Data and Metadata. 2025; 4:1103
doi: 10.56294/dm20251103

ORIGINAL

Methodological Proposal for the Design and Validation of Research Instruments Supported by Artificial Intelligence

Propuesta metodológica para el diseño y validación de instrumentos de investigación apoyados con Inteligencia Artificial

Richar Jacobo Posso Pacheco[1], Rosangela Caicedo-Quiroz[2], Giceya Maqueira-Caraballo[2], Julio Barzola-Monteses[2,3], Laura Cristina Barba Miranda[1], Jorge Ricardo Amancha Gabela[1]

[1]Ministerio de Educación. Quito, Ecuador.
[2]Universidad Bolivariana del Ecuador. Durán, Ecuador.
[3]Universidad de Guayaquil. Guayaquil, Ecuador.

Cite as: Posso Pacheco RJ, Caicedo-Quiroz R, Maqueira-Caraballo G, Barzola-Monteses J, Barba Miranda LC, Amancha Gabela JR. Methodological Proposal for the Design and Validation of Research Instruments Supported by Artificial Intelligence. Data and Metadata. 2025; 4:1103. https://doi.org/10.56294/dm20251103

Submitted: 15-10-2024 **Revised:** 22-03-2025 **Accepted:** 01-07-2025 **Published:** 02-07-2025

Editor: Dr. Adrián Alejandro Vitón Castillo

Corresponding Author: Richar Jacobo Posso Pacheco

ABSTRACT

Introduction: the validity of data collection instruments is essential to ensure the quality and replicability of scientific studies; traditional methods require time, resources, and expert participation, making validation difficult.
Objective: to develop a procedure for the design and validation of research instruments using Artificial Intelligence as a methodological support tool.
Method: an eight-phase model was designed, ranging from conceptual review and item formulation to linguistic evaluation, simulated rational validation, comprehension verification, internal consistency analysis, and structural optimization.
Results: the process demonstrated applicability, technical coherence, and practical utility. ChatGPT 4.5 enabled the automation of analyses and the generation of content aligned with theoretical constructs, optimizing the preliminary validation phases.
Conclusions: AI represents a viable alternative in resource-limited settings. While it does not replace classic empirical methods, it complements methodological rigor in key stages. Ethical and technical protocols must be established for its responsible use in scientific research.

Keywords: Validation; Instruments; Chatgpt; Research; Psychometrics.

RESUMEN

Introducción: la validez de los instrumentos de recolección de datos es esencial para garantizar la calidad y replicabilidad de los estudios científicos; los métodos tradicionales requieren tiempo, recursos y la participación de expertos, lo que dificulta su validación.
Objetivo: desarrollar un procedimiento de diseño y validación de instrumentos de investigación utilizando Inteligencia Artificial como herramienta de apoyo metodológico.
Método: se diseñó un modelo de ocho fases, desde la revisión conceptual y formulación de ítems, hasta la evaluación lingüística, validación racional simulada, verificación de comprensión, análisis de coherencia interna y optimización estructural.
Resultados: el proceso evidenció aplicabilidad, coherencia técnica y utilidad práctica. ChatGPT 4.5 permitió automatizar análisis y generar contenido alineado con constructos teóricos, optimizando las fases preliminares

de validación.
Conclusiones: la IA representa una alternativa viable en entornos con recursos limitados, no sustituye los métodos empíricos clásicos, complementa el rigor metodológico en etapas clave. Se debe establecer protocolos éticos y técnicos para su uso responsable en la investigación científica.

Palabras clave: Validación; Instrumentos; ChatGPT; Investigación; Psicometría.

