

LA VALIDACIÓN POR JUICIO DE EXPERTOS COMO ESTRATEGIA PARA MEDIR LA CONFIABILIDAD DE UN INSTRUMENTO

EXPERT JUDGMENT VALIDATION AS A STRATEGY FOR MEASURING THE RELIABILITY OF AN INSTRUMENT

Recibido: 25 de febrero 2022
Aceptado: 11 de marzo 2022

A. V., Balderas Sánchez¹
C., Cruz Navarro²
N., Zapata Garay³
J.M., Salazar Mata⁴

RESUMEN

La principal problemática de toda investigación es obtener información confiable; para lograrlo, se requiere contar con la validez de los instrumentos de recopilación de información, que permita garantizar la confiabilidad de los resultados, dejando que el investigador obtenga información útil que evalúe la hipótesis planteada toda investigación. Entendiendo como validez el grado en el que el instrumento permite medir lo que pretende medir; y la fiabilidad se refiere a que la información sea confiable, es decir, que sirva y aporte valor a la investigación. Estos dos aspectos ayudan a medir la efectividad del instrumento. El presente trabajo presenta los resultados de la validez del instrumento por juicio de expertos mediante la validación de contenido por el método de agregados individuales donde se consideraron las dos propiedades técnicas para indicar la calidad y utilidad del instrumento para medir los resultados de la investigación. Para la evaluación del instrumento se convocó a seis expertos, mismos que evaluaron diferentes aspectos como la presentación del cuestionario, la formulación y confiabilidad de los ítems y propuestas de mejora al instrumento; además, se evaluó los aspectos sociodemográficos y el nivel de adecuación de los ítems. Como resultado, se eliminó el aspecto "No. Control", considerando como "NO ADECUADO", en el apartado de datos sociodemográficos, y la reestructuración de un ítem de la primera dimensión del instrumento. Finalmente se obtuvo un índice de validez total (CVI) del 99%.

PALABRAS CLAVES:

Evaluación, Juicio de Expertos, Validez, Instrumento.

ABSTRAC

The main problem of any research is to obtain reliable information; to achieve this, the validity of the information collection instruments is carried out to guarantee the reliability of the results, allowing the researcher to obtain useful information that allows testing the hypothesis raised in his research. Understanding validity as the degree to which the instrument allows to measure what it intends to measure; and reliability refers to the fact that the information is reliable, that is, that it serves and adds value to the investigation. These two aspects help to measure the effectiveness of the instrument. This paper presents the results of the validity of the instrument by expert judgment through content validation by the method of individual aggregates where the two technical properties were considered to indicate the quality and usefulness of the instrument to measure the results of the investigation. For the evaluation of the instrument, six experts were summoned, who evaluated different aspects such as the presentation of the questionnaire, the formulation and reliability of the items and proposals for improvement of the instrument; In addition, the sociodemographic aspects and the level of adequacy of the items were evaluated. As a result, the "No. Control" aspect was eliminated, considering it "NOT ADEQUATE" in the sociodemographic data section, and the restructuring of an item from the first dimension of the instrument. Finally, a total validity index (CVI) of 99% was obtained.

¹ Profesor de Tiempo Completo. Tecnológico Nacional de México, Campus Ciudad Valles, alba.balderas@tecvalles.mx

² Profesor de Tiempo Completo. Tecnológico Nacional de México, Campus Ciudad Valles, claudia.cruz@tecvalles.mx

³ Profesor de Tiempo Completo. Tecnológico Nacional de México, Campus Ciudad Valles, nitgard.zapata@tecvalles.mx

⁴ Profesor de Tiempo Completo. Tecnológico Nacional de México, Campus Ciudad Valles, juan.salazar@tecvalles.mx

KEYWORDS:

Evaluation, Expert Judgment, Validity, Instrument.

INTRODUCCIÓN

La información confiable en un proceso de investigación es un elemento clave para garantizar la validez de los resultados obtenidos, así como, la confianza de que el instrumento utilizado esta evaluado y ajustado para cumplir con los estándares mínimos que nos permita aplicarlo de manera segura, logrando que la información obtenida sea útil para la toma de decisiones.

Para lograr lo anterior según Tristán (2008), se debe considera la confiabilidad y la validez del instrumento, para la primera es que el instrumento este construido adecuadamente y la segunda es que mida lo que tiene que medir. Rodero (2016) sugiere que para corroborar la validez de un instrumento se puede tomar en cuenta tres criterios principales: la validez de contenido, la validez de constructo y la validez de criterio.

