REVISTA PSICOLÓGICA HEREDIANA
ISSN 0034-8574
2025, enero-junio | Vol. 18, núm. 1, e8509

ARTÍCULO ORIGINAL
ORIGINAL ARTICLE / ARTIGO ORIGINAL

DOI: https://doi.org/10.20453/rph.v18i1.8509

Validez y confiabilidad de un cuestionario para evaluar el progreso académico, las habilidades técnicas y las actitudes profesionales en estudiantes de Enfermería a mitad de carrera: estudio piloto

Validity and reliability of a questionnaire to assess academic achievement, technical skills and professional attitudes in nursing students halfway through their degree: a pilot study

Validade e confiabilidade de um questionário para avaliar o progresso acadêmico, as habilidades técnicas e as atitudes profissionais de estudantes de Enfermagem na metade do curso: estudo piloto

Ana Elvira López y Rojas[1],
Ana Lucy Siccha-Maccassi[1],
Zenaida Rojas-Apaza[2],
Daniel José Blanco-Victorio[2],
Mayte Luzmila Casaño Meza[2],
Elizabeth Dany Araujo-Robles[2],
Geraldine Salazar-Vargas[2]

[1] Universidad Nacional del Callao. Callao, Perú.

[2] Universidad Peruana Cayetano Heredia. Lima, Perú.

Recibido: 05-04-2025
Aceptado: 04-06-2025
En línea: 30-06-2025

Artículo de acceso abierto
© Los autores

RESUMEN

El objetivo de este estudio fue evaluar la validez y la confiabilidad de un cuestionario diseñado para medir el progreso académico, las habilidades técnicas y las actitudes profesionales en estudiantes de Enfermería a mitad de carrera en una universidad peruana. Participaron 51 estudiantes seleccionadas mediante muestreo por conveniencia. Se aplicó un cuestionario autoadministrado de 30 ítems, organizado en seis dimensiones: Introducción a la Enfermería, Bioquímica, Tecnología del Cuidado, Cuidado Integral de la Familia, Epidemiología y Enfermería en Atención del Adulto I y II. Las respuestas fueron dicotomizadas como incorrectas (0) o correctas (1). La confiabilidad global fue aceptable (KR-20 = 0,7627) y todas las dimensiones superaron el valor de 0,72. Epidemiología presentó la mayor confiabilidad (KR-20 = 0,7713), mientras que Tecnología del Cuidado obtuvo el valor más bajo (KR-20 = 0,7235). La dificultad promedio fue moderada-alta (0,7075), con Bioquímica como la dimensión de mayor exigencia y Cuidado Integral de la Familia como la de menor dificultad. El análisis factorial confirmatorio y los índices globales de ajuste evidenciaron limitaciones en la estructura interna; sin embargo, el análisis por ítems mostró estadísticos Infit y Outfit dentro de rangos productivos para la medición. Además, se observó una adecuada correspondencia entre la habilidad de las participantes y la dificultad de los ítems. La validez convergente y divergente fue favorable en la mayoría de las dimensiones. En conclusión, el cuestionario presenta evidencia inicial de validez y confiabilidad para evaluar competencias académicas y profesionales en estudiantes de Enfermería a mitad de carrera. Se recomienda revisar los ítems con menor escalabilidad y ampliar la validación en muestras más grandes y diversos contextos institucionales.

Palabras clave: estudiantes de enfermería; validez; confiabilidad; medición educativa; competencias profesionales.

Citar como:
López y Rojas, A. E., Siccha-Maccasi A. L., Rojas-Apaza, Z., Blanco-Victorio, D. J., Casaño Meza, M. L., Araujo-Robles, E. D. y Salazar-Vargas, G. (2025). Validez y confiabilidad de un cuestionario para evaluar el progreso académico, las habilidades técnicas y las actitudes profesionales en estudiantes de Enfermería a mitad de carrera: estudio piloto. *Revista Psicológica Herediana, 18*(1), e8509. https://doi.org/10.20453/rph.v18i1.8509

ABSTRACT

This study aimed to evaluate the validity and reliability of a questionnaire designed to assess academic achievement, technical skills and professional attitudes among nursing students halfway through their degree at a Peruvian university. A total of 51 students were selected through convenience sampling. A 30-item self-administered questionnaire was applied and organized into six dimensions: Introduction to Nursing, Biochemistry, Care Technology, Comprehensive Family Care, Epidemiology, and Adult Nursing Care I and II. Responses were dichotomized as incorrect (0) or correct (1). Overall reliability was acceptable (KR-20 = 0.7627), and all dimensions exceeded 0.72. Epidemiology showed the highest reliability (KR-20 = 0.7713), whereas Care Technology showed the lowest value (KR-20 = 0.7235). The average item difficulty was moderately high (0.7075), with Biochemistry being the most demanding dimension and Comprehensive Family Care the least difficult. Confirmatory factor analysis and global fit indices revealed limitations in the internal structure; however, item-level analysis showed Infit and Outfit statistics within productive measurement ranges. An adequate correspondence was also observed between participant ability and item difficulty. Convergent and divergent validity were favorable in most dimensions. In conclusion, the questionnaire provides initial evidence of validity and reliability for assessing academic and professional competencies in nursing students halfway through their degree. Future studies should review items with lower scalability and validate the instrument in larger samples and diverse institutional contexts.

