

Validación de un instrumento para evaluar competencias en investigación en médicos: estudio psicométrico con modelo de Rasch y grupo de expertos

Validation of an instrument to assess research competencies in physicians: a psychometric study using the Rasch model and expert focus group

Daniel Baltasar O'Besso Monsalve^{1,a}, Rodrigo Manuel Espinoza Rojas^{1,b},
Manuel Sebastian Costa Chavez^{1,c}, Luis Felipe Rojas Cama^{1,d}

¹ Facultad de Medicina, Universidad de Piura. Lima, Perú.

^a Estudiante de medicina

^b Estudiante de medicina

^c Estudiante de medicina

^d Magister en docencia e investigación en salud

An Fac med. 2025;86(1):47-54./ DOI: <https://doi.org/10.15381/anales.v86i1.29479>

Correspondencia:

Daniel Baltasar O'Besso Monsalve
daniel.obesso@alum.udelp.edu.pe

Recibido: 20 de noviembre 2024

Aprobado: 4 de marzo 2025

Publicación en línea: 28 de marzo 2025

Conflicto de interés: Las autores declaran no tener conflictos de interés

Fuente de financiamiento: Autofinanciado.

Contribuciones de autoría: DBOM: Conceptualización, metodología, análisis formal (estadístico, matemático, computacional), gestión de datos, redacción (primer borrador), redacción, revisión y edición, visualización y supervisión. RMER: Conceptualización, metodología, análisis formal (estadístico, matemático, computacional), gestión de datos, redacción (primer borrador), redacción, revisión y edición, visualización y supervisión. MSCC: Conceptualización, metodología, análisis formal (estadístico, matemático, computacional), gestión de datos, redacción (primer borrador), redacción, revisión y edición, visualización y supervisión. LFRC: Conceptualización, metodología, provisión de recursos, redacción (primer borrador), redacción, revisión y edición, visualización y supervisión.

El presente estudio forma parte de la tesis: Daniel Baltasar O'Besso Monsalve, Rodrigo Manuel Espinoza Rojas, Manuel Sebastian Costa Chavez. Validación de un instrumento para evaluar competencias en investigación en médicos: Estudio Psicométrico con modelo Rasch y Grupo de Expertos [tesis para optar al título de médico cirujano]. Lima: Facultad de Medicina Humana, Universidad de Piura; 2025.

Citar como: O'Besso D, Espinoza R, Costa M, Rojas L. Validación de un instrumento para evaluar competencias en investigación en médicos: estudio psicométrico con modelo de Rasch y grupo de expertos. An Fac med. 2025;86(1):47-54. DOI: <https://doi.org/10.15381/anales.v86i1.29479>



Resumen

Introducción. La validación de instrumentos de investigación es fundamental para asegurar la validez y confiabilidad de los resultados, lo que a su vez respalda la calidad de los estudios. **Objetivo.** Realizar un análisis psicométrico de la dimensión teórica de la «Prueba para evaluar las competencias en investigación» (PECI) y proponer una versión mejorada de esta dimensión. **Métodos.** Realizamos un análisis de bases secundarias con datos obtenidos de un estudio previo que aplicó este instrumento a médicos nombrados y contratados de Hospitales de Lima, Perú. Utilizamos la teoría de respuesta al ítem y el modelo de Rasch sobre la dimensión teórica del cuestionario. Además, tuvimos tres sesiones con expertos en investigación científica biomédica para recoger sugerencias y recomendaciones para mejorar los ítems de esta dimensión, las cuales fueron incorporadas a una versión final sugerida. **Resultados.** En la evaluación de los datos obtenidos en el estudio previo, la media del índice de dificultad fue 0,611 y del índice de discriminación 0,294. Todos los valores de infits y outfit estuvieron dentro del rango de 0,7 a 1,3. Ningún valor Q3 fue $\geq 0,5$. La PECI en su dimensión conocimiento teórico no cumplió el supuesto de unidimensionalidad. Recopilamos comentarios de cinco expertos y generamos una propuesta de mejora para la dimensión teórica. Nueve de los diez ítems requirieron cambios en su formulación y uno fue reformulado por completo. **Conclusión.** Este estudio presenta una propuesta mejorada de la dimensión teórica del instrumento PECI. Demostramos que instrumentos que fueron usados en estudios publicados son perfectibles en su validez y confiabilidad.

Palabras clave: Estudio de Validación; Psicometría; Encuestas y Cuestionarios; Grupos Focales (fuente: DeCS BIREME).

Abstract

Introduction. The validation of research instruments is essential to ensure the validity and reliability of the results, which in turn supports the quality of studies. **Objective.** To conduct a psychometric analysis of the theoretical dimension of the "Test to Assess Research Competencies" (PECI) and propose an improved version of this dimension. **Methods.** We performed a secondary data analysis using information from a previous study that applied this instrument to appointed and contracted physicians in hospitals in Lima, Peru. We applied item response theory and the Rasch model to the theoretical dimension of the questionnaire. Additionally, we held three sessions with experts in biomedical scientific research to gather suggestions and recommendations for improving the items in this dimension, which were incorporated into a proposed final version. **Results.** In the evaluation of the data obtained from the previous study, the average difficulty index was 0.611, and the discrimination index was 0.294. All infit and outfit values were within the range of 0.7 to 1.3. No Q3 value was ≥ 0.5 . The PECI, in its knowledge dimension, did not meet the unidimensionality assumption. We gathered feedback from five experts and developed a proposal to improve the theoretical dimension. Nine out of the ten items required changes in their wording, and one was completely reformulated. **Conclusion.** This study presents an improved version of the theoretical dimension of the PECI instrument. We demonstrated that instruments used in published studies could be further refined in terms of their validity and reliability.

Keywords: Validation Study, Psychometrics, Educational Measurement, Focus Groups (source: MeSH NLM).

INTRODUCCIÓN

La validación de los instrumentos de investigación es esencial para asegurar la confiabilidad y exactitud de los resultados en cualquier estudio^(1,2). Un instrumento validado respalda los hallazgos y reduce sesgos de medición que afecten la validez interna del estudio. Usualmente, los investigadores eligen instrumentos con validez de contenido, de constructo y de criterios, además, de medidas de confiabilidad aceptables; sin embargo, también resulta relevante realizar estudios psicométricos complementarios para aumentar la calidad de los instrumentos o tener una mejor aproximación sobre la validez de resultados de estudios previos^(3,4).

En 2022, Rojas y Contreras, publicaron una investigación que evalúa las competencias en investigación de médicos de hospitales de Lima⁽⁵⁾. Este instrumento fue tomado y adaptado de los estudios de González et al., Goleman y Boyatzis, quienes identificaron tres dimensiones de competencias en investigación: cognitiva, personal y social⁽⁶⁻⁸⁾. Los autores incorporaron la competencia teórica basada en los estudios de Rivas Tovar y Campos Céspedes cuyo contenido fue adaptado para formular 10 preguntas objetivas^(9,10). Este instrumento adaptado denominado PECEI "Prueba para evaluar las competencias en investigación" fue sometido a una prueba piloto en 32 médicos para evaluar la consistencia interna de los ítems en las competencias personal y social, sin embargo, no se realizó la validez del contenido de la competencia teórica.