Para la validez del contenido se determina el grado en que el instrumento creado podrá medir lo que tiene que medir a través de las tres dimensiones, los 30 ítems y los 4 niveles o ponderaciones, mismas que los expertos deberán evaluar para determinar la confianza de que los ítems son los más adecuados para evaluar el proyecto.

Para la validación por juicio de expertos, como mínimo se deben reunir de 5 a 7 expertos para considerar la evaluación como válida según Almanasreh (2019), aunque algunos autores sugieren un número mayor de expertos, no hay un acuerdo unánime para su determinación. De acuerdo con lo que describe Cabero y Llorente (2013), “la selección del número de expertos depende de aspectos como la facilidad para acceder a ellos o la posibilidad de conocer expertos suficientes sobre la temática objeto de la investigación”. Por otra parte, Escobar Pérez y Cuervo-Martínez (2008) sugieren que el número de expertos a considerar en un juicio depende del nivel de experiencia y el nivel de conocimiento.

Una vez elegidos y conformados se les envía individualmente el formato de evaluación, mismo que se concentra para determinar el índice de validez del instrumento, sin embargo, cada experto evalúa la relevancia y claridad de cada aspecto, además de emitir recomendaciones de claridad pertinencia y relevancia de cada ítem.

Además de lo anterior, cada experto recibe el propósito de la investigación, el objetivo del instrumento, la población y muestra, y todas aquellas variables que representen un valor para evaluar cada aspecto. Además, se envía un instrumento de validación que contiene los aspectos a considerar en la evaluación. Finalmente se concentra la información evaluada logrando identificar los ítems que requieren ser eliminados, ajustados o acordes al objeto de la investigación.

Como contexto para la evaluación, se estableció que el instrumento será aplicado a una muestra representativa de 74 estudiantes, muestra obtenida del grupo de estudiantes que estuvieron a cargo de los 6 docentes del proyecto, los cuales implementaron estrategias de incorporación de herramientas de software libre en la formación de sus estudiantes de la carrera de Ingeniería en Sistemas Computacionales, esto, servirá como una acción de evaluación de resultados del proyecto de investigación educativa “Impacto del uso del software libre en la formación de estudiantes de la carrera de ISC”.

METODOLOGÍA

Para el desarrollo del presente trabajo se siguió el procedimiento para “evaluar la validez de contenido un instrumento de recopilación de información”, bajo las siguientes consideraciones:

Objetivo

Diseñar y evaluar la validez de contenido, considerando los tres aspectos, claridad, pertinencia y relevancia de los ítems que lo componen, que permita medir la confiabilidad del “impacto del software libre en el logro de las competencias específicas y genéricas de los estudiantes de la carrera de Ingeniería en Sistemas Computacionales”.

Muestra

De acuerdo con las recomendaciones de Almanasreh (2019), para evaluar la validez de contenido se requiere de 5 a 7 participantes, razón por la que se seleccionó 6 expertos en el tema de interés, todos ellos docentes con estudios de posgrado ejerciendo la docencia e investigación en el área afín al tema de la investigación.

Instrumento

El instrumento a evaluar consta de una presentación, el objetivo de la investigación y los datos sociodemográficos del entrevistado; en su constructo consta de tres dimensiones, cada una de ellas considera 30 ítems o preguntas, con 4 posibles respuestas según escala de confiabilidad de Likert (Maldonado, 2012). Para facilitar su evaluación y aplicación, el instrumento fue elaborado en una herramienta digital facilitando la valoración por el panel de expertos, datos sociodemográficos y los cuestionamientos a realizar.

Adicional al instrumento se utilizará una matriz de evaluación que permite al experto cuantificar los aspectos que determinen la validez de contenido del instrumento considerando su propio juicio.

Proceso

Para la realización de la evaluación por juicio de expertos se inició con una selección de expertos cuya trayectoria ha sido reconocida en el campo de estudio afín a la investigación. Según lo describen Escobar-Pérez y Cuervo-Martínez (2008) para la elección de expertos se debe tener “una opinión informada de personas con trayectoria en el tema, que son reconocidas por otros como expertos cualificados en éste, y que pueden dar información, evidencia, juicios y valoraciones”. El promedio de expertos que deben ser considerados es de 10-15, aunque puede bastar con 5 a 7 para considerar como válido el documento.

