo que se debe de consultar la tabla de valores críticos propuesta por Aiken (1985). En este orden, la significancia estadística para el coeficiente V de Aiken tiene por hipótesis nula que un valor del contraste de 0.50, que representa la variación aleatoria; siendo este nivel inaceptable para fines prácticos ya que se considera que está por debajo de las recomendaciones para la construcción de pruebas para cualificar la magnitud de un coeficiente de validez (Anastasi & Urbina, 1997; Charter, 2003; Cicchetti & Sparrow, 1981; Cicchetti, 1994; Lindley Bartram & Kennedy, 2005; Nunnally & Bernstein, 1995; Prieto & Muñoz, 2000). Respecto al valor mínimo aceptado para el coeficiente existen diversas propuestas, destacando las de Aiken (1985), quien propuso que el valor mínimo para aceptar un ítem como válido debería de ser de 0.69, Cicchetti (1994) de 0.50, Charter (2003) de 0.70, Penfield y Giaccobi (2004), Merino-Soto y Livia-Segovia (2009), proponen conservar ítems con V superior a un valor de 0.75. De acuerdo con Penfield y Giaccobi (2004) así como Bulger y Housner (2007) recomiendan eliminar ítems con una V menor 0.80, sugiriendo como ítems muy adecuados aquellos con V superior a 0.80. Penfield y Giaccobi (2004) efectuaron un análisis extensivo del coeficiente, aportando una modificación al aparato matemático y proponiendo el cálculo de los intervalos de confianza para el coeficiente.

La propuesta del cálculo de los intervalos de confianza resulta pertinente ya que como lo menciona Charter (2003) su determinación es altamente recomendada ya que estos representan una forma pertinente de expresar el grado de imprecisión o incertidumbre asociada con los resultados cuantitativos de algún cálculo en general. Específicamente, el cálculo del intervalo de confianza para la V de Aiken permite determinar si la magnitud obtenida del coeficiente es superior a una establecida como mínimamente aceptable de acuerdo con Aiken (1980, 1985, 1996); esto es, los valores mínimos sobre los que decidir qué ítems se deberían de aceptar o rechazar. Merino-Soto y Livia-Segovia (2009) indican que el cálculo de los intervalos de confianza para este coeficiente permite probar si la magnitud obtenida del coeficiente es superior a una que es establecida como mínimamente aceptable (0.70) para concluir sobre la validez de contenido de los ítems.

La Razón de Validez de Contenido (RVC) fue propuesta por Lawshe (Lawse, 1975) y es un modelo específico de cuantificación de la representatividad de los ítems. Para efectuar su cálculo, se emplea una escala de evaluación que consta de tres elementos (esencial, útil, pero prescindible e innecesario). El RVC oscila entre -1 y +1, siendo las puntuaciones positivas las que indican una mejor validez de contenido. Lawse (1975) propone los valores mínimos del RVC de acuerdo con el número de expertos empleados, destacando que los valores mínimos aceptables para un número reducido de expertos suelen ser demasiado exigentes (cerca de uno), mientras que cuando se emplea un gran número de jueces, los valores tienden a ser demasiado escuetos (cerca de 0.40). Una ventaja de este coeficiente es el cálculo del índice de validez de contenido (IVC), que es referido como el promedio de los ítems que pasaron el supuesto del valor mínimo aceptable. Davis (1992) propone un valor superior a 0,80 para definir el conjunto de ítems como adecuado. Tristán (2008) efectuó una modificación al modelo original de Lawse para el RVC, y propone que la Razón de Validez de Contenido (RVC'), la cual define que, para aceptar un ítem, este debe de presentar un valor superior a 0.5823. La modificación que efectúa obedece a que los valores mínimos de RVC, se ven afectados por el número de expertos. Polit, Beck y Owen (2007) refieren diversas ventajas de este coeficiente destacando su facilidad de cálculo, facilidad de interpretación, el aporte de información tanto a nivel de ítem como de instrumento. Estos autores destacan que este coeficiente no es aplicable en juicios con menos de cinco expertos.

