

Validation of Instruments for Monitoring and Clinical Assessment During the Internal Medicine Clerkship Through Expert Judgment

¹⁻ Viorkis Pérez Ortiz, ²⁻ Nicolás Sotero Rodríguez León, ³⁻ Fanny Soraya Reyes, ⁴⁻ Johanna Elizabeth Alcivar Alcivar, ⁵⁻ María Fernanda Joza Vera, ⁶⁻ Valeria Sophia Navas Benavides

¹Universidad Técnica de Manabí.

Instituto Universitario del Rosario (IUNIR) Doctorado en Ciencias de la Salud.

Email. viorkis.perez@utm.edu.ec

ORCID: <https://orcid.org/0000-0002-5921-2484>

²Instituto Universitario Italiano de Rosario (IUNIR).

Universidad Nacional de Educación a Distancia (UNED).

nrleon2010@yahoo.com.ar

ORCID: <https://orcid.org/0000-0003-3335-641X>

³Universidad Técnica de Manabí.

Email. fanny.reyes@utm.edu.ec

ORCID: <https://orcid.org/0000-0001-6782-1653>

⁴Universidad Técnica de Manabí.

Email. johanna.alcivar@utm.edu.com

ORCID: <https://orcid.org/0009-0007-0768-028X>

⁵Universidad Técnica de Manabí.

Email. maria.joza@utm.edu.ec

ORCID: 0000-0003-1786-8079

⁶Universidad Técnica de Manabí

Email. valerianavas1993@gmail.com

ORCID: <https://orcid.org/0009-0000-3489-491X>

HOW TO CITE:

Viorkis Pérez Ortiz, Nicolás Sotero Rodríguez León, Fanny Soraya Reyes, Johanna Elizabeth Alcivar Alcivar, María Fernanda Joza Vera, Valeria Sophia Navas Benavides (2026). Validation of Instruments for Monitoring and Clinical Assessment During the Internal Medicine Clerkship Through Expert Judgment. International Journal of Special Education, 41(12s), 901-911

ABSTRACT:

Introduction: Clinical assessment during the Internal Medicine clerkship requires valid instruments that enable the objective evaluation of students' learning progress and competency development. At the Universidad Técnica de Manabí, no methodologically validated instruments were available for this purpose. Objective: To validate, through expert judgment, the content of a questionnaire designed for medical students and a semi-structured interview guide for clinical faculty aimed at assessing monitoring and clinical evaluation during the Internal Medicine clerkship. Methodology: A methodological instrument validation study was conducted between January and March 2026 at the Universidad Técnica de Manabí. The instruments were developed based on a literature review and the principles of competency-based medical education. Seven experts holding doctoral degrees and with experience in medical education, university teaching, clinical assessment, and research evaluated the clarity, relevance, pertinence, and coherence of the items using a five-point ordinal scale. Content validity was determined using Aiken's V coefficient. Main Results: The questionnaire for medical students achieved a mean Aiken's V coefficient of 0.854, while the semi-structured interview guide for clinical faculty obtained a

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.
Open access publication under the
terms and conditions
of the Creative Commons Attribution
(CC BY)
license (<http://creativecommons.org/licenses/by/4.0/>).

mean coefficient of 0.890. All items exceeded the minimum acceptability criterion, demonstrating strong expert agreement and confirming the content representativeness of both instruments. Conclusion: Both instruments demonstrated adequate content validity, constituting methodologically robust tools for monitoring and assessing clinical performance during the Internal Medicine clerkship. Keywords: Medical Education; Validation Studies; Reproducibility of Results; Clinical Competence; Internship and Residency.

Introduction

The assessment process in medical education has undergone significant methodological transformations in recent decades, evolving from approaches focused on measuring theoretical knowledge via summative assessments toward innovative models centered on the development of clinical competencies. This necessary evolution has fostered more comprehensive assessment strategies; however, the challenge remains to train physicians capable of meeting the growing demands of healthcare systems with quality, ethics, and scientific rigor (1).

Clinical training is widely recognized as an essential component of the medical curriculum, as it integrates knowledge acquired during the preclinical phase with the development of patient care skills, clinical reasoning, diagnostic capabilities, and communication competencies within real-world healthcare settings (2-6).