El experto al emitir su juicio podrá sugerir la mejora del cuestionario, qué criterios debe considerar para que cumplan con el objetivo del estudio, además sugerirá cambio o eliminación de ítems, así como la validez y la pertinencia de las preguntas y tiempos de respuesta.

Para el estudio se seleccionó a un grupo de expertos que cumplen con amplia trayectoria y conocimiento en el área a evaluar, a continuación, se describen los perfiles de los expertos elegidos:

1. Experto 1. Área de conocimiento: Sistemas Computacionales, Nivel de estudios: Doctorado, Experiencia: Docente e Investigador
2. Experto 2. Área de conocimiento: Sistemas Computacionales, Nivel de estudios: Maestría, Experiencia: Egresado, Asesor
3. Experto 3. Área de conocimiento: Sistemas Computacionales, Nivel de estudios: Ingeniería, Experiencia: Docente, Gestor, Directivo
4. Experto 4. Área de conocimiento: Tecnologías de la Información, Nivel de estudios: Maestría, Experiencia: Docente e Investigador
5. Experto 5. Área de conocimiento: Sistemas Computacionales, Nivel de estudios: Maestría, Experiencia: Docente, Directivo e Investigador
6. Experto 6. Área de conocimiento: Sistemas Computacionales, Nivel de estudios: Doctorado, Experiencia: Docente, Directivo e Investigador

Una vez seleccionados, se envió vía correo electrónico a cada uno las indicaciones, explicando su propósito, alcance, objetivo y adjuntando el formato de evaluación y el formulario en línea.

Considerando como eje el método de Juicio de Expertos agregados individuales, se solicita a cada experto de manera particular, dé una estimación directa del instrumento. Posterior a ello se concentró la información obtenida en los siguientes aspectos evaluados:

- ❖ Para la VALORACIÓN GLOBAL se evaluó: Presentación del cuestionario, Formulación de los Ítems, Propuestas de mejora y la Validación global del instrumento. Se consideró una escala de medición de cuatro niveles: NO ADECUADO, POCO ADECUADO, ADECUADO, MUY ADECUADO.
- ❖ Para la VALORACIÓN SOCIODEMOGRÁFICA se evaluó: Aspectos personales y de identidad. La escala de medición propuesta: NO ADECUADO, POCO ADECUADO, ADECUADO, MUY ADECUADO.
- ❖ Finalmente, para la VALORACIÓN DEL NIVEL DE ADECUACIÓN DE LOS ÍTEMS se consideró: Las tres dimensiones que consta el instrumento, los 10 ítems en cada una de ellas. Se evaluó la CLARIDAD, PERTINENCIA Y RELEVANCIA de cada pregunta.

Una vez revisados y concentrados se analizó la coincidencia entre expertos y se evaluó si la redacción, estructura y contenido del instrumento es favorable, es decir que las preguntas son congruentes, claras en su redacción y no tendenciosas, además, se identificaron los aspectos en los cuales exista alguna diferencia de opinión entre los jueces.

RESULTADOS

El instrumento utilizado para evaluar por juicio de expertos consta de tres aspectos: La valoración global del instrumento, los datos sociodemográficos necesarios para la caracterización de la investigación y finalmente el nivel de adecuación de cada ítem o pregunta con respecto al objeto de estudio.

Partiendo de los resultados de la VALORACIÓN GLOBAL del instrumento, los expertos revisaron la forma de cómo se presenta el cuestionario, la claridad y la coherencia en la redacción de las preguntas, considerando la revisión, la formulación y la propuesta de mejora; finalizando con una valoración global del instrumento. En la siguiente tabla se describe la estructura de este aspecto (Tabla1):

Tabla 1. Aspecto 1. Valoración global
Fuente: Elaboración propia

VALORACIÓN GLOBAL. ASPECTOS A EVALUAR
1. PRESENTACIÓN DEL CUESTIONARIO
1.1. Valore si la presentación visual del cuestionario es adecuada
1.2 Valore si la introducción del cuestionario es adecuada
1.3 Valore las instrucciones para completar el cuestionario
1.4 Valore la adecuación del consentimiento informado que se incluye
2. FORMULACIÓN DE LOS ÍTEMS
2.1. ¿Considera que los ítems que se recogen son acordes con los constructos que se quieren medir?
2.2. ¿Considera que los ítems responden a las finalidades de la investigación?
3. PROPUESTA DE MEJORA
3.1. ¿Considera que hay algún aspecto que no está contemplado en el cuestionario? ¿Cuál? (datos sociodemográficos, diseño, constructos...)
3.2. ¿De acuerdo con su lectura el cuestionario, es comprensible?
3.3. ¿Considera que el cuestionario informa sobre el objeto de la investigación?
3.4. ¿Considera que hay algún aspecto que sobra o que genere redundancia?
4. VALORACIÓN GLOBAL DEL CUESTIONARIO
4.1 ¿Qué puntuación global daría al cuestionario?