Hernández-Nieto (2002) propuso el Coeficiente de validez de contenido (CVC), el cual permite valorar el grado de acuerdo de los expertos. En este sentido, de los lineamientos para emplear este coeficiente se recomienda la participación de entre tres y cinco expertos y se debe de emplear una escala tipo Likert de cinco alternativas. Su cálculo se efectúa con la media (promedio) para uno de los ítems entre la puntuación máxima que el ítem podría alcanzar. Además de este cálculo, se debe calcular el error asignado a cada ítem (Pei), reduciendo el posible sesgo introducido por alguno de los jueces, Por lo tanto, el CVC final se calcularía aplicando el CVC inicial menos error asignado a cada ítem ($CVC = CVC_i - Pei$). Hernández-Nieto (2002) recomienda mantener únicamente aquellos ítems con un coeficiente superior a 0.80, mientras que Balbinotti (2004) indica conservar aquellos ítems con un coeficiente superior a 0.70.

Polit y Beck (2006) propusieron el índice de validez de contenido (IVC), el cual ha sido usado para estimar la representatividad, comprensión, ambigüedad y claridad para los diferentes ítems que integran un instrumento. Para su empleo, se requiere usar una escala Likert de uno a cuatro, siendo cuatro la máxima valoración que pueden emitir los jueces. El cálculo para cada ítem se efectúa dividiendo el número de jueces que evaluaron el ítem con tres a cuatro entre el número total de jueces. Es importante señalar que este índice es que permite el cálculo de del índice global de validez de contenido, refiriendo que este se calcula con el promedio de los ítems que fueron conservados, los cuales deben de presentar un valor superior a 0.80.

4. Conclusiones

Como se observó, la validación de contenido de un instrumento articula una serie de elementos que requieren de una atención especial. En este sentido, la caracterización y selección de expertos que integraran el juicio, el empleo de una escala que permita la evaluación cuantitativa y el análisis de los resultados de esta evaluación mediante coeficientes pertinentes para tal efecto, representan fases de alta atención y de aplicación de criterios o métodos formales. Por lo anterior, se propone que la selección de los expertos está fundamentada bajo procesos estructurados (biograma, coeficiente de competencia experta, empleo de variables cuantitativas) que denoten su pertinencia y experiencia en el constructo a evaluar, considerando un número mínimo de 10 expertos para efectuar el juicio de expertos, el empleo de una escala numérica para evaluar la pertinencia, relevancia y redacción de los ítems y finalmente el uso de coeficientes, en los que los criterios para conservar o desechar ítems están previamente delimitados. Respecto a este punto, se recomienda el empleo del coeficiente V de Aiken y sus intervalos de confianza, en el afán de brindar certidumbre respecto al valor obtenido.

El hecho de no considerar estos aspectos (selección de los jueces, empleo de una escala numérica de evaluación y el análisis de la evaluación cuantitativa), podría determinar la inclusión de ítems que no presentan pertinencia ni relevancia para el constructo a medir, determinando la generación de conclusiones equivocadas (Mendoza & Garza, 2009). A este respecto Prieto y Delgado (2010) indican que "difícilmente se podrá justificar la calidad de las medidas si los ítems no representan de forma suficiente las diferentes facetas de las manifestaciones del constructo. Si eso sucede, el constructo estará infrarrepresentado. La evidencia de que las respuestas a los ítems están influidas por variables ajenas al constructo de interés constituye una de las principales amenazas a la validez produciendo la denominada varianza irrelevante al constructo" (p. 72). Por su parte, Yaghmale (2003) destaca el empleo de una escala de evaluación para evaluar la pertinencia y relevancia de los ítems del instrumento por parte de los jueces y el empleo de coeficientes pertinentes para un adecuado análisis cuantitativo son procesos fundamentales, ya que declaraciones tales como: "la validez del instrumento había sido probado por los comentarios de los expertos" o "validez del contenido se determinó a través de una revisión de la literatura o los expertos del panel" actualmente son inaceptables e inválidas.