Keywords: nursing students; validity; reliability; educational measurement; professional competence.

RESUMO

O objetivo deste estudo foi avaliar a validade e a confiabilidade de um questionário elaborado para medir o progresso acadêmico, as habilidades técnicas e as atitudes profissionais de estudantes de Enfermagem na metade do curso em uma universidade peruana. Participaram 51 estudantes selecionadas por meio de amostragem por conveniência. Foi aplicado um questionário autoaplicável com 30 itens, organizado em seis dimensões: Introdução à Enfermagem, Bioquímica, Tecnologia do Cuidado, Cuidado Integral da Família, Epidemiologia e Enfermagem no Atendimento ao Adulto I e II. As respostas foram dicotomizadas como incorretas (0) ou corretas (1). A confiabilidade global foi aceitável (KR-20 = 0,7627) e todas as dimensões superaram o valor de 0,72. A Epidemiologia apresentou a maior confiabilidade (KR-20 = 0,7713), enquanto a Tecnologia do Cuidado obteve o valor mais baixo (KR-20 = 0,7235). A dificuldade média foi moderada-alta (0,7075), sendo a Bioquímica a dimensão de maior exigência e o Cuidado Integral da Família a de menor dificuldade. A análise fatorial confirmatória e os índices globais de ajuste evidenciaram limitações na estrutura interna; no entanto, a análise por itens mostrou os índices Infit e Outfit dentro de faixas aceitáveis para a medição. Além disso, observou-se uma correspondência adequada entre a habilidade das participantes e a dificuldade dos itens. A validade convergente e divergente foi favorável na maioria das dimensões. Em conclusão, o questionário apresenta evidências iniciais de validade e confiabilidade para avaliar competências acadêmicas e profissionais em estudantes de Enfermagem na metade do curso. Recomenda-se revisar os itens com menor escalabilidade e ampliar a validação em amostras maiores e em diversos contextos institucionais.

Palavras-chave: estudantes de enfermagem; validade; confiabilidade; medição educacional; competências profissionais.

INTRODUCCIÓN

La educación en Enfermería tiene como finalidad formar estudiantes capaces de alcanzar las competencias profesionales necesarias antes de incorporarse a la práctica clínica. Una formación de calidad constituye la base para preparar profesionales de la salud competentes, capaces de brindar una atención segura, ética y centrada en las necesidades de la población. En este proceso, tanto la formación teórica como la práctica son componentes esenciales del currículo, ya que permiten integrar conocimientos, habilidades técnicas y actitudes profesionales en escenarios reales de atención (Nilsson et al., 2019; Purabdollah et al., 2023; Suliman et al., 2023).

La Organización Mundial de la Salud define las competencias como el conjunto de conocimientos, habilidades y comportamientos necesarios para desempeñar adecuadamente roles y responsabilidades profesionales (World Health Organization, 2015). En Enfermería, estas competencias constituyen indicadores relevantes del desempeño y permiten valorar la capacidad para ejecutar tareas propias del cuidado profesional (American Association of Colleges of Nursing, 2026; Lin et al., 2017). Por ello, los planes de estudio deben orientarse a la formación de profesionales capaces de integrar la teoría y la práctica, desarrollar autonomía en la valoración integral de las personas y asumir con responsabilidad su rol dentro del equipo de salud (Lee y Sim, 2019; Löfgren et al., 2023).

En los últimos años, la educación basada en competencias ha adquirido especial relevancia en Enfermería. Este enfoque prioriza los resultados de aprendizaje y las competencias que el estudiante debe demostrar, más que el tiempo de permanencia en un programa formativo (Galura y Warshawsky, 2022; Giddens y Mansfield, 2023; Jenkins-Weintaub et al., 2023). En consecuencia, la evaluación del progreso académico y profesional requiere instrumentos válidos, confiables y pertinentes al contexto educativo.

La evaluación de estudiantes de Enfermería durante su formación representa un desafío para las instituciones, especialmente cuando se busca establecer niveles esperados de competencia y reducir la variabilidad en los métodos de evaluación. La literatura destaca la necesidad de contar con herramientas estandarizadas, con evidencias de validez y confiabilidad, así como de fortalecer la capacitación de docentes y supervisores clínicos para garantizar evaluaciones objetivas y justas (Altmiller, 2019; Finstad et al., 2022; Hughes et al., 2019).

En este contexto, disponer de un instrumento que permita medir el progreso académico, las habilidades técnicas y las actitudes profesionales resulta fundamental para identificar fortalezas y áreas de mejora en la formación de estudiantes de Enfermería. Por ello, el objetivo del presente estudio fue evaluar la validez y la confiabilidad de un cuestionario diseñado para medir dichas dimensiones en estudiantes de Enfermería a mitad de carrera.

MATERIALES Y MÉTODOS

Participantes

Se realizó un estudio transversal en estudiantes que se encontraban a mitad de la carrera de Enfermería. Se incluyeron 51 estudiantes de la Escuela de Enfermería de la Universidad Nacional del Callao durante el año 2023, seleccionadas mediante muestreo por conveniencia. Los criterios de inclusión fueron: asistencia regular, matrícula en cursos correspondientes a la mitad de la carrera y firma del consentimiento informado.