Como resultado podemos decir que aun cuando el 100% de los jurados considera que la presentación y formulación del instrumento global como ADECUADO, dos de los seis expertos sugiere una propuesta de mejora en el instrumento, específicamente en lo que se refiere a la omisión de un dato sociodemográfico y la mejora en la redacción de dos preguntas.

Para la valoración SOCIODEMOGRÁFICA, se analizó la pertinencia y la importancia de los datos solicitados al encuestado, como se muestra en la siguiente tabla (Tabla2):

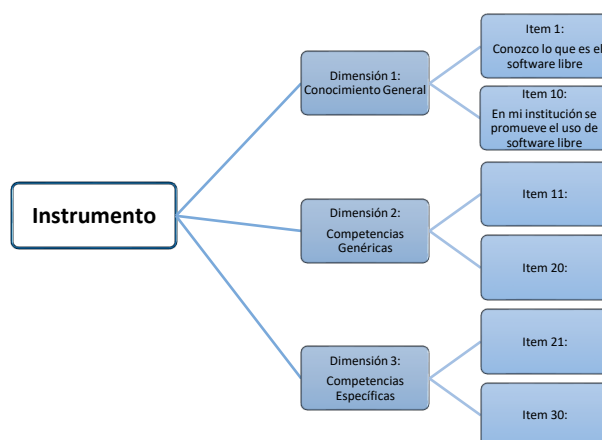
Tabla 2. Aspecto 2. Datos sociodemográficos
Fuente: elaboración propia

ATOS SOCIODEMOGRÁFICOS SOLICITADOS
1. NO. DE CONTROL
2. SEMESTRE
3. EDAD

ATOS SOCIODEMOGRÁFICOS SOLICITADOS
4. SEXO
5. ÁREA DE INTERÉS (BASES DE DATOS, REDES, PROGRAMACIÓN, GESTIÓN DE PROYECTOS, SEGURIDAD, INTERNET DE LAS COSAS, OTRA)
6. CORREO ELECTRÓNICO

Como resultado de los aspectos anteriores, el 100% de los expertos considero que el primer dato solicitado es irrelevante para la investigación, el resto de los datos se evaluaron como MUY ADECUADOS, concluyendo que deberá quedar excluido del instrumento el dato “No. de Control” por considerarse no necesario.

Finalmente, para la VALORACIÓN DEL NIVEL DE ADECUACIÓN DE LOS ÍTEMS se consideró una matriz para determinar la claridad, pertinencia y relevancia de cada pregunta. A continuación se muestra la estructura general de los aspectos que consta el instrumento, dimensiones, ítems o preguntas y niveles de valoración.



Gráfica 1. Estructura general del instrumento
Fuente: Elaboración propia

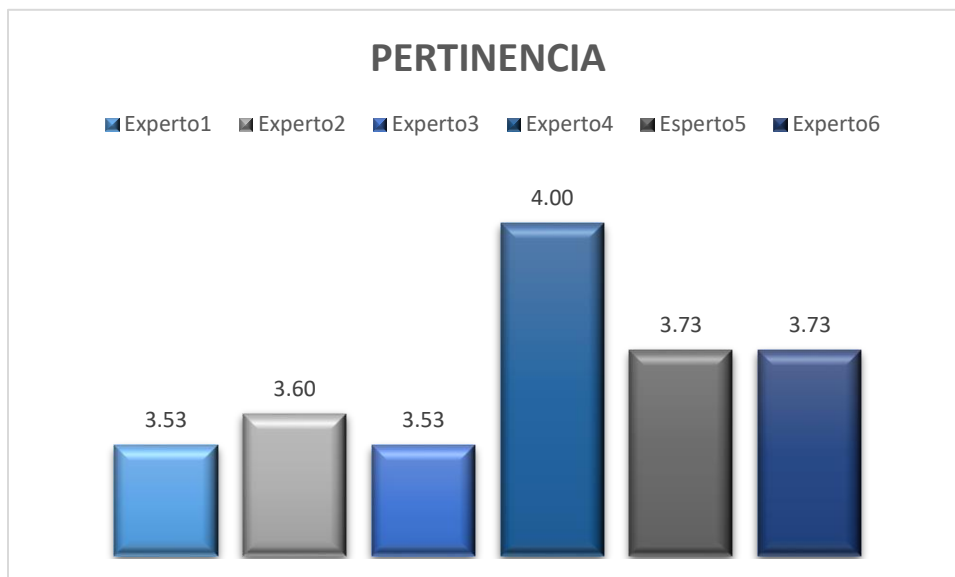
Para evaluar el nivel de adecuación de los ítems se consideran tres aspectos importantes: la claridad, la pertinencia y la relevancia de cada pregunta. La valoración considera una estala de Likert (Maldonado, 2012) de 4 niveles. A continuación, se presentan los resultados de la evaluación de la Claridad de las preguntas, considerando el rango de nada claro, no muy claro, claro y muy claro.