Validar un instrumento que mida estas competencias es crucial para atender la creciente demanda de médicos competentes en la investigación científica. A su vez, evaluar estas competencias es fundamental, ya que su análisis puede mejorar la formación de los profesionales médicos y consecuentemente fortalecer su capacidad para generar producción científica de calidad⁽⁵⁾.

No contamos con antecedentes que hayan validado esta dimensión específica de la PECEI (la teórica). Por ello, consideramos un aporte, completar la validación del instrumento con la dimensión teórica. El objetivo de este estudio fue realizar

un análisis psicométrico de la dimensión de competencia teórica en investigación de un instrumento utilizado en un estudio publicado en una revista arbitrada⁽⁵⁾. Además, a partir de esos resultados recogimos sugerencias y comentarios de expertos en investigación para proponer una versión mejorada del instrumento.

MÉTODOS

Ámbito y diseño de estudio

Realizamos un estudio descriptivo, psicométrico con análisis de datos secundarios proveniente del estudio de Rojas y Contreras quienes aplicaron la PECEI a 100 médicos⁽⁵⁾. El proceso comenzó con la obtención del instrumento (paso 1), seguido por la obtención y limpieza de la base de datos utilizada en la publicación previa (paso 2). Luego, calculamos el índice de dificultad y el índice de discriminación (paso 3). Posteriormente, utilizamos el modelo Rasch (paso 4) y se organizó una reunión con expertos en investigación (paso 5). Durante esta reunión, se discutieron los resultados estadísticos y se recolectaron opiniones para mejorar la formulación de los ítems de la dimensión teórica (paso 6). Tras ello, se realizó el procesamiento de los resultados (paso 7), lo que permitió el diseño de una versión mejorada de la dimensión teórica (paso 8).

Descripción del instrumento PECEI

La PECEI fue adaptada por Rojas y Contreras en 2022, basada en informes de expertos y modelos de competencias en investigación de Goleman y Boyatzis^(7,8). Consta de 4 dimensiones: teórica, cognitiva, personal y social. Las dimensiones teórica y cognitiva se evaluaron binariamente (correcto = 1 / incorrecto = 0), con puntajes de 0 a 10 y 0 a 3, respectivamente. Las dimensiones personal y social se midieron en escala Likert (1-5), con rangos de 14 a 70 y 19 a 90, respectivamente. La puntuación total varió de 32 a 173 puntos.

Las competencias cognitiva, personal y social fueron basadas en estudios realizados por González et al., entre el 2008 y 2012^(6,11,12). La competencia teórica estuvo basada en los trabajos de Rivas Tovar y Campos Céspedes, quienes formularon 10 preguntas para evaluar la dimensión

mediante sub competencias^(9,10). La competencia teórica es una dimensión que mide el nivel de habilidad sobre investigación científica. En este sentido un investigador que presente todas las subcompetencias, tendrá un nivel de habilidad destacable en la dimensión teórica.

Las subcompetencias se evaluaron con los ítems del cuestionario: identificar un problema (ítem 1), elaborar una adecuada pregunta de investigación (ítem 2), revisar el estado del arte (ítem 3), usar tecnologías de información y comunicación (ítem 4), identificar diseños de estudio (ítem 7), conocer los principios de conducta responsable de investigación (ítem 5 y 8), conocer pruebas de validación de instrumentos (ítem 6), pruebas estadísticas (ítem 9) e identificar de estilos de escritura (ítem 10)^(9,10).

Recolección de datos

Solicitamos permiso para acceder a la base de datos del estudio primario que contenía un registros de 100 participantes y las variables: edad, sexo, estado civil, universidad de pregrado y posgrado, especialidad, lugar de trabajo, estatus de labor docente, tiempo en años de contrato y las claves marcadas para la dimensión teórica de la PECEI.

Preparamos la base en Microsoft Excel para su análisis. Todas las variables que no fueran de utilidad para el análisis, la discusión con expertos y los objetivos del estudio fueron eliminadas. La versión final incluía el código de los participantes y las respuestas marcadas para los 10 ítems de la competencia teórica. Finalmente, codificamos las respuestas con un valor de 0 (incorrecto) o 1 (correcto).

Recolectamos los comentarios de los expertos en investigación de manera presencial y virtual utilizando la aplicación *Zoom*. Para el desarrollo de la reunión tomamos en cuenta los resultados del análisis descriptivo, la teoría de respuesta al ítem (TRI) y el análisis mediante modelo de Rasch. Realizamos tres sesiones: en dos participaron dos expertos y en una, una experta.

Presentamos los ítems con los indicadores calculados y luego, ítem por ítem se les solicitó a los expertos su retroalimentación. Estos dieron recomendacio-

nes sobre el enunciado y las alternativas, así como sugerencias sobre el cuestionario en general. En las dos sesiones dónde hubo dos expertos hubo retroalimentación y comentarios entre ellos. En la última sesión con solo una experta, presentamos los comentarios proporcionados por los expertos en las reuniones anteriores para que pueda apoyar o refutar las opiniones.

Grabamos las sesiones y recopilamos comentarios en una base de datos anonimizada. La información proporcionada por los expertos sirvió para realizar cambios al cuestionario. Tomamos en cuenta todos los comentarios proporcionados y aplicamos los cambios. En caso de discrepancias entre los expertos, seleccionamos aquella opinión que estuvo apoyada por los resultados del análisis psicométrico.

Perfil de los expertos

Los expertos fueron docentes investigadores de una universidad privada en Lima, Perú, seleccionados por conveniencia. Cuatro eran médicos: dos docentes en metodología, estadística y redacción científica; otro en genética y ciencias básicas; y uno en gestión en salud. La quinta, ingeniera industrial, investigadora y docente en análisis y estadística de datos. Tres eran miembros de un comité de ética en investigación.

Análisis de datos

Incluimos a los participantes del estudio de Rojas y Contreras que respondieron a todas las preguntas del cuestionario. Calculamos la media de la sumatoria de puntos de los ítems. Aplicamos la Teoría de Respuesta al Ítem mediante la aplicación del modelo de Rasch. Utilizamos el programa *Jamovi*, específicamente, el módulo *snowIRT*. Utilizamos el modelo de Rasch pues permite analizar la relación entre un rasgo latente y la probabilidad de contestar de manera correcta un ítem de un cuestionario que mide ese rasgo. Los datos obtenidos mediante este análisis son evidencia sólida para saber si un ítem mide de forma correcta el constructo que dice medir ⁽¹³⁾. En el caso de la dimensión teórica del PECl, evaluamos si los 10 ítems, de forma independiente, miden los conocimientos en investigación científica.