Instrumento

Se aplicó un cuestionario autoadministrado elaborado con la participación de docentes responsables de curso. El instrumento tuvo como propósito evaluar el progreso académico, las habilidades técnicas y las actitudes profesionales de estudiantes de Enfermería a mitad de carrera. Estuvo conformado por 30 preguntas distribuidas en seis dimensiones: i) Introducción a la Enfermería (P1-P5); ii) Bioquímica (P6-P9); iii) Tecnología del Cuidado (P10-P13); iv) Cuidado Integral de la Familia (P14-P18); v) Epidemiología (P19-P23); y vi) Enfermería en Atención del Adulto I y II (P24-P30). Cada pregunta presentó cinco alternativas de respuesta (A, B, C, D y E). Para el análisis psicométrico, las respuestas se dicotomizaron en incorrectas (0) y correctas (1).

Análisis de datos

La consistencia interna del cuestionario y de sus dimensiones se evaluó mediante el coeficiente de Kuder-Richardson 20 (KR-20), considerando aceptables los valores iguales o superiores a 0,70. Asimismo, se aplicó la teoría de respuesta al ítem (TRI), particularmente el modelo logístico de dos parámetros (2PL), con el fin de examinar la dificultad y la discriminación de los reactivos. También se evaluaron los estadísticos de ajuste individual Infit y Outfit, considerando como rango productivo para la medición los valores entre 0,5 y 1,5, y como rango aceptable estricto los valores entre 0,7 y 1,3.

La validez estructural interna se examinó mediante análisis factorial confirmatorio (AFC). Para evaluar el ajuste del modelo se consideró el estadístico M2, la prueba de chi-cuadrado, la razón entre chi-cuadrado y grados de libertad (χ^2/gl), el índice de ajuste comparativo (CFI), el índice de Tucker-Lewis (TLI), el error cuadrático medio de aproximación (RMSEA) y el residuo cuadrático medio estandarizado (SRMR). Se asumió un ajuste adecuado cuando CFI y TLI fueron ≥ 0,95, RMSEA y SRMR presentaron valores bajos y la razón χ^2/gl se mantuvo dentro de rangos aceptables. La validez convergente se consideró favorable cuando los ítems correlacionaron de manera suficiente con su propia dimensión, mientras que la validez divergente se valoró al observar co-

rrelaciones más altas con la dimensión correspondiente que con dimensiones diferentes. Los análisis se realizaron en STATA v.18 y RStudio, utilizando los paquetes ggmirt y cowplot.

Aspectos éticos

El estudio fue considerado exento por el Comité de Investigación Institucional de la Universidad Nacional del Callao. La participación fue voluntaria y se obtuvo consentimiento informado. Los datos fueron anónimos y almacenados mediante códigos encriptados, garantizando la confidencialidad y el cumplimiento de los principios éticos de la Declaración de Helsinki.

RESULTADOS

Se invitó a participar a todas las estudiantes que culminaron el quinto ciclo de Enfermería (n = 51). La totalidad de participantes fue de sexo femenino y sus edades oscilaron entre 19 y 23 años. El análisis de escalabilidad mostró valores bajos de Hj_min (<0,30) en algunas dimensiones. Se observaron valores reducidos en Introducción a la Enfermería (Hj_min = 0,27), asociados al ítem P3; Tecnología del Cuidado (Hj_min = 0,28), asociado al ítem P11; Epidemiología (Hj_min = 0,27), asociado al ítem P22; y Enfermería en Atención del Adulto I y II (Hj_min = 0,27). Estos hallazgos sugieren la necesidad de revisar la formulación y alineación conceptual de algunos reactivos. Los valores de confiabilidad, dificultad y correlación ítem-test se presentan en la tabla 1.

Tabla 1. Valores de confiabilidad, dificultad y correlación ítem-test

Dimensión	Ítems	Coeficiente KR-20	Dificultad de los ítems	Correlación ítem-test
Introducción a la Enfermería	P1, P2, P3, P4, P5	0,7382	0,7304	0,5153
Bioquímica	P6, P7, P8, P9	0,7304	0,7892	0,5449
Tecnología del Cuidado	P10, P11, P12, P13	0,7235	0,6863	0,5308
Cuidado Integral de la Familia	P14, P15, P16, P17, P18	0,7317	0,6275	0,4858
Epidemiología	P19, P20, P21, P22, P23	0,7713	0,7157	0,5565
Enfermería en Atención del Adulto I y II	P24, P25, P26, P27, P28, P29, P30	0,7371	0,7098	0,4916
Total	30	0,7627	0,7075	0,2869

El coeficiente KR-20 total fue de 0,7627, lo que indica una confiabilidad aceptable del instrumento en su conjunto, al superar el umbral mínimo recomendado de 0,70. Todas las dimensiones presentaron valores de KR-20 superiores a 0,72. Epidemiología fue la dimensión con mayor confiabilidad (KR-20 = 0,7713), mientras que Tecnología del Cuidado obtuvo el valor más bajo (KR-20 = 0,7235). En conjunto, estos resultados sugieren una consistencia interna adecuada en las dimensiones evaluadas.