Gráfica 2. Aspecto 3. Nivel de Adecuación de los ítems: Claridad
Fuente: elaboración propia

Como se observa en la Gráfica 2, en lo referente a la claridad de los cuestionamientos del instrumento, sí consideramos el puntaje obtenido por experto para el total de los ítems, se observan que son considerados como claros y muy claros, con una calificación promedio de casi 4 puntos. Sin embargo, la revisión individual por experto arroja dos ítems que son evaluados como no muy claros, el ítem 6 y el ítem 9 de la primera dimensión, en resultado a ello se sugiere sean ajustados en su redacción para mejorar la claridad y la comprensión.

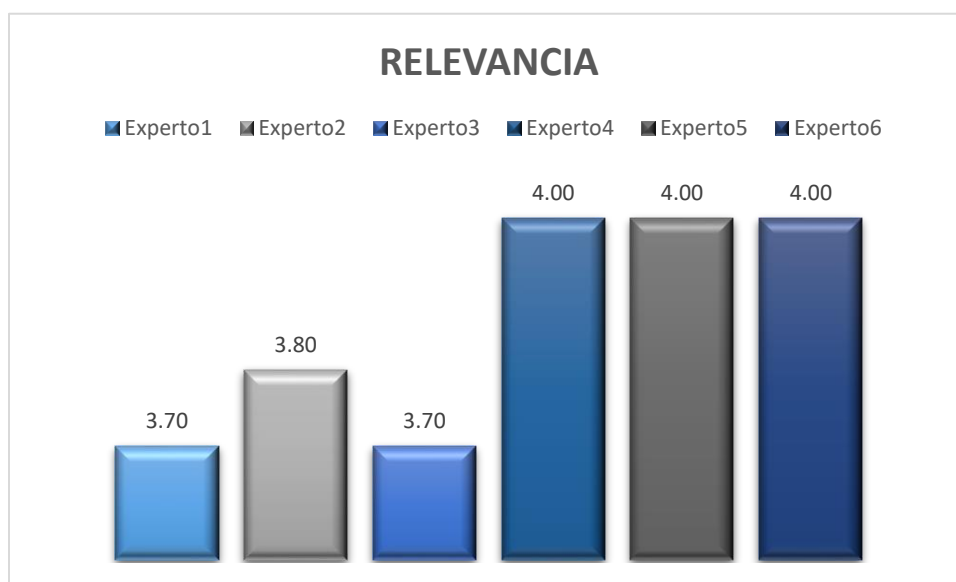
Para los resultados de la evaluación de la Pertinencia de las preguntas, se consideró el rango de nada pertinente, no pertinente, pertinente y muy pertinente.



Gráfica 3. Aspecto 3. Nivel de adecuación de los ítems: Pertinencia
Fuente: elaboración propia

Como se observa en la Gráfica 3, en lo referente a la pertinencia de los cuestionamientos del instrumento considerando el total del puntaje obtenido por experto para el conjunto de ítems, podemos observar que se destaca el experto4 con la puntuación más alta en lo que a pertinencia se refiere, ya que refiere que este tipo de acciones son muy necesarias para lograr posicionar a los estudiantes en un entorno laboral altamente competitivo. En generar los expertos determinan como pertinentes y muy pertinentes el total de los ítems. Igual que el análisis previo de calidad, en la revisión particular para el experto 5 y experto 6 consideran como no pertinente al ítem 26, se sugiere adecuar el cuestionamiento para denotar realmente la importancia o el sentido de la pregunta.