INTRODUCTION

In scientific research, the validity of data collection instruments guarantees the quality of the results;[1] validity allows us to determine whether an instrument accurately measures the construct it is intended to measure;[2] reliability ensures the consistency of measurements over time.[3] When these conditions are not met, the interpretation of the data is erroneous, the conclusions fragile, and the replicability of the study fails.[4] This methodological concern is shared by psychology, education, and the social sciences, where instrumental rigor is considered an essential condition for obtaining scientific knowledge. This methodological concern is shared by psychology, education, and the social sciences, where instrumental rigor is considered an essential condition for obtaining scientific knowledge.[5]

The validation of instruments follows a well-structured process that involves operational definitions of variables, item design, expert review, pilot testing, statistical evaluation, correlation analysis, and internal consistency testing.[6,7,8] Coefficients such as Cronbach's alpha or the Kuder-Richardson 20 index are used to assess internal consistency.[9,10] At the same time, qualitative methods, including expert judgment and the Delphi method, complement the content and construct validity of the study.[11,12]

Validation is a rigorous process that guarantees the methodological quality of a research instrument; however, it involves high costs, a considerable investment of time, and the participation of highly specialized professionals.[13] This poses an obstacle for independent researchers, teachers without funding, or research teams with limited resources, making it difficult for them to implement technical procedures that ensure the scientific soundness of their study.

In this scenario, artificial intelligence (AI), such as ChatGPT, is a technology that can support the methodological processes of scientific research due to its ability to analyze semantic patterns, identify logical inconsistencies, reformulate items, detect ambiguities, and generate answers based on defined theoretical constructs.[14] Currently, its use has been extended to educational, evaluative, and editorial tasks, demonstrating its effectiveness in academic writing, thematic analysis, and instrument design.[15,16] Although, paradoxically, empirical studies are still very scarce, their applications are being explored.[17,18]

Despite these advances, the application of AI in the validation of research instruments has not been widely published. There are no standardized protocols that scientifically validate the results generated by artificial intelligence, and its use in academic production is questioned due to the lack of depth and control over the sources consulted.[19] This lack of knowledge creates a knowledge gap that limits the widespread use of these technologies in the psychometric assurance of research instruments.

From a scientific and methodological perspective, incorporating AI into validation processes represents an innovative solution. Especially in contexts where access to expert panels is not available, where validation timelines are tight, or where rapid adjustments are required, ChatGPT offers an agile, automatable, and replicable alternative, provided it is used under well-defined ethical and methodological guidelines.[20] Moreover, its ability to analyze text from multiple perspectives allows addressing aspects such as the level of readability or appropriateness to the level of the informant, thus extending the scope of traditional validation.

Against this background, this article aimed to develop a validation procedure for a research instrument using ChatGPT as a methodological support tool. It is proposed to verify the semantic coherence of the items through automated linguistic analysis of the model and to contrast the results with assessments obtained through expert judgment, thus determining the validity, accuracy, and usefulness of using artificial intelligence in psychometric assurance processes.

METHOD

This study was based on instrumental research and methodological development, which did not seek to apply the procedure empirically but rather to design and methodologically validate it through conceptual frameworks supported by classical psychometrics and automated language processing. The proposed validation model addresses the growing need for efficient, replicable, and well-founded tools to enhance methodological rigor in educational studies.

The development of the methodology was carried out in eight phases between January and April 2025; it began with a documentary review of the state of the art in classical psychometrics, the design of instruments

and applications of natural language processing in research; this action was fundamental to establish the conceptual framework that underpinned each of the proposed phases.

For this purpose, an iterative process was implemented to construct the phases, which included defining their purposes and technical foundations, as well as elaborating operational prompts for interaction with ChatGPT 4.5. Internal simulations and continuous adjustments were also carried out, testing the logical coherence and feasibility of each step with various item structures and constructs, thus contributing to the preliminary validation of educational instruments from an innovative approach.[21,22] This allowed critical aspects of the ChatGPT 4.5 process to be addressed.

This made it possible to address critical aspects of the validation process, such as identifying lexical ambiguities, detecting deficient formulations, and verifying the theoretical alignment of the items with the construct's dimensions.[23,24] The consistency of the results obtained through the replicability of the process, combined with the technical relevance of the responses generated by the AI, conclusively supports the applicability and reliability of the procedure developed.

To ensure the argumentative, structural, and technical validity of the proposed procedure, a complementary review process was implemented by a panel of eight experts. These specialists, with an established track record in scale validation, instructional design, and scientific consultancy, conducted an exhaustive review of each phase of the procedure, the technical criteria underlying it, and the prompts designed for interaction with ChatGPT 4.5.