Evaluamos los supuestos requeridos para utilizar el modelo de Rasch. Para evaluar el supuesto de unidimensionalidad realizamos el análisis de componentes principales (ACP) mediante la rotación *varimax* ⁽¹⁴⁾. Previamente, realizamos la prueba de esfericidad de Bartlett y la medida de idoneidad de la muestra para el análisis factorial con la prueba de Kayser-Meyer-Olkin (KMO-MSA).

La prueba de esfericidad de Bartlett verifica si la matriz de correlaciones se adecua al ACP. Se plantean dos hipótesis, nula (H_0) y alternativa (H_1). La H_0 afirma que la matriz de correlaciones es una matriz identidad dónde las correlaciones entre variables son cercanas a 0. La H_1 afirma que existen correlaciones entre las variables y, por lo tanto, es posible reducir la dimensionalidad de los datos. Rechazaremos la H_0 con un valor $p < 0,05$ ⁽¹⁵⁾.

La prueba KMO-MSA mide la proporción de varianza común entre las variables, indicando mayor idoneidad para el análisis a medida que el valor KMO es más alto. Calculamos el valor global para toda la dimensión teórica y consideramos un valor menor a 0,5 como indicador de que los datos no se ajustan al modelo ⁽¹⁴⁾.

Corroboramos el principio de independencia local, el cual propone que los ítems de una prueba no se deben correlacionar entre ellos. Para ello, utilizamos el coeficiente de correlación Q3 y el valor MADaQ3, que es la media de los valores Q3. Consideraremos que no existe correlación con un valor menor a 0,5 ⁽¹⁶⁾; este corte también fue usado para MADaQ3.

Calculamos los *infits* y *outfits* para cada ítem. Estos valores son residuales de la respuesta individual, los cuales se definen como la diferencia entre la información observada comparada con lo esperado por el modelo. Los valores cercanos al cero o mucho mayores a la unidad reflejan que la información no se ajusta con el modelo de Rasch ⁽¹⁷⁾. Consideramos como valores que indican ajuste al modelo los comprendidos entre 0,7 y 1,3 ⁽¹⁸⁾.

Utilizamos *Jamovi* versión 2.4.8 para graficar las curvas características del ítem (CCI). La CCI correlaciona un nivel determinado del rasgo con la probabilidad de responder correctamente un ítem ⁽¹³⁾.

Para cada uno de los niveles de dominio de la dimensión teórica se obtiene una probabilidad de contestar correctamente cada ítem ⁽¹⁹⁾.

Obtuvimos el índice de dificultad y el índice de discriminación utilizando *Microsoft Excel*. Para el cálculo del índice de dificultad tomamos como numerador el número de participantes que respondieron de manera correcta el ítem y el número total de respuestas como denominador. Respecto al índice de discriminación, dividimos a la muestra a la mitad. Para establecer estos dos grupos ordenamos los puntajes totales de los 10 ítems de los participantes. El grupo superior fue aquella mitad con los puntajes más altos y el grupo inferior, los que tuvieron peor desempeño. Finalmente, aplicamos la siguiente fórmula: (N° de participantes que respondieron correctamente al ítem en el grupo superior - N° de participantes que respondieron correctamente al ítem en el grupo inferior) / N° total de participantes.

Para la interpretación del índice de dificultad usamos los siguientes puntos de corte: $\leq 0,34$ (difícil), desde 0,35 hasta 0,59 (medianamente difíciles), desde 0,60 hasta 0,84 (medianamente fáciles) y $\geq 0,85$ (fáciles) ⁽²⁰⁾. Para la interpretación del índice de discriminación usamos los intervalos: $\geq 0,40$ (muy buen ítem), desde 0,39 hasta 0,30 (razonablemente bueno, pero con posibilidad de mejora), desde 0,29 hasta 0,20 (ítems regulares, usualmente necesitan mejora), $\leq 0,19$ (ítem de pobre calidad, necesita mejora o ser rechazado) ⁽²¹⁾.

Aspectos éticos

La investigación fue aprobada por el Comité Institucional de Ética en Investigación de la Universidad de Piura, expediente N° PREMEO6202325. La base de datos analizada estuvo anonimizada.

RESULTADOS

Excluimos a 4 médicos del análisis por tener al menos una respuesta sin marcar. El promedio de los puntajes en el cuestionario fue 6,11 y la mediana 6,63. 63 de 96 médicos obtuvieron un puntaje ≥ 6 . Cuatro obtuvieron un puntaje de 10, siete un puntaje de 9 y diez un puntaje

de 8. Tres obtuvieron el puntaje total más bajo que fue de 2.

Índice de dificultad y discriminación de los ítems

El promedio del índice de dificultad fue 0,611. Clasificamos un ítem como fácil (#1), cuatro como medianamente fáciles (#2, 4, 5, 7), cuatro como medianamente difíciles (#3, 8, 9, 10) y uno como difícil (#6). El promedio del índice de discriminación fue 0,294. Dos ítems fueron muy buenos (#8 y 10), cuatro razonablemente buenos (#3, 6, 7 y 9), uno regular (#5) y tres de pobre calidad (#1, 2, 4) (Tabla 1).

Análisis de teoría de respuesta al ítem

El valor global de la prueba KMO-MSA fue 0,480 y la prueba de esfericidad de Bartlett resultó en $p < 0,001$. El ACP mostró la existencia de dos dimensiones (Tabla 2). El promedio de los valores Q3 fue 0,106 (MADaQ3). Ningún valor Q3 tuvo un valor $\geq 0,5$ y algunas correlaciones fueron negativas (Tabla 3). Todos los infits y outfits dentro del rango de 0,7 y 1,3 (Tabla 1).

Curvas características de los ítems

Siete ítems (#1, 2, 3, 4, 5, 7, 9) tuvieron una probabilidad del 50% de ser contestados de forma correcta con un nivel de habilidad promedio en la dimensión teórica (Figura 1).

Reuniones con expertos y modificación del cuestionario

Realizamos satisfactoriamente las 3 reuniones con los expertos en investigación y recolectamos todos los comentarios y sugerencias. Realizamos una propuesta de mejora de cada ítem con sus respectivas alternativas según lo discutido en las 3 reuniones (Tabla 4). De manera general, los expertos comentaron que era mejor tener 4 alternativas antes que 5. En consecuencia, los expertos recomendaron eliminar una alternativa (distractor) de cada ítem. 9 de los 10 enunciados recibió algún tipo de modificación y, sólo un ítem, el #10, fue reformulado por completo.

DISCUSIÓN

Los resultados indican que la dimensión teórica no se ajustó adecuadamente al ACP, ya que no se cumple el criterio de unidimensionalidad. Sin embargo, se cumple el principio de independencia local pues no se identificaron valores Q3 $\geq 0,5$. Algunas correlaciones fueron negativas. El índice de dificultad promedio fue de 0,611, mientras que el índice de discriminación promedio fue de 0,294. Además, siete ítems presentaron una probabilidad del 50% de ser respondidos correctamente con un nivel de habilidad promedio. Nueve de los diez enunciados fueron mo-

dificados de alguna manera, y solo un ítem fue completamente reformulado.