Es fundamental mencionar que la validez de contenido es un componente importante de la validez de constructo, porque provee evidencia respecto al grado en el cual los elementos de un instrumento de evaluación son relevantes y pertinentes del constructo objetivo, denotando que la validez de contenido proporciona información sobre los datos obtenidos de un instrumento de evaluación y las inferencias que se pueden extraer de esos datos (Hayness et al., 1995; Mendoza & Garza, 2009).

Referencias bibliográficas

Abad, F. J., Olea, J., Ponsoda, V., & García, C (2011). *Medición en ciencias sociales y de la salud*. Madrid, España: Síntesis

Abdolmohammadi, M., & Shanteau, J. (1992). Personal Attributes of Expert Auditors. *Organizational Behavior and Human Decision Processes*, 53(2),

- Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement* 40, 955-959. Recuperado de: <https://doi.org/10.1177/001316448004000419>
- Aiken, L. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*. 45(1), 131-142. Recuperado de: <https://doi.org/10.1177/0013164485451012>
- Aiken, L. R. (1996). *Tests psicológicos y evaluación*. México: Prentice Hall.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed). New York: McMillian.
- Barroso-Osuna, J., Cabero-Almenara, J., & Vázquez-Martínez, A. I. (2012). Formación desde la perspectiva de los entornos personales de aprendizaje. *Revista de Innovación educativa*, 4(1). Recuperado de: http://www.udgvirtual.udg.mx/apertura/num16/1_formacion.htm
- Beck, C.T., & Gable, R.K. (2001). Ensuring content validity: An illustration of the process. *Journal of Nursing Measurement*, 9, 201-215
- Bohrstedt, G.W. (1976). *Evaluación de la confiabilidad y validez en la medición de actitudes*. México D.F.: Editorial Trillas.
- Blasco, J. E., López, A., & Mengual, S. (2010). Validación mediante el método Delphi de un cuestionario para conocer las experiencias e interés hacia las actividades acuáticas con especial atención al windsurf. *Ágora para la educación física y el deporte*, 12(1), 75-96.
- Brill, J. M., Bishop, J., & Walker, A. (2006). The competencies and characteristics required of an effective project manager: A web-based Delphi study. *Educational Technology Research*, 54(2), 115-140. Recuperado de: doi: 10.1007/s11423-006-8251-y
- Bulger, S.M., & Hurner, L. D. (2007). Modified delphi investigation of exercise science in physical education teacher education. *Journal of Teaching in Physical Education*. 26, 57-80. Recuperado de: <https://doi.org/10.1123/jtpe.26.1.57>
- Cabero-Almenara, J. (2001). *Tecnología educativa. Diseño y utilización de medios en la enseñanza*. Barcelona: Paidós
- Cabero-Almenara, J., & Barroso-Osuna, J. (2013). La utilización del juicio de experto para la evaluación de TIC: el coeficiente de competencia experta. *Bordón, Revista de Pedagogía*, 65(2), 25-38. Recuperado de: <http://recyt.fecyt.es/index.php/BORDON/article/view/brp.2013.65202/11419>
- Cabero-Almenara, J., & Infante-Moro, A. (2014). Empleo del método Delphi y su empleo en la investigación en Comunicación y Educación. *Revista Electrónica de Investigación Educativa*, 48, 1-16. Recuperado de: http://edutec.rediris.es/Revelec2/Revelec48/pdf/Edutece_n48_Cabero-Infante.pdf