In this context, the Internal Medicine clinical clerkship represents a pivotal stage, as students actively participate in comprehensive patient assessment, the analysis of clinical problems, and therapeutic decision-making under faculty supervision, thereby progressively strengthening their professional competencies and their identity as future physicians (7-10).

The complexity of this process requires mechanisms that allow for the continuous assessment of the student's progress and timely guidance of their learning. In line with current trends in health education, these processes should be based on formative feedback, ongoing performance monitoring, and active student participation (11-15). However, their

implementation requires methodologically sound tools that provide valid information about the dimensions that make up both monitoring and clinical evaluation.

The lack of properly validated tools can compromise the objectivity of the evaluation process, limit the quality of pedagogical feedback, and hinder the systematic monitoring of competency development during the externship. These considerations will establish that instrument validation is an essential stage in educational research, providing evidence in a clear, relevant, pertinent, and coherent manner regarding the construct that is intended to be measured.

At the Technical University of Manabí, there are no validated instruments available to assess the monitoring and clinical evaluation during the Internal Medicine externship from the perspective of students and teachers. Considering the above, the present study aimed to validate, through expert judgment, the content of a questionnaire directed at medical students and a semi-structured interview guide for clinical teachers, both designed to evaluate monitoring and clinical assessment during the Internal Medicine externship.

Methodology**Study design, context, and execution period**

This study was a methodological research of an instrumental type, with a quantitative approach and cross-sectional design, aimed at developing and validating content for instruments used in monitoring and clinical evaluation during the Internal Medicine internship at the Technical University of Manabí (UTM), Ecuador. The

research was conducted between January and March 2026 and was focused on obtaining content validity evidence through expert judgment, as an initial stage of the validation process for educational assessment instruments.

Development of the instruments

The development of the instruments was carried out through a systematic process based on a review of scientific literature related to medical education, competency-based assessment, clinical evaluation, formative feedback, and performance monitoring during the externship. Additionally, contemporary models of medical education focused on the progressive development of professional competencies and continuous assessment of clinical learning were considered.

Based on the identified evidence, the construct of interest and its conceptual dimensions were defined, from which the first version of two complementary instruments was developed. Subsequently, the items were reviewed by the research team to verify their alignment with the study's objectives, the clarity of the language, and the coherence between each dimension and the proposals.

Description of the instruments

Two instruments were developed.

The first consisted of a structured questionnaire aimed at Medical students to assess their perception of the follow-up and clinical evaluation during the Internal Medicine internship. The instrument was composed of two sections. The first gathered general information about the participants (age, gender, academic period, and duration of the internship). The second included 25 items distributed across five dimensions: clarity and relevance of the evaluation model, feedback and performance improvement, integration of evaluation strategies, traceability of the evaluation process, and satisfaction with the evaluation model. All items used a five-point Likert scale response format (1 = strongly disagree to 5 = strongly agree).

The second instrument was a semi-structured interview guide aimed at clinical teachers,

consisting of 10 open-ended questions organized into five thematic areas: assessment of clinical performance, feedback, student participation in the evaluation process, the role of the supervising teacher, and proposals for improving the evaluation process. Its purpose was to qualitatively explore teachers' perceptions regarding pedagogical follow-up and clinical assessment during the externship.

Panel of experts and selection criteria

The content validity of both instruments was conducted through expert judgment using purposive sampling, a strategy widely used in methodological studies because it prioritizes the selection of specialists with specific experience in the phenomenon under study.

The panel was composed of seven experts, all holding a doctoral degree (PhD) and experience in medical education and university teaching. Five of them also had experience in clinical evaluation, and two had a solid background in scientific research related to medical training. The selection criteria included a fourth-level academic degree, university teaching experience, participation in clinical evaluation processes, scientific production, and professional experience related to the construct being evaluated..

Content validation procedure

Each expert received electronically the preliminary version of both instruments, along with the study's objective, the conceptual definition of the evaluated dimensions, and a rating matrix specifically designed for the validation process.

Each item was evaluated considering four criteria: clarity, pertinence, relevance, and coherence.