Nuestros hallazgos muestran que un instrumento al someterse a un riguroso análisis puede ser mejorado. Por lo tanto, debe validarse correctamente luego de ser adaptado o creado. Las propiedades psicométricas de un instrumento deben ser evaluadas y publicadas, ya que esto garantiza su calidad y proporciona respaldo para su uso por otros autores⁽³⁾. Nuestra intención con este análisis fue estudiar las propiedades psicométricas de un instrumento utilizado en un estudio publicado, y que afrontó la revisión por pares. Solo con la interpretación de estos resultados estadísticos evidenciamos que el instrumento podría mejorarse, lo cual fue consistente con la revisión de los expertos.

Según los resultados, los datos respetan el principio de independencia local y los *infits* y *outfits*, y se ajustan al modelo de Rasch. Sin embargo, no se cumple el principio de unidimensionalidad y la prueba KMO-MSA tuvo un valor menor a 0,5, lo cual demuestra una muy pobre adecuación al ACP⁽¹⁴⁾. De este resultado interpretamos que no todos los ítems miden lo mismo y esto se puede observar en que algunos valores Q3 tuvieron correlaciones negativas entre ítems. A pesar de esto, continuamos con el análisis mediante la TRI y el modelo de Rasch. Consideramos que el motivo de que la dimensión teórica

Tabla 1. Índice de dificultad, índice de discriminación, *infits* y *outfits* por ítem.

N° ítem	Índice de dificultad	Interpretación	Índice de discriminación	Interpretación	<i>Infit</i> ^a	<i>Outfit</i> ^a
1	0,948	Fácil	0,104	Pobre calidad, necesita mejorar o ser eliminado.	0,958	0,667
2	0,729	Medianamente fácil	0,083	Pobre calidad, necesita mejorar o ser eliminado.	1,149	1,194
3	0,531	Medianamente difícil	0,313	Razonablemente bueno pero con posibilidad de mejora.	1,044	1,049
4	0,833	Medianamente fácil	0,167	Pobre calidad, necesita mejorar o ser eliminado.	1,019	1,034
5	0,677	Medianamente fácil	0,271	Regular calidad, usualmente necesitan mejora.	0,973	0,939
6	0,208	Difícil	0,375	Razonablemente bueno pero con posibilidad de mejora.	0,909	0,833
7	0,698	Medianamente fácil	0,313	Razonablemente bueno pero con posibilidad de mejora.	0,957	0,951
8	0,365	Medianamente difícil	0,479	Muy buen ítem.	0,988	1,021
9	0,583	Medianamente difícil	0,375	Razonablemente bueno pero con posibilidad de mejora.	1,027	1,056
10	0,542	Medianamente difícil	0,458	Muy buen ítem.	0,958	0,952

^a *Infit* y *Outfit*: valores entre 0,7 y 1,3 indican ajuste al modelo.

Tabla 2. Análisis de componentes principales con cargas por ítem.

Ítems	Dimensiones	
	1	2
1	-	0,432
2	-	-
3	0,429	-
4	-	0,418
5	0,854	-
6	-	0,761
7	0,800	-
8	-	0,554
9	0,415	0,393
10	-	0,583

no sea unidimensional radica en que Rojas y Contreras adaptaron 9 sub-competencias a una sola dimensión ^(9,10).

En nuestra opinión y la de los expertos, la competencia teórica está enfocada en la redacción de un protocolo de investigación dejando de lado otros aspectos como la validez interna, sesgos, recojo de datos, entre otros ⁽²²⁾. 5 de los 10 ítems tratan sobre redacción del protocolo de investigación (#1, 2, 4, 5, 10). El resto de los ítems tienen como temática la búsqueda bibliográfica (#3), la validación de un instrumento (#6), los diseños de estudio (#7), ética en investigación (#8) y análisis estadístico (#9), todos con un solo ítem agrupados en un solo factor. Estos hallazgos son consistentes con la naturaleza compleja y multidimensional que tienen los conocimientos en investigación científica.

Cinco ítems fueron medianamente fáciles o fáciles y cinco medianamente difíciles o difíciles. Ocho ítems tuvieron necesidad de mejora. Interpretamos de las CCI que 7 ítems tenían una probabilidad de respuesta correcta mayor al 50% con un nivel de habilidad promedio en conocimientos en investigación.

A pesar de que el instrumento fue usado en un estudio publicado y que fue revisado por pares, la metodología usada conllevó a la modificación de 9 de 10 ítems. Los expertos recomendaron eliminar uno de los distractores, fundamen-