- Cabero-Almenara, J., & Llorente Cejudo, M. C. (2013). La aplicación del juicio de experto como técnica de evaluación de las tecnologías de la información (TIC). *Revista de Tecnología de Información y Comunicación en Educación*, 7 (2) pp.11-22. Recuperado de: <http://tecnologiaedu.us.es/tecnoedu/images/stories/jca107.pdf>
- Charter, R. A. (2003). A breakdown of reliability coefficients by test type and reliability method, and the clinical implications of low reliability. *Journal of General Psychology*, 130(3), 290-304. Recuperado de: <dx.doi.org/10.1080/00221300309601160>
- Cicchetti, D.V., & Sparrow, S.S. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*, 86, 127-137. Recuperado de: <https://www.ncbi.nlm.nih.gov/pubmed/7315877>
- Cicchetti, D. V. (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessments*, 6, 284-290. Recuperado de: https://www.researchgate.net/publication/15882978_Developing_Criteria_for_Establishing_Interrater_Reliability_of_Specific_Items_Applications_to_Asses
- Clark-Carter, David (2002). *Investigación cuantitativa en Psicología*. México D. F. Editorial Oxford.
- Claeys, C., Nève, J., Tulkens, P. M., & Spinewine, A. (2012). Content validity and inter-rater reliability of an instrument to characterize unintentional medication discrepancies. *Drugs Aging*, 29, 577-591. Recuperado de: <https://doi.org/10.1007/BF03262275>
- Crespo, T. (2007). *Respuestas a 16 preguntas sobre el empleo de expertos en la investigación pedagógica*. Lima, Perú: San Marcos.
- Creswell, J. W. (2003). *Research Design: Qualitative, quantitative, and mixed method approaches*. Thousand Oaks California: Sage Publications.
- Cruz-Ramírez, M., & Martínez-Cepena, M. C. (2012). Perfeccionamiento de un instrumento para la selección de expertos en las investigaciones educativas. *REDIE. Revista Electrónica de Investigación Educativa*, 14, 167-179. Recuperado de <http://www.redalyc.org/articulo.oa?id=15525013012>
- Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, 5, 194-197. Recuperado de: https://www.researchgate.net/publication/248867629_Instrument_review_Getting_the_most_from_a_panel_of_experts
- Domingo, J., & Fernández, M. (1999). *Técnicas para el desarrollo personal y formación del profesorado*. Bilbao: Universidad de Deusto.
- Drauden, G. M., & Peterson, N. G. (1974). *A domain sampling approach to job analysis*. Test Validation Center. St. Paul: Minn.
- Escobar-Pérez, J. y Cuervo-Martínez, A. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en Medición*, 6, 27-36. Disponible en http://www.humanas.unal.edu.co/psicometria/files/7113/8574/5708/Articulo3_Juicio_de_expertos_27-36.pdf
- Fitch K., Bernstein S. J., Aguilar, M. D., Burnand, B., LaCalle, J. R., Lazaro, P., ... Kahan, J. P. (2001) *The RAND/UCLA Appropriateness Method User's Manual*: RAND corporation. California: UCLA
- Furr, R. M., & Bacharach, V. R. (2008). *Psychometrics: An introduction*. Thousand Oaks California : Sage Publications.
- Gamero, K; Flores, C; Arias, W L; Ceballos, K D; Román, A; Marquina, E; (2016). Estandarización del Test de Dependencia al Celular para estudiantes universitarios de Arequipa. *Persona*, 19, 179-200. Recuperado de: <http://www.redalyc.org/articulo.oa?id=147149810011>