The assessment was carried out using a five-category ordinal scale, where:

- 1: Does not meet
- 2: Partially meets
- 3: Adequately meets
- 4: Fully meets
- 5: Completely meets

Additionally, the experts provided qualitative observations and recommendations aimed at improving the wording, understanding, organization, and conceptual accuracy. These observations were analyzed by the research team and considered for making adjustments to the form and content before obtaining the final version of the instruments. No items were eliminated during the validation process.

Statistical analysis

Content validity was determined using Aiken's V coefficient, which is widely recommended for quantifying the degree of agreement among experts regarding the appropriateness of the items in an instrument.

The coefficient was calculated using the following expression.:

$$V = \frac{\sum s}{n(c - 1)}$$

Where:

- $s = r - l$
- r corresponds to the score given by each expert.
- l represents the minimum score on the scale.
- n corresponds to the number of experts ($n = 7$).
- c represents the number of response categories ($c = 5$).

For each item, independent coefficients were calculated for the criteria of clarity, pertinence, relevance, and coherence, in addition to a global coefficient obtained by averaging the four values. The interpretation of Aiken's V was carried out considering the following criteria:

- ≥ 0.90 : excellent content validity.
- $0.80-0.89$: high validity.
- $0.70-0.79$: acceptable validity.
- < 0.70 : item subject to review.

Additionally, the qualitative observations provided by the experts were analyzed descriptively to support the reformulation of those elements that presented suggestions for

improvement. The data were organized and processed using Microsoft Excel®, which was used for calculating Aiken's V coefficient, and IBM SPSS Statistics version 27, employed for the descriptive analysis of the scores obtained.

Ethical considerations

The study was approved by the Human Research Ethics Committee (CEISH) of the Technical University of Manabí, in accordance with institutional regulations for research involving human participants. Although the research did not involve clinical interventions or the handling of biological samples, the participation of experts in the validation process required the corresponding ethical review and approval.

All experts participated voluntarily, following informed consent, with guarantees regarding confidentiality, the anonymity of assessments, and the exclusive use of the information for scientific purposes. The study was conducted in accordance with ethical principles for research involving human subjects and applicable national and international regulations.

Results

The content validity of the two instruments designed for the monitoring and evaluation of the Internal Medicine externship was estimated using the V Aiken coefficient, based on the assessment provided by a panel of seven experts in medical education, scientific research, university teaching, and clinical evaluation. Each expert rated the items on the four established methodological criteria (clarity, pertinence, relevance, and coherence) using a four-category ordinal scale. For the interpretation of the coefficients, the cutoff points proposed in the specialized literature were considered: values ≥ 0.70 as acceptable, ≥ 0.80 as adequate, and ≥ 0.90 as excellent (Aiken, 1985).

Validez de contenido del cuestionario dirigido a estudiantes

The questionnaire aimed at Medical students, consisting of 25 items distributed across five dimensions, showed Aiken's V coefficients by criterion ranging from 0.714 to 0.964, with overall averages of 0.856 in clarity, 0.846 in relevance,

0.841 in pertinence, and 0.874 in coherence (Table 1). The total V coefficient per item ranged from 0.812 to 0.902, with a general average of 0.854, indicating an adequate level of content validity for the entire instrument. Only items 10 and 23 recorded, specifically, coefficients of 0.714 in relevance and coherence criteria respectively,

falling within the acceptable range; however, their total coefficients (0.839 and 0.812) remained within the appropriate range, so their elimination was not considered necessary, but rather a partial reformulation according to the qualitative observations of the panel.

Table 1. Aiken's V coefficients by criterion and item aimed at Medical students (k = 25, n = 7 experts)