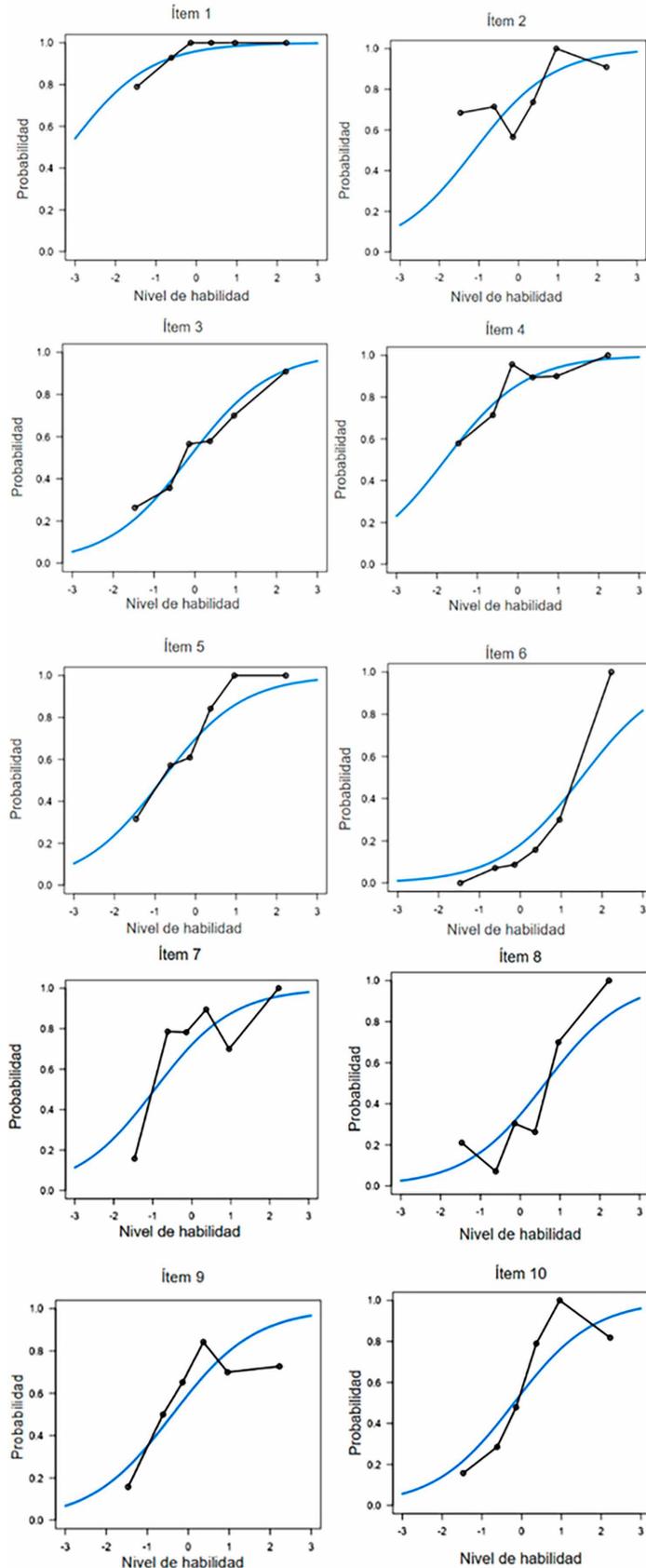


Figura 1. Curvas características de los ítems originales de la dimensión teórica.

Tabla 3. Matriz de correlación Q3 entre los ítems de la dimensión teórica.

Ítem	1	2	3	4	5	6	7	8	9	10
1	—									
2	-0,121	—								
3	0,048	-0,242	—							
4	-0,098	-0,150	-0,167	—						
5	0,009	0,041	-0,118	0,035	—					
6	0,026	0,028	-0,004	-0,042	0,008	—				
7	0,136	-0,160	-0,067	0,196	0,041	-0,361	—			
8	-0,026	-0,124	0,043	-0,283	-0,191	0,458	-0,169	—		
9	0,067	-0,288	-0,158	0,003	-0,041	-0,123	0,117	-0,099	—	
10	0,039	-0,215	-0,037	0,009	-0,086	0,084	0,026	-0,064	-0,077	—

Consideramos que valores por $\geq 0,5$ no respetan el supuesto de independencia local.

tando esta decisión en la presencia de distractores inadecuados o poco funcionales en todas las preguntas, los cuales parecían artificiales y comprometían la efectividad de la evaluación. Finalmente, aplicamos los cambios sugeridos por los expertos para generar una nueva versión del cuestionario.

La cantidad de modificaciones requeridas es la esperable dentro del proceso de validación de un instrumento. En otros estudios de diseño de instrumento, el juicio de expertos para la validación de contenidos suele generar modificaciones en el instrumento, lo cual no significa que esté mal diseñado si no, que es perfectible. El análisis mediante el modelo de Rasch brindó información a los expertos para contextualizar cada ítem y con dicha evidencia brindar sugerencias objetivas y basadas en evidencia.

Es imprescindible validar un instrumento antes de utilizarlo. Sin garantías de su validez y confiabilidad no se puede saber si mide realmente la variable que desea medir ⁽⁴⁾. Existen investigaciones que han evaluado psicométricamente instrumentos para medir conocimientos en investigación científica, sin embargo, estos no fueron realizados en médicos ^(23,24). No obstante, constituyen demostraciones de la pertinencia de estas validaciones antes de aplicar el instrumento en un nuevo estudio. Es fundamental destacar que el desarrollo y validación de un cuestionario

deben sustentarse en una base bibliográfica sólida y en un análisis riguroso ^(3,4).

Un estudio de 2023 diseñó un instrumento para medir competencias investigativas en docentes y estudiantes universitarios. Tras una revisión bibliográfica, elaboraron un cuestionario validado por expertos y fue ajustado tras una prueba piloto con 25 participantes, antes de aplicarlo a 304 personas. Evaluaron la confiabilidad con el alfa de Cronbach, el método de Hoyt y el de dos mitades de Rulon, y la validez con juicio de expertos y análisis factorial exploratorio (AFE). A diferencia de la PECE, que tiene una dimensión con nueve subdimensiones, este estudio presentó tres dimensiones con más ítems. Aunque la metodología y el contenido difieren del presente estudio, comparten el uso del análisis psicométrico utilizando las cargas por ítem del AFE en su caso y el ACP en el nuestro. Ambos métodos evidencia una estructura latente del instrumento ⁽²³⁾.

En 2023, otro estudio diseñó un instrumento para evaluar competencias en la elaboración de proyectos de investigación en estudiantes de enfermería. Siguiendo cinco etapas que fueron la validación por 15 expertos, la prueba piloto, el análisis factorial y técnicas como test-retest y el cálculo del alfa de Cronbach para evaluar la confiabilidad ⁽²⁴⁾. A diferencia del estudio de la PECE, este requirió múltiples etapas de validación, ya

que parte de las dimensiones de la PECE habían sido previamente estudiadas ⁽⁵⁾. No obstante, este artículo resalta la importancia de un análisis exhaustivo al desarrollar o adaptar un cuestionario, como lo evidencia la consulta con expertos, el análisis psicométrico y la extracción de dimensiones mediante ACP.

La calidad de los instrumentos utilizados en revistas científicas indizadas debe estar respaldada por un proceso riguroso de validación y ser descrita de manera detallada en la sección de métodos de los artículos originales. Además, el proceso de revisión por pares debe asegurar el cumplimiento del reporte de esta información, la cual es fundamental para valorar el riesgo de sesgo en la medición de los constructos estudiados ⁽²⁵⁾. Existen múltiples pasos, enfoques y estrategias que se pueden seguir para estudiar la confiabilidad y validez de un instrumento; por ejemplo, analizando la consistencia interna, realizando pruebas piloto, juicio de expertos y el análisis factorial exploratorio y confirmatorio ^(26,27). A su vez, se pueden implementar técnicas matemáticas y estadísticas variadas ^(28,29).

El estudio presentó limitaciones, como la imposibilidad de reunir a los cinco expertos simultáneamente y de corroborar el cumplimiento del supuesto de unidimensionalidad mediante una prueba piloto, para corroborar los ajustes realizados por los expertos. Como aspectos

Tabla 4. Ítem original, modificado y los cambios realizados junto a los comentarios de los expertos.