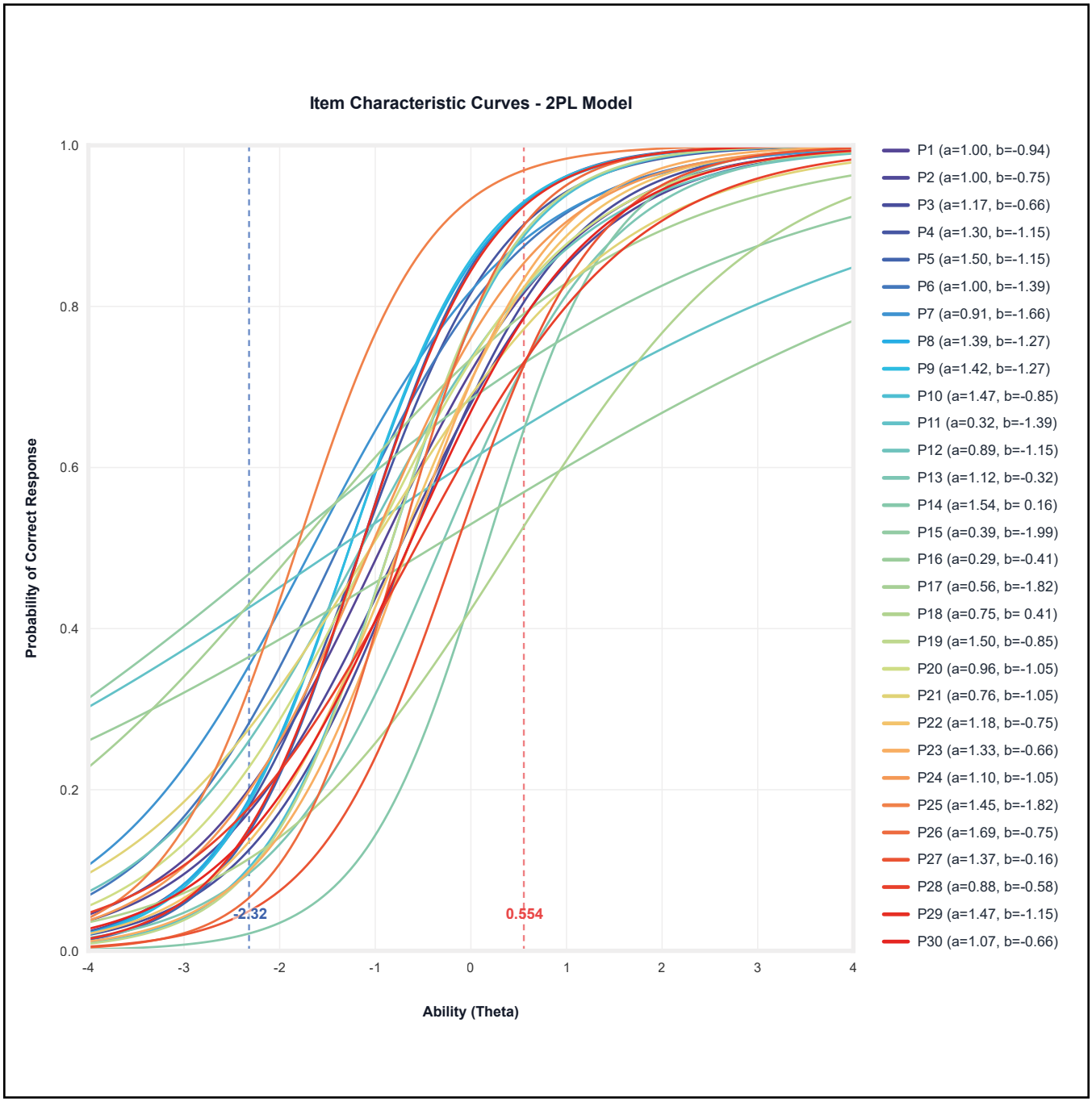
La dificultad promedio general fue de 0,7075, lo que indica que el cuestionario presenta una dificultad moderada-alta. La dimensión con mayor dificultad fue Bioquímica (0,7892), mientras que la de menor dificultad fue Cuidado Integral de la Familia (0,6275). El rango de dificultad entre dimensiones (0,6275-0,7892) evidencia variabilidad adecuada en el nivel de exigencia de las áreas evaluadas.

A nivel de ítems, el reactivo más fácil fue el P15 ($\theta = -2,32$), mientras que el más difícil fue el P18 ($\theta = 0,554$), tal como se observa en las curvas características de los reactivos (figura 1). Respecto a la discriminación, la correlación ítem-test total promedio fue de 0,2869, valor relativamente bajo que podría reflejar cierta heterogeneidad en el constructo global evaluado. En cambio, las correlaciones ítem-dimensión fueron más sólidas, con

valores promedio entre 0,4858 y 0,5565. Epidemiología presentó la mejor correlación promedio (0,5565), lo que sugiere mayor coherencia entre sus reactivos.

En conjunto, los resultados indican que el instrumento presenta confiabilidad aceptable, niveles adecuados de dificultad y mayor consistencia interna a nivel dimensional que global. No obstante, la correlación ítem-test total relativamente baja sugiere la conveniencia de revisar la estructura general del instrumento, especialmente si se pretende interpretar un puntaje total como representación de un único constructo.

Figura 1
Curvas características de los reactivos



Los resultados del AFC evidenciaron limitaciones sustanciales en el ajuste global del modelo con todos los ítems. Aunque la razón χ^2/gl sugirió un nivel aceptable de parsimonia, los demás indicadores mostraron deficiencias en la capacidad del modelo para reproducir la estructura de covarianza observada. El estadístico M2 fue de 588,69 con 300 grados de libertad y un valor $p < 0,001$, lo que indica diferencias significativas entre el modelo y los datos. Los índices de ajuste también fueron desfavorables (RMSEA = 0,138; SRMR = 0,188; TLI = 0,578; CFI = 0,301), por lo que el ajuste global debe interpretarse como problemático.