The comments and feedback provided by this expert panel were crucial in strengthening the terminological consistency of the procedure, clarifying the scope and objectives of each stage, and adjusting the operational instructions to ensure their systematic and accurate application. The evaluation was conducted through successive rounds of feedback, which facilitated discussion until a unanimous technical consensus was reached. This expert validation process ensured that the proposed procedure complies with the epistemological principles of documentary validation and can be effectively replicated in diverse educational contexts, always under acceptable ethical and methodological standards.

RESULTS

As a result of this construction and validation process, the 'Artificial Intelligence Assisted Research Instrument Validation Procedure' is presented, a methodological sequence of eight phases designed to optimize and strengthen the rigor in the construction and evaluation of measurement tools.

Phase 1: Conceptual review and keyword extraction

The validation process begins with the rigorous definition of the construct to be measured. This phase is based on a systematic review of the state of the art to delimit the theoretical dimensions that comprise it. Based on this conceptual analysis, we proceed to identify the fundamental semantic units that allow us to operationalize the construct. With the support of ChatGPT 4.5, an automated mechanism is integrated for the assisted extraction of keywords from academic definitions from specialized sources.

This action strengthens the alignment between the theoretical domains and the future items of the instrument, ensuring that each dimension is represented in a comprehensive manner. The language model is instructed to interpret the disciplinary content and generate an organized list of key terms, accompanied by their logical justification concerning the construct.

Applied technical prompt: "From this academic definition of [construct], identify at least 10 keywords representing its main dimensions. Justify each word with its conceptual link to the main construct".

This phase responds directly to the principle of content validity by ensuring a representative semantic basis for the subsequent construction of the instrument.

Phase 2: Initial item formulation

With the keywords obtained, we proceed to the initial formulation of the items. This stage aims to translate the theoretical dimensions of the construct into measurable statements structured according to the principles of psychometric writing: syntactic clarity, semantic uniqueness, conceptual relevance, and appropriateness to the level of the informant. ChatGPT 4.5 is used here as a generative tool aimed at producing items that faithfully reflect each dimension of the construct in language that is accessible and free of unnecessary technicalities.

The model enables the production of differentiated, non-redundant statements that meet the criteria of face validity. Each item is formulated in terms of its theoretical dimension, ensuring logical coherence and conceptual coverage.

Technical Prompt applied: "Write three items measuring dimension [X] of the construct [Y], using clear, direct, and non-technical language. The items should be representative of the dimension and adapted for secondary school students".

This phase helps ensure the functional adequacy of the instrument, thereby strengthening its content

validity from an operational perspective.

Phase 3: Automated linguistic and semantic evaluation

Once the preliminary items have been generated, a technical evaluation of the language is conducted, focusing on grammatical clarity, semantic coherence, readability, and expressive neutrality. This phase addresses the need to minimize ambiguities, formulation errors, or interpretative biases, which could compromise the quality of the instrument. ChatGPT 4.5 is configured as an expert linguistic evaluator, capable of identifying lexical inconsistencies and proposing reformulations while preserving the original intention of the item.

This automated analysis operates on each utterance individually, optimizing reading comprehension and improving the communicative quality of the instrument without altering its theoretical validity.

Technical Prompt applied: "Evaluate this item in terms of grammatical clarity, semantic ambiguity, readability, and neutrality. Indicates any errors or problems and provides a corrected version while maintaining the original meaning".

The result is a refined version of the instrument that respects the principle of semantic validity, the key to ensuring consistent interpretations by informants.

Phase 4: Rational content validation (simulated)

The fourth phase of the procedure corresponds to rational validation, understood as the assessment of the degree of correspondence between each item and the theoretical dimension it is intended to represent. This judgment, traditionally made by experts in the field, is simulated using precise instructions to ChatGPT 4.5, which operates as a conceptual evaluator based on the criteria defined by the researcher.