Ítem original	Ítem modificado	Cambios realizados y comentarios de expertos
<p>1. Se desea realizar una investigación, ¿Cuál es el primer paso para iniciar la elaboración de un protocolo de investigación?: (marque una)</p> <p><input type="checkbox"/> La elaboración del objetivo general.</p> <p><input type="checkbox"/> Tener un asesor estadístico.</p> <p><input type="checkbox"/> Contar con artículos científicos.</p> <p><input type="checkbox"/> Identificar una problemática a resolver.</p> <p><input type="checkbox"/> Tener una base de datos.</p>	<p>1. Se desea realizar una investigación, ¿Cuál es el primer paso para iniciar la redacción de un protocolo de investigación?: (marque una)</p> <p><input type="checkbox"/> La elaboración del objetivo general.</p> <p><input type="checkbox"/> Tener un asesor estadístico.</p> <p><input type="checkbox"/> Contar con artículos científicos.</p> <p><input type="checkbox"/> Identificar una problemática a resolver.</p>	<p>En el enunciado, cambiamos "elaboración" por "redacción". El cambio en el enunciado fue para mejorar la claridad ya que dos expertos lo consideraron confuso.</p> <p>Eliminamos la alternativa 5 por recomendación de los expertos. Además, ningún participante del estudio primario la marcó.</p> <p>Dos expertos comentaron que la pregunta era accesible para médicos sin formación en investigación y esto explica el índice de dificultad.</p>
<p>2. Identifique la característica de una adecuada pregunta de investigación: (marque una)</p> <p><input type="checkbox"/> Debe describir las variables demográficas.</p> <p><input type="checkbox"/> Describe las variables principales.</p> <p><input type="checkbox"/> Es similar al título más signos de interrogación.</p> <p><input type="checkbox"/> Es igual al objetivo general en condicional.</p> <p><input type="checkbox"/> Describe el año de la base de datos.</p>	<p>2. Identifique la característica de una buena estructura de una pregunta de investigación en salud: (marque una)</p> <p><input type="checkbox"/> Debe describir las variables demográficas.</p> <p><input type="checkbox"/> Describe las variables principales.</p> <p><input type="checkbox"/> Es similar al título más signos de interrogación.</p> <p><input type="checkbox"/> Es igual al objetivo general en condicional.</p>	<p>Los expertos comentaron que la pregunta generaba confusión. No estaba claro si preguntaba por la estructura (Modelo PICO) o por la calidad (Criterios FINER). Para ello, incluimos la frase "buena estructura de una pregunta de investigación en salud" en el enunciado.</p> <p>Eliminamos la alternativa 5. Los expertos la consideraron un mal distractor pues no tenía relación con la pregunta. Además, ningún participante del estudio primario la marcó.</p>
<p>3. Para realizar una búsqueda bibliográfica acertada, marque la mejor estrategia: (marque una)</p> <p><input type="checkbox"/> Buscar en google académico.</p> <p><input type="checkbox"/> Usar descriptores lógicos y lenguaje natural.</p> <p><input type="checkbox"/> Usar conectores lógicos y palabras clave.</p> <p><input type="checkbox"/> Realizar búsqueda avanzada y palabras usuales.</p> <p><input type="checkbox"/> Buscar tesis y artículos indexados.</p>	<p>3. Para realizar una búsqueda bibliográfica acertada, la mejor estrategia es utilizar: (marque una)</p> <p><input type="checkbox"/> Descriptores lógicos y lenguaje natural.</p> <p><input type="checkbox"/> Conectores lógicos y palabras clave.</p> <p><input type="checkbox"/> Búsqueda avanzada y palabras usuales.</p> <p><input type="checkbox"/> Tesis y artículos indexados.</p>	<p>Complementamos el enunciado con la frase "es utilizar".</p> <p>Eliminamos los conectores de las alternativas. Según los expertos esto evitaría que se utilice método de resolución de exámenes.</p> <p>Eliminamos la alternativa 1. Los expertos la consideraron un mal distractor. Además, fue marcada por solo 4 personas en el estudio primario.</p>
<p>4. Seleccione una herramienta tecnológica que permite organizar las fuentes bibliográficas: (marque una)</p> <p><input type="checkbox"/> Publisher.</p> <p><input type="checkbox"/> Zotero.</p> <p><input type="checkbox"/> SPSS.</p> <p><input type="checkbox"/> Word.</p> <p><input type="checkbox"/> Excel.</p>	<p>4. Seleccione la mejor herramienta tecnológica que permite gestionar las fuentes bibliográficas: (marque una)</p> <p><input type="checkbox"/> Publisher®.</p> <p><input type="checkbox"/> Zotero®</p> <p><input type="checkbox"/> APA.</p> <p><input type="checkbox"/> Excel®.</p>	<p>Reformulamos el enunciado a "Seleccione la mejor herramienta tecnológica que permite gestionar las fuentes bibliográficas". Los expertos mencionaron que todas las alternativas podrían ser correctas porque el enunciado pregunta por programas para organizar fuentes bibliográficas. Sugirieron el uso específico de la palabra "Gestionar".</p> <p>Los expertos sugirieron que debíamos incluir el símbolo de marca registrada en las alternativas.</p> <p>Cambiamos Word por APA como distractor. Sustentaron que el cambio de "Word" por "APA" plantea un mejor distractor.</p> <p>Eliminamos la alternativa 3.</p>
<p>5. Para evitar el plagio durante la elaboración de la base teórica debe realizar lo siguiente: (marque una)</p> <p><input type="checkbox"/> Insertar ideas de los autores directamente.</p> <p><input type="checkbox"/> Parafrasear y citar a sus autores.</p> <p><input type="checkbox"/> Citar y copiar el contenido de los artículos.</p> <p><input type="checkbox"/> Entrecomillar lo copiado.</p> <p><input type="checkbox"/> Añadir contenido de tesis.</p>	<p>5. Para evitar el plagio durante la redacción de la base teórica en una investigación médica se debe realizar lo siguiente: (marque una)</p> <p><input type="checkbox"/> Insertar ideas de los autores directamente.</p> <p><input type="checkbox"/> Parafrasear y citar a sus autores.</p> <p><input type="checkbox"/> Citar y copiar el contenido de los artículos.</p> <p><input type="checkbox"/> Entrecomillar lo copiado.</p>	<p>En el enunciado, cambiamos "elaboración" por "redacción" y agregamos la frase "investigación en medicina". Según los expertos, esto servirá para contextualizar a un paso específico del proceso investigativo y clarificar el tipo de investigación pues, no solo se investiga en medicina.</p> <p>Eliminamos la alternativa 5 porque la consideraron un mal distractor en comparación al resto.</p>
<p>6. Elija una prueba para validar un instrumento de recolección de datos: (marque una)</p> <p><input type="checkbox"/> Prueba de consistencia</p> <p><input type="checkbox"/> Prueba de rachas.</p> <p><input type="checkbox"/> Sensibilidad y especificidad.</p> <p><input type="checkbox"/> Juicio de expertos.</p> <p><input type="checkbox"/> Rho Spearman.</p>	<p>6. De los siguientes, ¿Cuál es el mejor método para contribuir en la validación de un instrumento de recolección de datos?: (marque una)</p> <p><input type="checkbox"/> Prueba de consistencia.</p> <p><input type="checkbox"/> Prueba de rachas.</p> <p><input type="checkbox"/> Juicio de expertos.</p> <p><input type="checkbox"/> Rho Spearman.</p>	<p>Reformulamos el enunciado a "De los siguientes, ¿Cuál es el mejor método para contribuir en la validación de un instrumento de recolección de datos?: (marque una)" Los expertos lo consideraron importante para evitar confusiones.</p> <p>Eliminamos la alternativa 3. Los expertos no llegaron a un consenso sobre qué alternativa eliminar. Decidimos eliminar la 3 porque fue la menos marcada en el estudio primario.</p>
<p>7. Es un diseño que presenta aleatorización, medición antes y después de la intervención y controles: (marque una)</p> <p><input type="checkbox"/> Transversal.</p> <p><input type="checkbox"/> Longitudinal.</p> <p><input type="checkbox"/> Ensayo Clínico aleatorizado.</p> <p><input type="checkbox"/> Casos y controles.</p> <p><input type="checkbox"/> Cohortes.</p>	<p>7. En investigación médica, es un diseño que presenta aleatorización, medición antes y después de la intervención y controles: (marque una)</p> <p><input type="checkbox"/> Longitudinal.</p> <p><input type="checkbox"/> Experimento.</p> <p><input type="checkbox"/> Casos y controles.</p> <p><input type="checkbox"/> Cohortes.</p>	<p>En el enunciado, agregamos la frase "En investigación médica". Este cambio tuvo el objetivo de contextualizar la pregunta.</p> <p>Eliminamos la alternativa 1. Los expertos la consideraron un mal distractor por ser el único tipo de investigación que no incluía temporalidad.</p> <p>Cambiamos la alternativa 3 a "Experimento" por sugerencia de los expertos. De este modo todas las alternativas tienen una extensión similar sin modificar el significado de la respuesta.</p>
<p>8. Dejar que el paciente decida acerca de su tratamiento o del ingreso a un estudio, cumple con la ética en el principio de: (marque una)</p> <p><input type="checkbox"/> Respeto.</p> <p><input type="checkbox"/> No maleficencia.</p> <p><input type="checkbox"/> Justicia.</p> <p><input type="checkbox"/> Autonomía.</p> <p><input type="checkbox"/> Beneficencia.</p>	<p>8. Dejar que el paciente decida acerca de su tratamiento o del ingreso a un estudio científico, ¿con qué principio de bioética se correlaciona?: (marque una)</p> <p><input type="checkbox"/> Respeto.</p> <p><input type="checkbox"/> Justicia.</p> <p><input type="checkbox"/> Autonomía.</p> <p><input type="checkbox"/> Beneficencia.</p>	<p>En el enunciado, especificamos "estudio científico" y agregamos la frase "¿Con qué principio de la bioética se correlaciona?".</p> <p>Los expertos comentaron que este cambio especificaba a qué aspecto de la bioética hacía referencia la pregunta.</p> <p>Eliminamos la alternativa dos.</p>
<p>9. Seleccione la prueba estadística que permite la comparación de medias: (marque una)</p> <p><input type="checkbox"/> Chi cuadrada.</p> <p><input type="checkbox"/> Rho spearman.</p> <p><input type="checkbox"/> Kruskal Wallis.</p> <p><input type="checkbox"/> OR y RR.</p> <p><input type="checkbox"/> T de Student.</p>	<p>9. Seleccione la prueba estadística que permite la comparación de medias: (marque una)</p> <p><input type="checkbox"/> Chi cuadrada.</p> <p><input type="checkbox"/> Rho spearman.</p> <p><input type="checkbox"/> Riesgo relativo.</p> <p><input type="checkbox"/> T de Student.</p>	<p>Eliminamos la alternativa 3.</p> <p>Sustituimos la alternativa 4 por solo una prueba "Riesgo relativo". Los expertos mencionaron que usar abreviaturas no era correcto. Además, es la única alternativa que presenta dos pruebas.</p>
<p>10. Seleccione una característica del estilo APA: (marque una)</p> <p><input type="checkbox"/> Citas numeradas.</p> <p><input type="checkbox"/> Números en orden.</p> <p><input type="checkbox"/> Lista de autores en orden alfabético.</p> <p><input type="checkbox"/> Similar a vancouver.</p> <p><input type="checkbox"/> Tablas y gráficos a colores.</p>	<p>10. Seleccione una característica del estilo Vancouver: (marque una)</p> <p><input type="checkbox"/> Citas con el apellido del autor y año.</p> <p><input type="checkbox"/> Citas numeradas en el texto.</p> <p><input type="checkbox"/> Lista de autores en orden alfabético.</p> <p><input type="checkbox"/> Citas con notas al pie de página.</p>	<p>Reformulamos totalmente el ítem 10 a una pregunta relacionada con el estilo Vancouver.</p>