Figura 2
Ajuste individual de los ítems según estadísticos Infit y Outfit

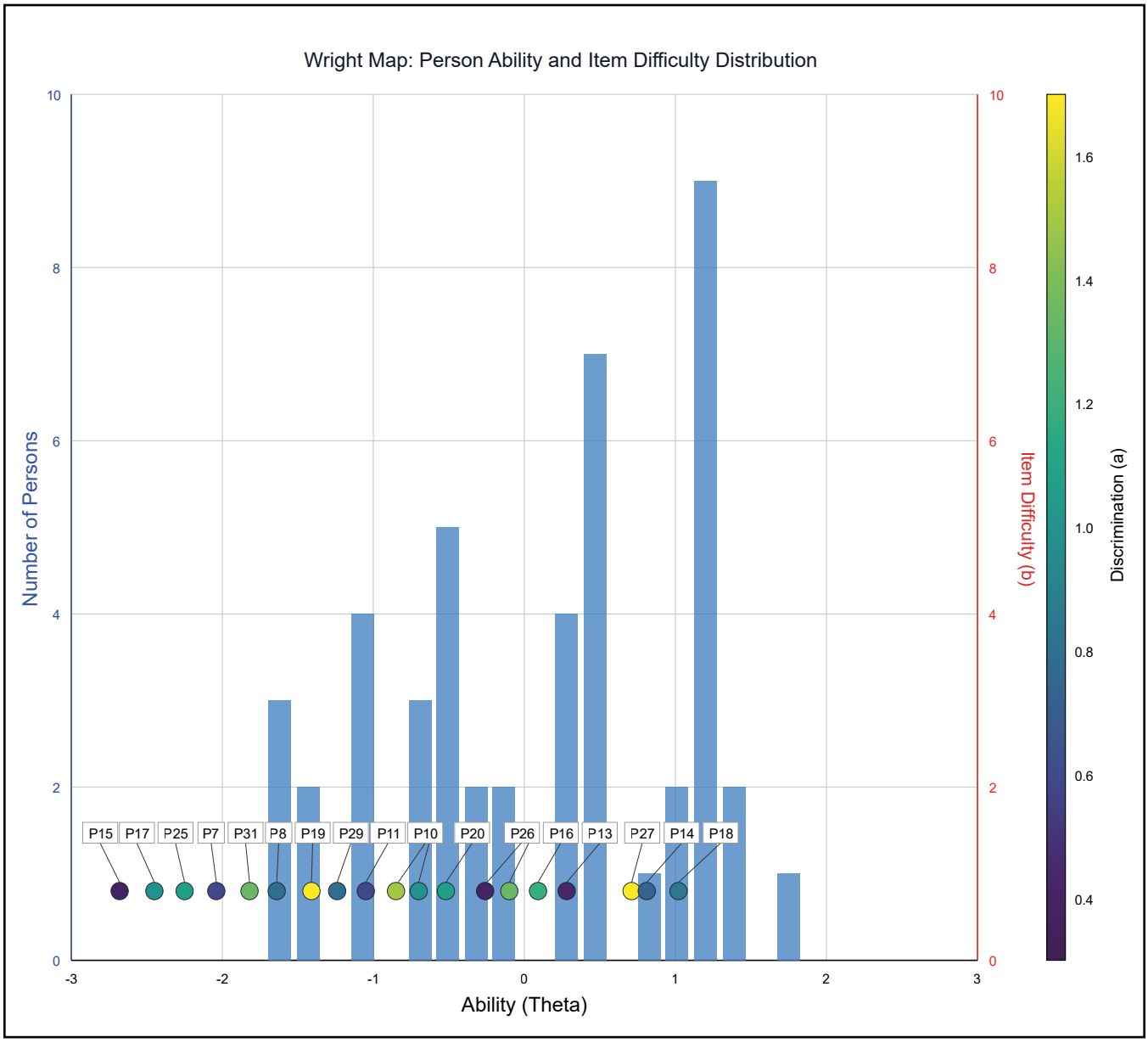


No obstante, el análisis a nivel de ítems mostró resultados más favorables. Los estadísticos de ajuste individual Infit y Outfit se mantuvieron dentro del rango productivo para la medición (0,5-1,5) en todos los reactivos, con valores Infit cercanos a 1,0. Los ítems P15 y P17 mostraron un ligero sobreajuste (Outfit < 0,7), lo que sugiere que podrían ser demasiado predecibles o redundantes. La mayoría de reactivos se ubicó dentro del rango aceptable de ajuste (0,7-1,3), lo que indica un funcionamiento adecuado de los ítems dentro del modelo TRI. El ítem P10, por su mayor tendencia al desajuste, podría beneficiarse de una revisión de redacción y claridad conceptual; en tanto, P15 y P17 podrían revisarse por posible redundancia si se busca abreviar el instrumento (figura 2).