Finalmente, los resultados de la evaluación de la Relevancia se consideraron al igual que los aspectos anteriores un escalamiento de Likert (Maldonado, 2012) para facilitar su evaluación utilizando el rango de nada relevante, no relevante, relevante y muy relevante.



Gráfica 4. Aspecto 3. Nivel de adecuación de los ítems: Relevancia
Fuente: elaboración propia

Como resultado de la medición de los ítems en cuanto a su relevancia, se puede observar que el 100% de los expertos consideraron que los cuestionamientos son relevantes y muy relevantes como se observa en la puntuación obtenida en la gráfica anterior (Gráfico 4).

Como observaciones generales al instrumento los expertos opinaron que el tema a evaluar es muy relevante para la investigación y fortalece en gran medida las estrategias para lograr el perfil de egreso de los estudiantes de la carrera de Ingeniería en Sistemas Computacionales.

Con los resultados anteriores, una vez concentrados se obtuvo un CVI (Índice de Validez de Contenido) y el CVR (Razón de validez de contenido) para cada ítem de 1 en cuanto los tres aspectos evaluados, excepto el ítem 6 con un CVI de 0.66 y CVR de 0.3, y el ítem 9 con un CVI de 0.83 y un CVR de 0.7 de la primera dimensión para la CLARIDAD del cuestionamiento, así como el ítem 26 de la tercera dimensión con un CVI de 0.66 y un CVR

de 0.3 en cuanto a la PERTINENCIA de éste. Para el CVI parcial por aspecto se obtuvieron para CLARIDAD un 0.98, RELEVANCIA un 1 y PERTINENCIA un 1, finalmente se obtuvo un CVI total del 0.99 considerando que el instrumento es válido para su aplicación.

CONCLUSIONES

Siempre que construimos un instrumento para la recopilación de información debemos ser cuidadosos al considerar los factores que pueden afectar la confiabilidad y validez del mismo, como la improvisación, la falta de validación del instrumento, el no tener claro el contexto de aplicación, las condiciones y las estrategias para la aplicación.

Cuando se inicia con el proceso de construcción del instrumento, se debe diseñar un mayor número de ítems de los necesarios, con la finalidad de revisarlos e ir adecuando cada uno para el fin que está siendo construido, se debe considera recomendable que se redacten cuestionamientos en un 30% superior a los necesarios con la finalidad de mejorar la confiabilidad del instrumento.

Finalmente, el instrumento obtenido cumple el 99% de validez, sin embargo, requirió eliminar un aspecto en el apartado de datos sociodemográficos, al igual que mejorar la redacción para la claridad de 2 ítems y replantear la pertinencia de otro. Aun con lo anterior, la confiabilidad de la información a obtener estará sujeta a la correcta aplicación del instrumento.

BIBLIOGRAFÍA

- Almanasreh, E., et al. (2019). Research in Social and Administrative Pharmacy, 15(2), 214-221
- Cabero Almenara, Julio y Llórente Cejudo, María del Carmen. (2013). La aplicación del juicio de experto como técnica de evaluación de las tecnologías de la información (TIC). Eduweb. Revista de Tecnología de Información y Comunicación en Educación, vol. 7, núm. 2, pp.11-22. Recuperado de <http://servicio.bc.uc.edu.ve/educacion/eduweb/v7n2/art01.pdf>.
- Escobar Pérez, Jazmine y Cuervo Martínez, Ángela. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. Avances en Medición, vol. 6, núm. 1, pp. 27-36. Recuperado de http://www.humanas.unal.edu.co/psicometria/files/7113/8574/5708/Articulo3_Juicio_de_expertos_27-36.pdf. [Links]
- Maldonado Luna, S. M. (2012). Manual Práctico Para El Diseño De La Escala Likert (Maldonado 2012). Xihmai, 2(4). <https://doi.org/10.37646/xihmai.v2i4.101>
- Robles Garrote, Pilar y Rojas, Manuela del Carmen. (2015). La validación por juicio de expertos: dos investigaciones cualitativas en lingüística aplicada. Revista Nebrija de Lingüística Aplicada, núm. 18. Recuperado de https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_55002aca89c37.pdf. [Links]
- Rodero, C., Díaz, A., Guerrero, M., & Duque, M. (2016). Análisis de la fiabilidad y validez

de un cuestionario docente. *Revisión*, 9(1).

Tristán-López, A. (2008). Modificación al modelo de Lawshe para el dictamen cuantitativo de la validez de contenido de un instrumento objetivo. *Avances en medición*, 6(1), 37-48.