Las alternativas resaltadas en negrita corresponden a las respuestas correctas.

positivos, las reuniones con grupos reducidos permitieron que los expertos tuvieran más tiempo para explayarse y realizar más sugerencias sobre los ítems. Además, identificamos áreas de mejora de la dimensión teórica mediante el modelo de Rasch. Finalmente, realizamos ajustes dirigidos a optimizar el instrumento, validando el contenido de este con los hallazgos derivados del análisis de teoría de respuesta al ítem.

En conclusión, al analizar la dimensión teórica de la PECE observamos que esta dimensión se adecua parcialmente al modelo de Rasch, al cumplirse 2 de los 3 supuestos. En base a los hallazgos de la teoría de respuesta al ítem aplicado al instrumento, el juicio de expertos recomendó modificaciones en todos los ítems. Demostramos que incluso instrumentos usados en estudios publicados son susceptibles de mejora, en función de la calidad de validación previo a su uso en una investigación. Nuestro estudio es relevante para el desarrollo de nuevos instrumentos que evalúan las competencias en investigación en los médicos. Esperamos que esta nueva versión del instrumento pueda usarse y aplicarse en investigaciones futuras.

AGRADECIMIENTOS

Al Dr. Franco Romaní de la Unidad de Investigación de la Facultad de Medicina Humana de la Universidad de Piura por su apoyo con el análisis estadístico y la revisión del manuscrito.