La figura 3 presenta el mapa de personas-ítems. Se observó una distribución ligeramente sesgada hacia valores positivos de habilidad, con la mayoría de participantes ubicadas entre −1,0 y +1,0 logits y una concentración relevante alrededor de +1,0 logits. Los parámetros de dificultad oscilaron entre −1,99 y +0,41 logits (M = −0,94; DE = 0,55). Los ítems de mayor dificultad fueron P18, P14 y P27, mientras que los más fáciles fueron P15, P17 y P25.

Los parámetros de discriminación oscilaron entre 0,29 y 1,69 (M = 1,09; DE = 0,38). Los ítems P26, P14, P5, P19 y P29 mostraron mayor discriminación (a > 1,47), mientras que P16, P11, P15, P17 y P18 presentaron menor discriminación (a < 0,75). Asimismo, se observó una correlación positiva débil entre dificultad y discriminación (r = 0,19). Estos hallazgos sugieren una correspondencia adecuada entre la distribución de habilidades de las participantes y la dificultad de los ítems, así como una cobertura razonable del constructo evaluado.

Figura 3
Mapa de personas-ítems



La validez convergente evalúa el grado en que los ítems de una misma dimensión se relacionan con su propio dominio, mientras que la validez divergente verifica si dichas relaciones son mayores que las observadas con dimensiones diferentes. En este estudio, 24 de 25 ítems evaluados (96,0 %) presentaron correlaciones superiores a 0,30 con el puntaje de su propia dimensión. De igual manera, 24 de 25 ítems (96,0 %) mostraron mayor correlación con su dimensión correspondiente que con otras dimensiones. Estos resultados respaldan una adecuada coherencia interna y patrones de correlación consistentes con la estructura teórica del instrumento.

Tabla 2. Resumen de evidencias de validez convergente y divergente

Evidencia evaluada	Criterio considerado	Resultado	Interpretación
Validez convergente	Correlación del ítem con su propia dimensión > 0,30	24/25 ítems (96,0 %)	Evidencia favorable
Validez divergente	Mayor correlación con la dimensión propia que con dimensiones diferentes	24/25 ítems (96,0 %)	Evidencia favorable

DISCUSIÓN

El presente estudio tuvo como objetivo evaluar la validez y la confiabilidad de un cuestionario para medir el progreso académico, las habilidades técnicas y las actitudes profesionales de estudiantes de Enfermería a mitad de carrera. El instrumento, conformado por 30 ítems distribuidos en seis dimensiones, se alinea con la necesidad de contar con herramientas de evaluación pertinentes para la formación basada en competencias en ciencias de la salud (Albarqouni et al., 2018; Kumaravel et al., 2020; Ranganathan y Caduff, 2023).

La confiabilidad del cuestionario fue aceptable, con un KR-20 total de 0,7627. Además, todas las dimensiones presentaron coeficientes superiores a 0,72, lo que indica una consistencia interna adecuada. Epidemiología mostró el valor más alto, mientras que Tecnología del Cuidado presentó el valor más bajo, aunque dentro de niveles aceptables. Estos resultados sugieren que el cuestionario puede utilizarse como una medida inicial confiable para valorar conocimientos y competencias en las dimensiones evaluadas.

La validez convergente y divergente también fue favorable. La mayoría de los ítems se relacionó de manera más intensa con su propia dimensión que con otras, lo que respalda la coherencia conceptual del instrumento. No obstante, los valores bajos de escalabilidad observados en algunas dimensiones indican que ciertos reactivos deben ser revisados. En particular, los ítems con menor escalabilidad podrían requerir ajustes en su redacción, claridad conceptual o correspondencia con los resultados de aprendizaje de cada curso.

Estos hallazgos coinciden parcialmente con validaciones de cuestionarios en estudiantes de Enfermería orientados a medir constructos vinculados con la formación profesional (Hosseini et al., 2025; Huang et al., 2022; Kumaravel et al., 2020). Los cuestionarios son herramientas útiles para recopilar información educativa; sin embargo, su calidad depende de un diseño cuidadoso, de la claridad de los ítems y de la acumulación de evidencias de validez y confiabilidad en distintos contextos.

Un estudio comparable fue realizado por Asiandi et al. (2021), quienes evaluaron una escala de valores profesionales en Enfermería estructurada en dimensiones como cuidado, activismo y profesionalismo. Por su parte, Srivastava et al. (2024) desarrollaron un cuestionario de 36 ítems para valorar competencias clínicas percibidas en egresados de Enfermería, organizado en secciones relacionadas con conocimientos conceptuales, procedimientos básicos, procedimientos avanzados y procedimientos vinculados con infusión y transfusión, reportando alta confiabilidad.

En el presente estudio, los índices globales del modelo mostraron un ajuste insuficiente (M2 = 588,69; gl = 300; p < 0,001; RMSEA = 0,138; SRMR = 0,188; TLI = 0,578; CFI = 0,301). Estos resultados sugieren que la estructura propuesta no reproduce completamente las relaciones observadas entre los ítems. Esta limitación puede explicarse por la complejidad del constructo evaluado, la posible presencia de dimensiones adicionales no incluidas en el modelo inicial, la dependencia local entre algunos reactivos o el tamaño reducido de la muestra. Además, las dimensiones fueron definidas a partir de la organización curricular, y no mediante un análisis factorial exploratorio previo.