Ítem	V Clarity	V Pertinence	V Relevance	V Coherence	V Total
Ítem 1	0,857	0,929	0,857	0,821	0,866
Ítem 2	0,929	0,929	0,893	0,857	0,902
Ítem 3	0,893	0,786	0,786	0,857	0,83
Ítem 4	0,821	0,857	0,821	0,786	0,821
Ítem 5	0,821	0,821	0,857	0,821	0,83
Ítem 6	0,857	0,786	0,857	0,857	0,839
Ítem 7	0,821	0,893	0,786	0,857	0,839
Ítem 8	0,821	0,857	0,786	0,786	0,812
Ítem 9	0,821	0,857	0,893	0,857	0,857
Ítem 10	0,821	0,893	0,714	0,929	0,839
Ítem 11	0,821	0,893	0,786	0,893	0,848
Ítem 12	0,821	0,786	0,821	0,821	0,812
Ítem 13	0,821	0,857	0,821	0,821	0,83
Ítem 14	0,929	0,786	0,821	0,929	0,866
Ítem 15	0,929	0,786	0,786	0,964	0,866
Ítem 16	0,929	0,821	0,857	0,929	0,884
Ítem 17	0,893	0,893	0,786	0,964	0,884
Ítem 18	0,821	0,929	0,857	0,964	0,893
Ítem 19	0,893	0,821	0,929	0,929	0,893
Ítem 20	0,821	0,786	0,929	0,893	0,857
Ítem 21	0,821	0,857	0,964	0,929	0,893
Ítem 22	0,821	0,821	0,964	0,893	0,875
Ítem 23	0,893	0,786	0,857	0,714	0,812
Ítem 24	0,821	0,857	0,821	0,821	0,83
Ítem 25	0,893	0,857	0,786	0,964	0,875
Overall average	0,856	0,846	0,841	0,874	0,854

Note. V = Aiken's V coefficient. Interpretation: ≥ 0.70 acceptable; ≥ 0.80 adequate; ≥ 0.90 excellent (Aiken, 1985).

Content validity of the interview guide aimed at clinical teachers

The semi-structured interview guide directed at clinical teachers, consisting of 10 questions, obtained Aiken's V coefficients for criteria ranging from 0.786 to 1.000, with overall averages

of 0.889 in clarity, 0.854 in relevance, 0.914 in pertinence, and 0.904 in coherence (Table 2). The total V coefficient per item ranged from 0.848 to 0.973, with a general average of 0.890, a value higher than that obtained in the questionnaire directed at students. Question 8 achieved the highest coefficient in the study ($V = 0.973$),

including total agreement among experts on the relevance criterion ($V = 1.000$), while question 2

had the lowest value ($V = 0.848$), without falling below the appropriate threshold.

Table 2. Aiken's V coefficients by criterion and item of the interview guide for clinical teachers (k = 10, n = 7 experts)

Ítem	V Clarity	V Pertinence	V Relevance	V Coherence	V Total
Ask 1	0,929	0,893	0,857	0,929	0,902
Ask 2	0,821	0,821	0,929	0,821	0,848
Ask 3	0,857	0,893	0,964	0,929	0,911
Ask 4	0,857	0,821	0,857	0,929	0,866
Ask 5	0,893	0,857	0,857	0,929	0,884
Ask 6	0,893	0,821	0,893	0,964	0,893
Ask 7	0,821	0,821	0,929	0,857	0,857
Ask 8	0,964	0,964	1	0,964	0,973
Ask 9	0,929	0,857	0,893	0,857	0,884
Ask 10	0,929	0,786	0,964	0,857	0,884
Overall average	0,889	0,854	0,914	0,904	0,89

Note. V = Aiken's V coefficient. Interpretation: ≥ 0.70 acceptable; ≥ 0.80 adequate; ≥ 0.90 excellent (Aiken, 1985).

Síntesis comparativa de los coeficientes de validez de contenido

Al comparar ambos instrumentos (Tabla 3), la guía de entrevista para docentes presentó coeficientes promedio superiores a los del cuestionario para estudiantes en tres de los cuatro

criterios evaluados (pertinencia, relevancia y coherencia), así como en el coeficiente V total (0,890 frente a 0,854). En ambos casos, los promedios generales se ubicaron en la categoría de validez adecuada, sin que ningún coeficiente promedio descendiera por debajo de 0,80.

Tabla 3. Comparación de los coeficientes V de Aiken promedio por instrumento y criterio

Instrumento	V Clarity	V Pertinence	V Relevance	V Coherence	V Total
Student Questionnaire (k = 25)	0,856	0,846	0,841	0,874	0,854
Teacher Interview Guide (k = 10)	0,889	0,854	0,914	0,904	0,89

Note. V = average V coefficient of Aiken. k = number of items in the instrument.

The classification of the items according to their total V coefficient (Table 4) confirms the consistency of both instruments: 96.0% of the items in the student questionnaire and 70.0% of the questions in the teaching guide were placed in the appropriate category (0.80–0.89), while 4.0% and 30.0%, respectively, reached the excellent

category (≥ 0.90). No item from either instrument obtained a total coefficient below 0.70, so no elements were eliminated; only wording adjustments were made to items with lower scores, based on the qualitative observations provided by the expert panel.