- García, I., & Fernández, S. (2008). Procedimiento de aplicación del trabajo creativo en grupo de expertos, *Energética*, 29(2), 46-50. Recuperado de: <http://www.redalyc.org/articulo.oa?id=329127758006>
- García, A., Antúnez, A., & Ibáñez, S.J. (2016). Análisis del proceso formativo en jugadores expertos: validación de instrumento. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*. 16(61) 157-182. Recuperado de: <https://revistas.uam.es/rimcafd/article/view/3782/4072>
- García, L., & Fernández, S. (2008). Procedimiento de aplicación del trabajo creativo en grupo de expertos. *Energética*, 29(2), 46-50. Recuperado de <http://www.redalyc.org/articulo.oa?id=329127758006>
- Garrote, P. R., & Rojas, M. C. (2015). La validación por juicio de expertos: dos investigaciones cualitativas en Lingüística aplicada. *Revista Nebrija de Lingüística aplicada*. Nebrija, España. Recuperado de: https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_55002aca89c37.pdf
- González, N., Pirela, A., & Zerpa, M. L. (2012). La formación docente como investigador. *Una Responsabilidad Social Universitaria*. *Opción*, 28, 466-479. Recuperado de: <http://www.redalyc.org/articulo.oa?id=31025702009>
- Hambleton, R. K. (1980). Test score validity and standard setting methods. En R. A. Berk (Ed.), *Criterion-Referenced Measurement: The State of the Art*. Johns Hopkins University Press: Baltimore.
- Haynes, S. N., Richard, D. C. S., & Kubay, E. S. (1995). Content Validity in Psychological Assessment: A Functional Approach to Concepts and Methods. *Psychological Assessment*, 7(3), 238-247. Recuperado de: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.452.5453&rep=rep1&type=pdf>
- Hernández-Nieto, R. A. (2002), *Contributions to Statistical Analysis*. Mérida, Venezuela: Universidad de Los Andes.
- Hyrkäs, K., Appelqvist-Schmidlechner, K., & Oksa, L. (2003). Validating an instrument for clinical supervision using an expert panel. *International Journal of nursing studies*, 40(6), 619 -625.
- Jaramillo, S., & Osses, S. (2012). Validación de un Instrumento sobre Metacognición para Estudiantes de Segundo Ciclo de Educación General Básica. *Estudios pedagógicos (Valdivia)*, 38(2), 117-131. Recuperado de: <https://dx.doi.org/10.4067/S0718-07052012000200008>
- Jiménez, J., Salazar, W., & Morera, M. (2013). Diseño y validación de un instrumento para la evaluación de patrones básicos de movimiento. *European Journal of Human Movement*. 31, 87-97. Recuperado de: <http://www.redalyc.org/pdf/2742/274229586006.pdf>
- Juárez-Hernández, L. G., Tobón, S., & Cano-Jerónimo, A. (octubre, 2017). Caracterización y selección de expertos para la evaluación de un instrumento de investigación. Ponencia presentada en Congreso Internacional de Evaluación del Desempeño (VALORA), Cuernavaca, México.
- Kerlinger, F.N. (1986). *Foundations of behavioural research* (3rd ed), New York: Holt, Rinehart and Winston.
- Kerlinger, F.N., & Lee, H.B. (2002) *Investigación del Comportamiento: Métodos de Investigación en ciencias sociales*. México: McGraw-Hill Interamericana Editores.
- Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health System*