A pesar de estas limitaciones, el análisis individual de los ítems fue favorable. Los valores Infit y Outfit se ubicaron dentro del rango productivo para la medición, lo que indica que los reactivos funcionan de manera razonablemente coherente con el modelo. Esta aparente diferencia entre el ajuste global y el ajuste por ítems sugiere que el instrumento puede ser útil a nivel práctico, aunque su estructura interna requiere refinamiento psicométrico en investigaciones posteriores.

Hosseini et al. (2025) reportaron evidencias psicométricas más sólidas al evaluar una escala de valores profesionales en estudiantes de Enfermería con una muestra mayor (n = 439) y una estructura de cinco factores: cuidado, profesionalismo, activismo, responsabilidad y compromiso social. De manera similar, Huang et al. (2022), con una muestra de 246 participantes, informaron resultados consistentes en la evaluación de competencias de Enfermería. La diferencia en el tamaño muestral podría explicar, al menos parcialmente, las diferencias observadas respecto al ajuste del modelo.

También existen coincidencias parciales con estudios realizados en profesionales de Enfermería. Solera-Gómez et al. (2024) adaptaron y validaron un cuestionario de actitudes en una muestra española, evaluando imagen de la Enfermería, empatía médica, valores profesionales y habilidades de comunicación. Sus análisis confirmatorios respaldaron un modelo jerárquico de 21 ítems con buen ajuste global y consistencia interna adecuada. Por otro lado, Çalışkan y Şenyuva (2022) desarrollaron una escala para evaluar la competencia profesional de enfermeras y reportaron correlaciones ítem-total elevadas y un alfa de Cronbach de 0,98, en una muestra considerablemente mayor.

La principal limitación del presente estudio fue el tamaño reducido de la muestra y el muestreo por conveniencia, aspectos que limitan la generalización de los resultados. Además, la totalidad de participantes fue de sexo femenino, por lo que no fue posible comparar el desempeño del instrumento según sexo ni explorar diferencias por grupos etarios amplios. Futuras investigaciones deberían incluir muestras más grandes, distintas instituciones, análisis factorial exploratorio previo, evaluación de dependencia local y estudios longitudinales que permitan examinar la estabilidad temporal del instrumento.

De manera complementaria, sería valioso incorporar enfoques cualitativos para explorar la percepción de estudiantes y docentes sobre la pertinencia de los ítems, la claridad de las preguntas y la correspondencia entre el instrumento y los resultados de aprendizaje esperados. Esta retroalimentación permitiría refinar el cuestionario y fortalecer su utilidad como herramienta de evaluación curricular.

CONCLUSIONES

Los resultados sugieren que el cuestionario presenta evidencias iniciales favorables de confiabilidad y validez para evaluar el progreso académico, las habilidades técnicas y las actitudes profesionales en estudiantes de Enfermería a mitad de carrera. La consistencia interna fue aceptable tanto a nivel global como por dimensiones, y la mayoría de los ítems mostró un comportamiento adecuado en los análisis de ajuste individual.

Aunque el instrumento evidenció una adecuada correspondencia entre la dificultad de los ítems y la habilidad de las participantes, los índices globales de ajuste revelaron limitaciones en la estructura interna. Por ello, se recomienda revisar los ítems con menor escalabilidad, explorar posibles dimensiones adicionales y validar el cuestionario en muestras más amplias y heterogéneas.

En términos prácticos, el instrumento puede ser útil para que las instituciones educativas identifiquen fortalezas y áreas de mejora en la formación de estudiantes de Enfermería, orienten estrategias de enseñanza y fortalezcan la evaluación basada en competencias. No obstante, su uso debe considerarse como parte de un proceso continuo de validación y mejora psicométrica.

Conflicto de intereses:
Los autores declaran no tener conflicto de intereses.

Financiamiento:
Autofinanciado.

Contribución de autoría:
Todos los autores colaboraron en la misma medida y son responsables del contenido de este artículo.

Correspondencia:
Daniel Jose Blanco Victorio
✉ daniel.blanco.v@upch.pe

REFERENCIAS

Albarqouni, L., Hoffmann, T., Straus, S., Olsen, N. R., Young, T., Ilic, D., Shaneyfelt, T., Haynes, R. B., Guyatt, G. y Glasziou, P. (2018). Core competencies in evidence-based practice for health professionals: consensus statement based on a systematic review and Delphi survey. *JAMA Network Open*, *1*(2), e180281. https://doi.org/10.1001/jamanetworkopen.2018.0281

Altmiller, G. (2019). Content validation of quality and safety education for nurses prelicensure clinical evaluation instruments. *Nurse Educator*, *44*(3), 118-121. https://doi.org/10.1097/NNE.0000000000000656

American Association of Colleges of Nursing (2026). *The Essentials: Core competencies for professional nursing education.* https://www.aacnnursing.org/Portals/0/PDFs/Publications/Essentials-2021.pdf

Asiandi, A., Erlina, M., Lin, Y.-H. y Huang, M.-C. (2021). Psychometric evaluation of the Nurses Professional Values Scale-3: Indonesian version. *International Journal of Environmental Research and Public Health*, *18*(16), 8810. https://doi.org/10.3390/ijerph18168810

Çalışkan, F. y Şenyuva, E. (2022). A valid and reliable tool to assess nursing professional competences: the Nursing Professional Competence Scale. *Medical Journal of Bakirkoy*, *18*(2), 177-188. https://doi.org/10.4274/BMJ.galenos.2021.2021.11-2