Table 4. Distribution of items according to the level of content validity (total Aiken's V coefficient)

Instrument / Category	Frequency	Percentage
		%
<i>Student Questionnaire (k = 25)</i>		
Appropriate (0,80–0,89)	24	96
Excellent (≥ 0,90)	1	4
<i>Teacher Interview Guide (k = 10)</i>		
Appropriate (0,80–0,89)	7	70
Excellent (≥ 0,90)	3	30

Note: No item was classified in the unacceptable category ($V < 0.70$).

Discussion of Results

The coefficients obtained through Aiken's V demonstrated a high level of consensus among experts regarding the clarity, pertinence, relevance, and coherence of the observed factors, which supports that both instruments adequately represent the conceptual domains defined during their development. These findings are especially important in the context of current medical education, where the assessment of clinical learning has evolved from models focused on knowledge certification to continuous processes aimed at the progressive development of professional competencies (1-15).

The development of valid instruments is one of the essential components to strengthen clinical assessment systems. Recent studies agree that implementing competency-based evaluation models requires tools capable of gathering relevant information about student performance from multiple dimensions, integrating processes of monitoring, feedback, and educational decision-making (19,23,24).

In this regard, the instruments developed in this research address this need by incorporating dimensions related to the clarity of the evaluation process, teacher feedback, the integration of assessment strategies, the traceability of clinical performance, and satisfaction with the training process—factors widely recognized as fundamental components of contemporary clinical evaluation (10-15).

The average coefficients obtained for the questionnaire ($V=0.854$) and the interview guide ($V=0.890$) show a high consistency in the evaluations made by the expert panel. These results are consistent with recent methodological research aimed at developing instruments for medical education and other fields of health sciences, where values above 0.80 constitute sufficient evidence of adequate content validity (18,21,22,25,29-35).

Beyond the numerical value achieved, the observed consensus reflects that the conceptual structure of both instruments consistently represents the phenomenon they aim to assess, strengthening the methodological quality of the construction process..

An important aspect was that none of the concepts or terms needed to be eliminated during validation. This result suggests that the design phase was based on a thorough review of the literature and a precise definition of the constructs evaluated, allowing expert judgment to function mainly as a mechanism for methodological refinement through adjustments in wording and conceptual accuracy.

This behavior aligns with current recommendations on content validation, which emphasize that the purpose of expert judgment is not limited to accepting or rejecting items but to optimize the representativeness of the content before moving on to later stages of psychometric validation (16,18,21,29-37).

The simultaneous validation of a questionnaire for students and a semi-structured interview for teachers constitutes one of the main methodological contributions of this research. While the former allows for quantifying students' perceptions through standardized indicators, the latter facilitates understanding of pedagogical processes from the experience of those supervising clinical learning. This methodological complementarity is consistent with current trends in programmatic evaluation, which recommend integrating multiple sources of evidence to obtain a more comprehensive assessment of clinical competency development (10-15,19,23). The combination of both instruments provides a more robust evaluation strategy, promoting data triangulation and expanding the possibilities for interpreting the training process during the externship. From the perspective of medical education, having instruments specifically designed for the Internal Medicine externship represents a significant advancement for the systematic assessment of clinical learning. Several authors have noted that the quality of evaluation processes largely depends on the use of valid tools that allow for documenting student progress, guiding feedback, and supporting evidence-based curricular decisions (11-15, 23, 39-41). In this context, the developed instruments offer a methodological alternative tailored to the characteristics of the externship at the Technical University of Manabí and provide an objective basis for future research on clinical performance and the improvement of training processes.

The results obtained should be understood as a primary source of evidence within the validation process. The specialized literature agrees that validity is an accumulative process supported by multiple pieces of evidence, so the appropriate content validity achieved in this research establishes the necessary conditions to develop subsequent studies aimed at evaluating internal consistency, temporal stability, and the factorial structure of both instruments (21,22,27-37).

Consequently, this work not only provides methodologically sound tools for monitoring and

clinical evaluation during Internal Medicine externship, but also offers a solid foundation for future research aimed at strengthening competency-based assessment in medical training.