- Krippendorff, K. (1990). Metodología de análisis de contenido. Teoría y Práctica. Barcelona, España: Paidós.
- Landeta, J. (2002). El método Delphi: una técnica de previsión del futuro. Barcelona: Ariel.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28, 563-575. Recuperado de: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.460.9380&rep=rep1&type=pdf>
- Lindley, P., Bartram, D., & Kennedy, N (2005) EFPA review for the description and evaluation of psychological test: Test review form and notes for reviewers (version 3.41). Report of the Standing Committee on Test and Testing to General Assembly. Granada: EFPA.
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35, 382-385.
- Magnusson, D. (1982). Teoría de los test. México: México, D. F.: Trillas.
- Martín-Arribas, M. C. (2004). Diseño y validación de cuestionarios. *Matronas Profesión*, 5(17), 23-29. Recuperado de: http://enferpro.com/documentos/validacion_cuestionarios.pdf
- Martínez-Arias, R. (1996). Psicometría: Teoría de los tests psicológicos y educativos. Madrid: Editorial Síntesis.
- Mastaglia, B., Toyne, C., & Kristjanson, L.J. (2003). Ensuring content validity in instrument development: Challenges and innovative approaches. *Contemporary Nurse*, 14, 281-291. Recuperado de: 10.5172/conu.14.3.281
- Mendoza, J., & Garza, J.B. (2009). La medición en el proceso de investigación científica: Evaluación de validez de contenido y confiabilidad. *Innovaciones de negocios*, 6(11), 17-32. Recuperado de: <http://eprints.uanl.mx/12508/1/A2.pdf>
- Mengual, S. (2011). La importancia percibida por el profesorado y el alumnado sobre la inclusión de la competencia digital en educación Superior (Tesis doctoral). Universidad de Alicante., España.
- Merino, C., & Livia, J. (2009). Intervalos de confianza asimétricos para el índice la validez de contenido: Un programa Visual Basic para la V de Aiken. *Anales de Psicología*, 25(1), 169-171. Recuperado de: <http://www.redalyc.org/articulo.oa?id=16711594019>
- Mills. A., Butt. J., Maynard. I., & Hardwood. C. (2012). Identifying factors perceived to influence the development of elite youth football academy players. *Journal of Sport Sciences*, 30(15). 1593-1604. Recuperado de: <http://dx.doi.org/10.1080/02640414.2012.710753>
- Mitchell, J. V. (1986). Measurement in the larger context: Critical current issues. *Professional Psychology: Research and Practice*, 17, 544-550.
- Mussio, S. J., & Smith, M. K. (1973). Content validity: A procedural manual. Chicago: International Personnel Management Association.
- Nieves, I., Mayora, C., & Ojeda, V. (2013). Validación por juicio de expertos de un prototipo de software de lectura extensiva en inglés asistida por el computador. *Investigación y Postgrado*, 28(2), 9-35. Recuperado de: http://www.scielo.org.ve/scielo.php?script=sci_arttext&pid=S131600872013000200002&lng=es&tlng=es.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.
- Nunnally, J. C., & Bernstein, I. J. (1995). Teoría psicométrica (3ra ed.). México, D. F: McGraw-Hill
- Parrado-Lozano, Y. M., Sáenz-Montoya, X., Soto-Lesmes, V. I., Guáqueta-Parada, S. R., Amaya-Rey, P., Caro-Castillo, C. V., Parra-Vargas, M., & Triana-Restrepo, M. C. (2016). Validez de dos instrumentos para medir la relación interpersonal de la enfermera con el paciente y su familia en la unidad de cuidado intensivo. *Investigación en Enfermería: Imagen y Desarrollo*, 18(1) 115-128. Recuperado de: <http://www.redalyc.org/articulo.oa?id=145243501008>
- Penfield, R. D., & Giacobbi, P. R. (2004) Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science*, 8(4), 213-225.
- Polit, D.F., & Beck, C.T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29, 489-497.
- Polit, D.F., Beck., C.T., & Owen, S. (2007). Is the CVI an acceptable indicator of content validity?. *Research in Nursing & Health*, 30, 459-467.
- Powell, C. (2003). The Delphi technique: myths and realities. *Journal of Advanced Nursing*, 41(4), 376-382. Recuperado de: https://www.researchgate.net/publication/10905686_The_Delphi_technique_Myths_and_realities
- Prieto, G., & Delgado, A. R. (2010). Fiabilidad y validez. *Papeles del Psicólogo*, 31, 67-74. Recuperado de: [Recuperado de http://www.redalyc.org/articulo.oa?id=77812441007](http://www.redalyc.org/articulo.oa?id=77812441007)
- Prieto, G., & Muñiz, J. (2000). Un modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 77, 65-71. Recuperado de <http://www.redalyc.org/articulo.oa?id=77807709>
- Rubio, D. M., Berg-Weber, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research*, 27(2), 94-104.