Finstad, I., Knutstad, U., Havnes, A. y Sagbakken, M. (2022). The paradox of an expected level: the assessment of nursing students during clinical practice - A qualitative study. *Nurse Education in Practice*, *61*, 103332. https://doi.org/10.1016/j.nepr.2022.103332

Galura, S. y Warshawsky, N. (2022). Initial evaluation of a Doctor of Nursing Practice - Executive track program: The development of a three-year process to implement the new AACN Essentials. *Journal of Professional Nursing*, *42*, 276-280. https://doi.org/10.1016/j.profnurs.2022.07.014

Giddens, J. y Mansfield, J. A. (2023). A new era for nursing education: implications for practice. *Journal for Nurses in Professional Development*, *39*(1), 1-4. https://doi.org/10.1097/NND.0000000000000955

Hosseini, F. A., Zarshenas, L., Parvan, K., Shaygan, M., Tehranineshat, B. y Thomson, B. (2025). Assessment of the validity and reliability of the nurses' professional values scale-Revised in Persian version among nursing students. *BMC Medical Education*, *25*(1), 125. https://doi.org/10.1186/s12909-025-06677-y

Huang, S.-M., Fang, S.-C., Hung, C.-T. y Chen, Y.-H. (2022). Psychometric evaluation of a nursing competence assessment tool among nursing students: a development and validation study. *BMC Medical Education*, *22*(1), 372. https://doi.org/10.1186/s12909-022-03439-y

Hughes, L. J., Mitchell, M. L. y Johnston, A. N. (2019). Just how bad does it have to be? Industry and academic assessors' experiences of failing to fail - A descriptive study. *Nurse Education Today*, *76*, 206-215. https://doi.org/10.1016/j.nedt.2019.02.011

Jenkins-Weintaub, E., Goodwin, M. y Fingerhood, M. (2023). Competency-based evaluation: collaboration and consistency from academia to practice. *Journal of the American Association of Nurse Practitioners*, *35*(2), 142-149. https://doi.org/10.1097/JXX.0000000000000830

Kumaravel, B., Hearn, J. H., Jahangiri, L., Pollard, R., Stocker, C. J. y Nunan, D. (2020). A systematic review and taxonomy of tools for evaluating evidence-based medicine teaching in medical education. *Systematic Reviews*, *9*(1), 91. https://doi.org/10.1186/s13643-020-01311-y

Lee, J. E. y Sim, I. O. (2019). Gap between college education and clinical practice: experience of newly graduated nurses. *Nursing Open*, *7*(1), 449-456. https://doi.org/10.1002/nop2.409

Lin, C.-C., Wu, C.-J. J., Hsiao, Y.-C., Han, C.-Y. y Hung, C.-C. (2017). An exploratory factor analysis for developing and validating a scale of Nursing Students Competence Instrument. *Nurse Education Today*, *50*, 87-91. https://doi.org/10.1016/j.nedt.2016.12.007

Löfgren, U., Wälivaara, B.-M., Strömbäck, U. y Lindberg, B. (2023). The nursing process: A supportive model for nursing students' learning during clinical education - A qualitative study. *Nurse Education in Practice*, *72*, 103747. https://doi.org/10.1016/j.nepr.2023.103747

Nilsson, J., Mischo-Kelling, M., Thiekoetter, A., Deufert, D., Mendes, A. C., Fernandes, A., Kirchhoff, J. W. y Lepp, M. (2019). Nurse professional competence (NPC) assessed among newly graduated nurses in higher educational institutions in Europe. *Nordic Journal of Nursing Research*, *39*(3), 159-167. https://doi.org/10.1177/2057158519845321

Purabdollah, M., Zamanzadeh, V., Ghahramanian, A., Valizadeh, L., Mousavi, S. y Ghasempour, M. (2023). Competencies expected of undergraduate nursing students: a scoping review. *Nursing Open*, *10*(12), 7487-7508. https://doi.org/10.1002/nop2.2020

Ranganathan, P. y Caduff, C. (2023). Designing and validating a research questionnaire-Part 1. *Perspectives in Clinical Research*, *14*(3), 152-155. https://doi.org/10.4103/picr.picr_140_23

Solera-Gómez, S., Sancho-Cantus, D., Privado, J., Cunha-Pérez, C., Benedito-Monleón, A. y Llinares-Insa, L. I. (2024). Adaptation and validation of the Nursing Attitude Questionnaire in a Spanish population. *Nursing Reports*, *14*(4), 2758-2776. https://doi.org/10.3390/nursrep14040204

Srivastava, S., Pandurangan, H. y Kumar, A. (2024). Development and validation of a Comprehensive Nursing Competence Assessment Questionnaire (CNCAQ) to determine the perceived clinical competence of nursing graduates. *Nursing & Midwifery Research Journal*, *20*(2), 96-111. https://doi.org/10.1177/0974150X241247986

Suliman, M., Aljezawi, M., Tawalbeh, L., Ta'an, W., Abdalrhim, A., Albashtawy, M. y Alkhawaldeh, A. (2023). Assessing professional competencies among undergraduate nurses: an exploratory study. *Nursing Education Perspectives*, *44*(3), E10-E14. https://doi.org/10.1097/01.NEP.0000000000001074

World Health Organization (2015). *Roles and Responsibilities of Government Chief Nursing and Midwifery Officers: A* capacity-building *manual*. https://iris.who.int/handle/10665/351684