Strengths and Limitations of the Study

The main strength of this study lies in the development and validation of content for two complementary instruments to assess follow-up and clinical evaluation during the Internal Medicine externship, based on a literature review, a systematic construction process, and the assessment by a panel of experts with extensive experience in medical education, university teaching, clinical evaluation, and research.

The use of Aiken's V coefficient, complemented by the qualitative analysis of expert observations, provided solid evidence regarding the clarity, relevance, pertinence, and coherence of the instruments. However, the study was limited to content validation, so future research should provide additional evidence related to other properties of the teaching process, as well as evaluate its applicability in different medical training contexts.

Conclusions

This study successfully validated the content of a questionnaire aimed at medical students and a semi-structured interview guide for teachers through expert judgment, achieving high levels of agreement in the criteria of clarity, pertinence, relevance, and coherence, evaluated using Aiken's V coefficient. These results support the conceptual appropriateness of both instruments for assessing follow-up and clinical evaluation during Internal Medicine externship, providing a solid methodological foundation for their use in future studies and for obtaining new evidence of validity and reliability in different medical training contexts.

Conflict of Interest

The authors declare that there are no conflicts of interest, financial, academic, professional, or personal.

Bibliographic References

- Vengrovych OZ, Tymkiv I, Kozinchuk HV, Mishchuk V, Tymkiv I. Educación médica moderna: desafíos, necesidades y oportunidades. *Patología Clínica y Experimental*. 2025. doi:10.24061/1727-4338.xxiv.3.93.2025.14.
- Koufakis T, Gouveri E, Doumas M, Papanas N. From an empty glass to an almost full one and finally to one that never fills: the changing challenges of medical education. *Int J Low Extrem Wounds*. 2023;25:16-18. doi:10.1177/15347346231210406.
- Isayeva O, Shumylo M. Integración de la educación médica basada en competencias en los planes de estudio de medicina de pregrado: estrategias, desafíos y perspectivas de futuro. *Innovate Pedagogy*. 2025. doi:10.32782/2663-6085/2025/83.1.31.
- Shrivastava S, Marar AM, Bobhate P, Petkar P, Borkar S. Fortalecimiento de la formación clínica para estudiantes de medicina en la India: desafíos existentes y posibles soluciones. *J Clin Sci*. 2025. doi:10.4103/jcls.jcls_71_24.
- Jalali K, Marzaleh AM, Faghihi SA, Hayat A, Rezaee R. Factores que influyen en la resiliencia de la formación clínica en medicina general en Irán: un estudio cualitativo de análisis de contenido. *BMC Med Educ*. 2025;25. doi:10.1186/s12909-025-07000-5.
- Mathew C, Mathew J, Philip SS. El papel esencial de la formación clínica en la educación médica de pregrado: una revisión sistemática. *J Assoc Physicians India*. 2026;74(3):99-102. doi:10.59556/japi.74.1063.
- Cadena CVP, Guacho EAL, González MTG, Mercado A. Percepciones de estudiantes de Medicina de 5.º a 7.º semestre sobre el impacto del externado rotativo en su salud mental. *Arandu UTIC*. 2025;12(3):1678-1692.
- Montejo Galdámez JD. Radiología e imágenes diagnósticas como parte clave en la formación médica. *Rev Postgr Med*. 2022;1(1):55-70. doi:10.62267/rev.post.med.v1i1.7.
- Ramírez Rozo Á. Evaluación de las competencias profesionales según el plan de estudios del internado rotatorio en el Programa de Medicina de la Universidad Antonio Nariño [tesis]. Universidad Externado de Colombia; 2024.
- Bartman I, St-Onge C, Roy M, Gingerich A, Katsoulas E, Chahine S, et al. Retroalimentación de múltiples fuentes en la educación médica de pregrado: un estudio piloto. *Can Med Educ J*. 2025;16:25-31. doi:10.36834/cmej.79283.
- Schmude M, McCoog I, Adonizio T, Ellison HB, Howe A, Caleb AM. Uso de un portafolio electrónico para la evaluación longitudinal del desarrollo personal y profesional en la formación médica de pregrado. *Front Med*. 2025;11. doi:10.3389/fmed.2024.1505378.
- Faria TV, Pereira AS, De Souza DM, Girard DR, Leandro CB, De Oliveira Valentino TC, et al. Evaluación formativa en educación médica utilizando metodologías de aprendizaje activo: impactos en el aprendizaje autorregulado. *Rev Bras Desenvol*. 2025. doi:10.34117/bjdv11n12-062.
- Fuentes-Cimma J, Sluijsmans DMA, Riquelme A, Villagrán I, Isbej L, Olivares-Labbe MT, et al. Diseño de procesos de retroalimentación en el aprendizaje basado en el lugar de trabajo en la formación de pregrado en profesiones de la salud: una revisión exploratoria. *BMC Med Educ*. 2024;24. doi:10.1186/s12909-024-05439-6.
- Singh RA. Redefining evaluation: Programmatic assessment in Indian medical education. *Int J Recent Innov Med Clin Res*. 2024. doi:10.18231/j.ijrimcr.2024.044.
- Salamy AYMS, Claramita M, Suhoyo Y. La relevancia de la retroalimentación constructiva y la reflexión estudiantil para abordar el currículo basado en competencias: un estudio explicativo de las facultades de medicina en Indonesia. *Jurnal Pendidikan Kedokteran Indonesia*. 2024. doi:10.22146/jpki.87750.
- Maldonado-Suárez N, Santoyo-Telles F. Validez de contenido por juicio de expertos: integración cuantitativa y cualitativa en la construcción de instrumentos de medición. *REIRE Rev Innov Rec Educ*. 2024;17(2):1-19. doi:10.1344/reire.46238.
- Domínguez Celis G, Mendoza Anaya L, Reza López SA, Ishida Gutiérrez MC, Fierro Fierro O. Diseño y validación de un instrumento para evaluar factores intrínsecos y extrínsecos del rendimiento académico en estudiantes de medicina. *Rev Esp Educ Med*. 2026;7(2). doi:10.6018/edumed.696751.