REFERENCIAS BIBLIOGRÁFICAS

- Heale R, Twycross A. Validity and reliability in quantitative studies. *Evid Based Nurs*. 2015;18(3):66-7. DOI: 10.1136/eb-2015-102129.
- Souza AC de, Alexandre NMC, Guirardello EB. Psychometric properties in instruments evaluation of reliability and validity. *Epidemiol Serv Saude*. 2017; 26(3):649-659. DOI: 10.5123/S1679-49742017000300022.
- López R, Avello R, Palmero D, Sánchez S, Quintana M. Validación de instrumentos como garantía de la credibilidad en las investigaciones científicas. *Rev Cub Med Mil [Internet]*. 2019 [citado el 19 de noviembre de 2024]; 48(Suppl 1). Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0138-65572019000500011&lng=es. Epub 01-Dic-2019.
- Argibay J. C. Técnicas Psicométricas. cuestiones de validez y confiabilidad. *Subjetividad y Procesos Cognitivos [Internet]*. 2006 [citado el 19 de noviembre de 2024]; (8):15-33. Disponible en: <https://www.redalyc.org/articulo.oa?id=339630247002>
- Rojas-Cama L, Contreras C. Competencias en investigación y producción científica en médicos de hospitales e institutos de salud de Lima, Perú. *An Fac med*. 2022;83(2):95-103. DOI: <https://doi.org/10.15381/anales.v83i2.22240>
- González C, Tomimbeni S, Corigliani S, Gentes G, Ginocchio A, Morales MM. Evaluación de competencias requeridas para investigar [Internet]. En *Anuario de Investigaciones de la Facultad de Psicología*; 2012 [citado 2024, noviembre 27]. Disponible en: <https://revistas.unc.edu.ar/index.php/aifp/article/view/2904>
- Goleman D. *Inteligencia Emocional*. En: Penguin Random House Grupo USA; 2003. p. 582.
- Boyatzis R. *The Competent Manager: A Model for Effective Performance*. En: John Wiley & Sons; 1982. p. 342.
- Rivas Tovar L. Las nueve competencias de un investigador. *Investig Adm*. 2011;(108):34-54. Disponible en: <https://www.redalyc.org/articulo.oa?id=456045339003>
- Campos J, Madriz L, Brenes O, Rivera Y, Viales M. Competencias investigativas en el personal académico de la Escuela de Ciencias de la Educación de la Universidad Estatal a Distancia, Costa Rica. *Cuadernos de Investigación UNED Res J*. 2012;4(2):273-82. DOI: <https://doi.org/10.22458/urj.v4i2.16>
- Corigliani S, Tomimbeni S, Ginocchio A. La Evaluación De Las Competencias Cognitivas Para Investigar: Situaciones De Prueba [Internet]. 2009 [citado 27 de septiembre de 2024]; Disponible en: https://www.academia.edu/26982065/La_Evaluaci%C3%B3n_De_Las_Competencias_Cognitivas_Para_Investigar_Situaciones_De_Prueba.
- Tomimbeni S, González C, Corigliani S, Salvetti M. Concepciones de expertos sobre las competencias para investigar en psicología. *Enseñ. invest. psicol*. 2011;16(1):5-13. Disponible en: <https://www.redalyc.org/pdf/292/29215963001.pdf>
- Agustín T. Análisis de Rasch para todos. En: Agustín T (Editor). *Análisis de Rasch para todos. Una guía simplificada para evaluadores educativos*. México: Instituto de Evaluación e Ingeniería Avanzada. 2013. Disponible en: https://www.researchgate.net/profile/Wendel_Dos_Santos3/post/How_to_compare_frequency_of_yes_responses_on_a_multi-item_survey/attachment/59d6590c79197b80779aeb4e/AS3A540895440064512%401505970930169/download/LibroRaschXTodos.pdf
- Lloret-Segura S, Ferreres-Traver A, Hernández-Baeza A, Tomás-Marco I. El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *An. Psicol*. 2014;30(3):1151-69. DOI: 10.6018/analesps.30.3.199361.
- López-Aguado M, Gutiérrez-Provecho L. Cómo realizar e interpretar un análisis factorial exploratorio utilizando SPSS. *REIRE Rev Innovació Recer En Educ*. 2019;12(2):1-14. DOI: <https://doi.org/10.1344/reire2019.12.227057>.
- Tavakol M, Dennick R. Psychometric evaluation of a knowledge based examination using Rasch analysis: an illustrative guide: AMEE guide no. 72. *Medical Teacher*. 2013;35(1):e838-48. DOI: 10.3109/0142159X.2012.737488.
- Christensen K, Kreiner S, Mesbah M. *Rasch Models in Health*. En: Christensen K, Kreiner S, Mesbah M (Editores). *Rasch Models in Health*. Estados Unidos: John Wiley & Sons. 2013:386. DOI:10.1002/9781118574454.
- Bond T, Yan Z, Heene M. *Applying the Rasch Model*. En: Bond T, Yan Z, Heene M (Editores). *Applying the Rasch Model. Fundamental Measurement in the Human Sciences*. 4. Estados Unidos: Routledge. 2020:455. DOI: <https://doi.org/10.4324/9780429030499>.
- Matas-Terrón A. Introducción al análisis de la Teoría de Respuesta al Ítem. En: Matas-Terrón A, Aidesoc (Editores). *Introducción al análisis de la Teoría de Respuesta al Ítem*. Aidesoc [Internet]. 2010. [citado 27 de septiembre de 2024] Disponible en: https://riuma.uma.es/xmlui/bitstream/handle/10630/4711/TRI_aidesoc_2011.pdf?sequence=1&isAllowed=y
- Doran R. Basic measurement and evaluation of science instruction. En: Doran R (Editor). *Basic measurement and evaluation of science instruction*. Estados Unidos: National Science Teachers Association. 1980:131.
- Ebel R, Frisbie D. *Educational measurements*. En: Ebel R, Frisbie D (Editores). *Educational measurements*. Estados Unidos: Prentice-Hall. 1991:232.
- Indrayan A. Elements of medical research. *Indian J Med Res*. 2004;119(3):93-100.
- Rios P, Ruiz C, Paulos T, León R. Desarrollo de una escala para medir competencias investigativas en docentes y estudiantes universitarios. *Areté, Revista Digital del Doctorado en Educación*. 2023;9(17):147-169. DOI: <https://doi.org/10.55560/arete.2023.17.9.7>
- Landeros-Olvera E, Garita-Alonso M, Rangel-González R, Ramirez-Girón N. Diseño de un instrumento de valoración de competencias para construir proyectos de investigación con estudiantes de enfermería. *Index Enferm [Internet]*. 2023 [citado 18 de octubre de 2024];32(1). Disponible en: https://scielo.isciii.es/scielo.php?script=sci_abstract&pid=S1132-12962023000100013&lng=es&nrm=iso&tlng=es
- Carretero-Dios, H, Pérez C. Normas para el desarrollo y revisión de estudios instrumentales. *Int. J. Clin. Health Psychol [Internet]*. 2005;5(3):521-551. [citado el 6 de noviembre de 2024]. Disponible en: <https://www.redalyc.org/articulo.oa?id=33705307>
- Arribas, M. Diseño y validación de cuestionarios. *Matronas prof. [Internet]*. 2004 [citado el 6 de noviembre de 2024];5(17):23-29. Disponible en: https://www.enferpro.com/documentos/validacion_cuestionarios.pdf.
- Muñoz J, Fonseca-Pedrero E. Diez pasos para la construcción de un test. *Psicothema*. 2019;31:7-16. DOI:10.7334/psicothema2018.291.
- Pedrosa I, Suárez-Álvarez J, García-Cueto E. Evidencias sobre la Validez de Contenido: Avances Teóricos y Métodos para su Estimación. *Acción psicol*. 2014;10(2):3. DOI: <https://dx.doi.org/10.5944/ap.10.2.11820>.
- Soriano A. Diseño y validación de instrumentos de medición. *Diálogos*. 2015;13:19-40. DOI: <https://doi.org/10.5377/dialogos.v0i14.2202>.