The model analyzes whether the item content reflects a relevant, specific, and non-redundant way of assigning to a theoretical dimension, providing a reasoned justification for its acceptance, modification, or exclusion.

Technical Prompt applied: "Assess whether this item adequately represents dimension [X] of the construct [Y]. Justify your assessment by pointing out which aspect of the dimension is covered or what is lacking in the item".

This phase responds to the principle of content validity from a simulated approach, allowing the instrument to be refined before its empirical validation or traditional expert judgment.

Phase 5: Verification of understanding from the informant's perspective

An essential condition for the validity of an instrument is that its items are understandable to the target population. This phase seeks to anticipate the level of comprehension of the statements, simulating the interpretation that a typical respondent might make. ChatGPT 4.5 is incorporated as a cognitive simulation agent, trained to adopt the linguistic and cognitive profile of the target population.

The aim is to identify ambiguous terms, complex structures, or formulations that could lead to confusion or misinterpretation, particularly in populations with varying levels of education. This simulation allows proactive adjustments to be made before pilot testing is applied.

Technical Prompt applied: "Assume the role of a 13-year-old student reading this item. Describe in your own words what you understand. Indicate if there are any words, phrases, or concepts that are confusing".

This phase contributes to the face and application validity of the instrument, ensuring its accessibility and pragmatic appropriateness to the characteristics of the informants.

Phase 6: Assessment of conceptual internal consistency

The internal consistency of an instrument depends not only on its statistical consistency but also on the semantic and structural logic between items. This phase enables the identification of redundancies, overlaps, or thematic disconnections between statements, facilitating a theoretical reorganization of the instrument before any empirical analysis. ChatGPT 4.5 acts here as a semantic analyst, able to compare statements, group them according to conceptual similarity, and identify inconsistencies or redundancies.

The model processes the comprehensive content of the instrument, establishing relationships between items and suggesting a more coherent structure.

Technical Prompt applied: "Group the following items according to their thematic similarity. Indicate if any of them do not correspond to the groups formed and justify why".

This phase strengthens the conceptual internal validity, as it enables the anticipation of semantic deviations or coverage errors that could impact the overall interpretation of the results.

Phase 7: Assessment of the construct's theoretical structure

Structural validation aims to verify that the formulated items accurately represent the previously defined theoretical dimensions. At this stage, ChatGPT 4.5 is instructed to perform a supervised ranking of items,

assigning each item to the dimension it best represents and justifying this decision. This procedure enables the detection of underrepresented dimensions and the identification of misaligned items, ensuring correspondence between the construct structure and the instrument's organization.

The model applies principles of thematic classification, semantic inference, and logical content analysis to support this task systematically.

Technical Prompt applied: "The following items were designed to assess three dimensions of the construct [Y]: A, B, and C. Assign each item to a dimension, explain your decision, and point out if any of the items do not fit any of them."

Anticipatory structural validation enables the consolidation of the coherence of the underlying theoretical model, thereby strengthening construct validity at an early stage.

Phase 8: Evaluation of conceptual performance and optimization of the instrument

The final phase aims to optimize the instrument's structure by evaluating the conceptual performance of each item and eliminating redundant or uninformative items. ChatGPT 4.5 is implemented as a debugging assistant, applying criteria of semantic economy, thematic relevance, and conceptual density to propose a more efficient version of the instrument without compromising the coverage of the theoretical dimensions.

The model analyzes the items as a whole, determines their degree of conceptual differentiation, and suggests deletions or mergers based on the criteria of thematic saturation.

Technical Prompt applied: "Analyse this list of items and determine which items present less conceptual value or repeat information already covered. Propose a version of the instrument with fewer items but maintaining full coverage of the construct".

This phase enables the development of shorter, more manageable, and robust instruments, optimizing their applicability in real educational contexts while respecting the principles of documented validity and reliability.