- Merino Soto C. Coeficientes V de Aiken: diferencias en los juicios de validez de contenido. *MHSalud*. 2023;20(1):23-32. doi:10.15359/mhs.20-1.3.
- Qurratu D, Suhoyo Y, Budi A, Claramita M. Desarrollo de un instrumento de autoevaluación con componentes de evaluación programática para estudiantes de medicina de pregrado. *Eur J Educ Res*. 2023. doi:10.12973/eu-jer.12.2.649.
- Dyk LHV, Jordaan H. ¿Algo no lo es todo? Diseño de un instrumento que mide el entorno de aprendizaje clínico de forma holística. *Int J Pract Based Learn Health Soc Care*. 2024. doi:10.18552/ijpblhsc.v12i1.818.
- Martini-Blanquel HA. Validación de instrumentos clínicos: aspectos clave. *Aten Fam*. 2024. doi:10.22201/fm.14058871p.2024.388838.
- Samir AA, Abdelsayed K, Ahmed HAA, Mohamed AA, Bahr A, Al-Bawah N, et al. Una versión árabe validada del cuestionario de evaluación del aprendizaje clínico para estudiantes de medicina e internos. *Med Educ Online*. 2026;31. doi:10.1080/10872981.2026.2632522.
- Kang Y, Kim HW, Choi J. Desarrollo de un marco de evaluación de programas para mejorar la calidad de la educación médica de pregrado. *Korean J Med Educ*. 2025. doi:10.17496/kmer.24.035.
- Li F, Zhao H, Huang Q, Zhang C, Chen G. Desarrollo y validación de un marco de evaluación basado en competencias para estudiantes de medicina clínica en la era de la inteligencia artificial: un estudio de métodos mixtos en China. *BMC Med Educ*. 2026;26. doi:10.1186/s12909-026-08779-7.
- Rincón MD, Franco PCA, Torrado JAV, Díaz OLL. Validación de contenido de un cuestionario sobre la percepción del personal sanitario sobre las tecnologías. *Rev Cuidarte*. 2024;16. doi:10.15649/cuidarte.4145.
- Tomas N, Italo M, Eva B, Veronica L. Evaluación durante la formación clínica de estudiantes de enfermería mediante dos instrumentos de evaluación diferentes. *BMC Med Educ*. 2024;24. doi:10.1186/s12909-024-05771-x.
- Song X, Cleaves ES, Gluzman E, Kotlyar B, Russo RA, Schilling DC, et al. Una revisión exploratoria de las evaluaciones en la educación médica de pregrado: implicaciones para los programas de residencia y las facultades de medicina. *Acad Psychiatry*. 2025;49:263-273. doi:10.1007/s40596-025-02136-4.
- Dorrestein L, Ritter C, De Mol Z, Wichtel M, Cary J, Vengrin C, et al. Evidencia de validez para la evaluación de habilidades de comunicación en la educación de profesionales de la salud: una revisión exploratoria. *BMJ Open*. 2025;15. doi:10.1136/bmjopen-2024-096799.
- Durán RFA, Sernaque FRC, Sánchez JLG, Sirlopu JRR, Moncada JJR. Ruta metodológica para la validación de instrumentos en investigación de ingeniería. *Rev Cient Pakamuros*. 2026. doi:10.37787/z2whae41.
- Manajemen J, Islam P, Nur MS, Fadli M, Kunci K, Konten V, et al. Validez del contenido de los instrumentos de becas KIP-K en la educación superior utilizando la escala de validez del coeficiente de Aiken. *Al-Tanzim J Manaj Pendidik Islam*. 2022. doi:10.33650/al-tanzim.v6i3.3680.
- Syafitri A, Anggraini D, Parinduri WM, Rambe TR, Rambe N. Análisis de validez de contenido en instrumentos de evaluación final semestral de ciencias integradas en educación primaria. *SEJPGSD*. 2023. doi:10.24114/sejpgsd.v13i4.54309.
- Anggraini D, Khumaedi M, Widowati T. Validez y confiabilidad de los instrumentos de evaluación de la independencia de los estudiantes de estética básica. *J Res Educ Eval*. 2020. doi:10.15294/jere.v9i1.42558.
- Fleitas Díaz M, Pérez Ortiz V, Zambrano Arias EY, Andrade Montesdeoca JI, Benítez Pardillo T. Validación a través del juicio de expertos: importancia y contribución al rigor investigativo actual en el área de la salud. *Rev Publicando*. 2025;12. doi:10.51528/rp.vol12.id2517.
- De Castro IG, Closas AH. El instrumento de observación ante el panel de expertos. *Rev Fac Cienc Econ*. 2023;30(1):61-76. doi:10.30972/rfce.3016691.
- Montero Zamora EG, Segura Cueva KA, Montero Cadena OG, Montero Cadena SB. Evaluación de la validez y fiabilidad en estudios científicos: revisión sistemática de métodos y buenas prácticas. *Cienc Educ*. 2024;365-387. doi:10.5281/zenodo.14207509.
- Jornet Meliá JM, Perales Montolío J, González Such J. El concepto de validez de los procesos de evaluación de la docencia. *Rev Esp Pedagog*. 2020;78(276):233-252.