- Ruiz, C. (1998). Instrumentos de Investigación Educativa. Venezuela: Barquisimeto, CIDEG.
- Ruiz, C. (2002). Instrumentos de Investigación Educativa. Procedimientos para su Diseño y Validación. Tipografía y Litografía Horizontes, C.A. Venezuela.
- Scapolo, F., & Miles, I. (2006). Eliciting experts' knowledge: A comparison of two methods. *Technological Forecasting & Social Change*, 73, 679-704. Recuperado de: <http://dx.doi.org/10.1016/j.techfore.2006.03.001>
- Sireci, S. G. (1998a). The construct of content validity. *Social Indicators Research*, 45, 83-117.
- Sireci, S. G. (1998b). Gathering and analyzing content validity data. *Educational Assessment*, 5, 299-321. Recuperado de: http://dx.doi.org/10.1207/s15326977ea0504_2
- Sireci, S. G. (2003). Validity content. En: *Encyclopedia of psychological assessment*. Ballesteros, R. F. (Ed.) Londres, UK: Sage.
- Sokal, R. R., & Rohlf, F. J. (1981). *Biometry: The principles and practice of statistics in biological research*. San Francisco: W.H. Freeman.
- Ténière-Buchot, P. F. (2001). Décision, expertise, arbitraire et transparence: éléments d'un développement durable. *Le courrier de l'environnement de l'INRA*, 44, 41-52. Disponible en: <http://www7.inra.fr/lecourrier/assets/C44Teniere.pdf>
- Thomas, J.R., & Nelson, J.K. (2007). Métodos de investigación en actividad física. Barcelona: Paidotribo.
- Tristan, A. (2008). Modificación al modelo de Lawshe para el dictamen de validez de contenido de un instrumento objetivo. *Avances en Medición*. 6(1), 37-48. Recuperado de: http://www.humanas.unal.edu.co/psicometria/files/8413/8574/6036/Articulo4_Indice_de_validez_de_contenido_37-48.pdf
- Wiersma. L. D. (2001). Conceptualization and development of the sources of enjoyment in youth sport questionnaire. *Measurement in Physical Education and Exercise Science*, 5(3), 153-157.
- Wynd, C.A., Schmidt, B., & Schaefer, M.A. (2003). Two quantitative approaches for estimating instrument content validity. *Western Journal of Nursing Research*, 25(5), 508-518. Recuperado de: <https://doi.org/10.1177/0193945903252998>
- Williams, P. L., & Webb, C. (1994). The Delphi technique: a methodological discussion. *Journal of Advanced Nursing*, 19: 180-186. Recuperado de: [doi:10.1111/j.1365-2648.1994.tb01066.x](https://doi.org/10.1111/j.1365-2648.1994.tb01066.x)
- Wiersma. L. D. (2001). Conceptualization and development of the sources of enjoyment in youth sport questionnaire. *Measurement in Physical Education and Exercise Science*. 5(3). 153-157. Recuperado de: http://dx.doi.org/10.1207/S15327841MPEE0503_3
- Yaghmale, F. (2003). Content validity and its estimation. *Journal of Medical Education*, 3(1): 25-7. Recuperado de: <http://sw.sid.ir/FileServer/JE/1014200301S05>
- Zambrano, R., Gil, N., Lopera, E., Carrasco, N., Gutiérrez, A., & Villa, A. (2015). Validez y confiabilidad de un instrumento de evaluación de estrategias docentes en la práctica de la enseñanza universitaria. *Magister*, 27(1), 26-36. doi:10.1016/j.magis.2015.06.002
- Zayas, P. (2011). El proceso del análisis y la descripción con las especificaciones para confeccionar la matriz de las competencias y construir el perfil del cargo o de ocupación. Ejemplo de dependiente gastronómico en la rama turística. *Revista de Investigación y Desarrollo Local*, 4(9). Recuperado de: <http://www.eumed.net/rev/turydes/09/pmza.html>
- Zhu, W., Ennis, C. D., & Chen, A. (1998). Many-faceted Rasch modelling expert judgment in test development. *Measurement in Physical Education and*

Exercise Science, 2(1), 21-39. Recuperado de: http://blogbook.co/many_faceted_rasch_modeling_expert_judgment_in_test.pdf
Zúñiga-González, C. G, & Cárdenas-Aguilera, P. (2014). Instrumentos de evaluación: ¿Qué piensan los estudiantes al terminar la escolaridad obligatoria?. *Perspectiva Educativa, Formación de Profesores*, 53(1), 57-72. Recuperado de: <http://www.redalyc.org/articulo.oa?id=333329700005>

1. Centro Universitario CIFE, Cuernavaca, Morelos, México. Doctor en Ciencias Biológicas y de la Salud. luisgibrancife@gmail.com

2. Centro Universitario CIFE, Cuernavaca, Morelos, México. Web: www.cife.edu.mx, Doctor en Modelos Educativos y Políticas Culturales. Globalización e Identidad en la Sociedad del Conocimiento. Correo: stobon5@gmail.com

Revista ESPACIOS. ISSN 0798 1015
Vol. 39 (Número 53) Año 2018.

[Índice]

[En caso de encontrar algún error en este website favor enviar email a webmaster]

©2018. revistaESPACIOS.com · Derechos Reservados