- Herrera Guerra E, Bautista Arellanos R, Bonilla Ibañez P. Validez y confiabilidad de un instrumento para identificar factores que influyen en la adherencia al tratamiento en personas con factores de riesgo cardiovascular. *Salud UIS*. 2023;55. doi:10.18273/saluduis.55.e:23052.
- Benítez Pardillo T, Jiménez Espiñeira O, Molina EV, Ramírez Tinoco ME, Cobo Pozo KD. Desarrollo de competencias para la elaboración de fórmulas magistrales en estudiantes de Asistencia en Farmacia. *Rev Conecta Libertad*. 2022;6(3):77-87.
- Badrawi N, Hosny S, Ragab L, et al. Radical reform of the undergraduate medical education program in a developing country: the Egyptian experience. *BMC Med Educ*. 2023;23:143. doi:10.1186/s12909-023-04098-3.
- Al-Hashimi K, Said UN, Khan TN. Formative objective structured clinical examinations (OSCE) as an assessment tool in undergraduate medical education in the United Kingdom: a review of its utility. *Cureus*. 2023;15. doi:10.7759/cureus.38519.
- Höhne E, Recker F, Dietrich C, Schäfer V. Assessment methods in medical ultrasound education. *Front Med*. 2022;9. doi:10.3389/fmed.2022.871957.