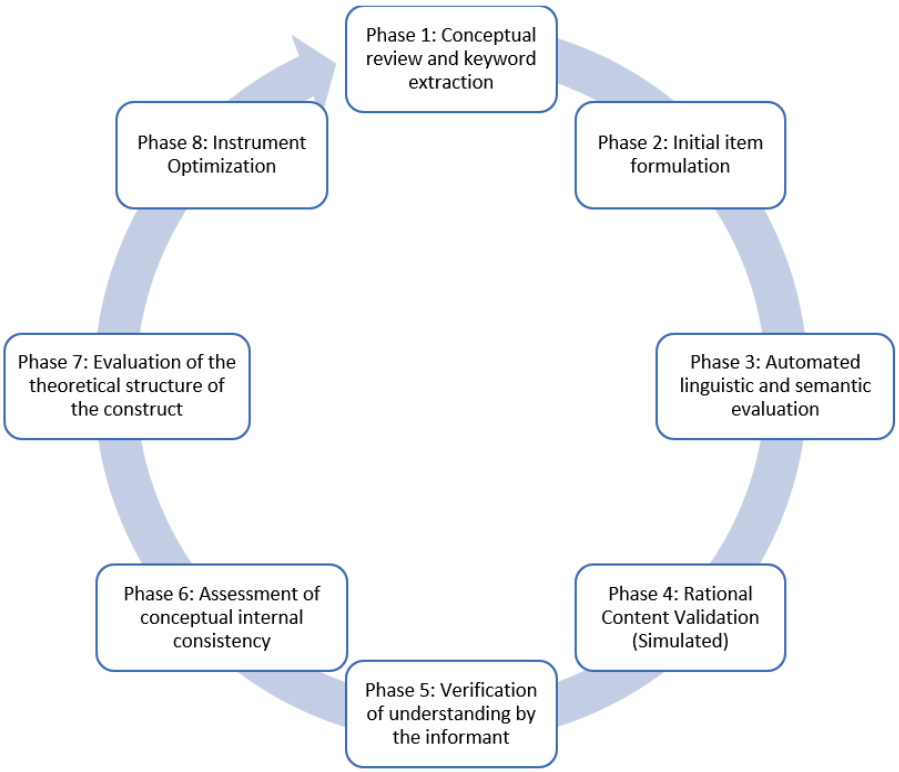

Figure 1. Artificial Intelligence-Assisted Validation Model for Research Instruments

DISCUSSION

Strengths of the AI-Assisted Approach

The AI-assisted instrument validation procedure, as proposed, offers several strengths that address the limitations of traditional methods. ChatGPT 4.5's ability to perform linguistic analyses, comprehension simulations, and conceptual assessments in an automated manner drastically reduces the time and effort required in the preliminary validation phases. This is valuable in contexts where timelines are tight or where quick adjustments to the instrument are needed.[25] The replicability of the process is remarkable.

The replicability of the process is remarkable because, by relying on standardized technical prompts and a consistent AI model, the procedure can be replicated by different researchers, thereby contributing to the transparency and verifiability of the results. This feature is fundamental to scientific rigor, as it allows other researchers to follow the same steps and obtain comparable results.[26]

Accessibility is a crucial strength for independent researchers; without funding or teams with limited resources, access to expert panels or large-scale pilot testing may be prohibitively expensive.[13] AI offers a viable, low-cost alternative for robust preliminary validation, democratizing access to advanced methodological tools and strengthening the scientific robustness of studies that might otherwise be lacking.[27]

Comparison with Traditional Validation Methods

The AI-assisted instrument validation procedure, specifically with ChatGPT 4.5, has distinct advantages and disadvantages compared to traditional validation methods such as expert judgment, pilot testing, and empirical statistical analysis.

In the field of qualitative and content validation, simulated expert judgment (Phase 4) and comprehension verification from the informant's perspective (Phase 5), utilizing AI, offer efficiency and accessibility. Traditional methods of expert judgment and pilot testing require an investment of time and resources to recruit specialists and participants, coordinate sessions, and analyze feedback. AI, in contrast, enables rapid and automated evaluation, reducing costs and timelines, which is particularly beneficial for independent researchers or teams with limited resources. AI simulation lacks the depth of contextual reasoning, intuition, and ability to identify complex cultural or linguistic nuances that a human expert or real participant can provide. Human interaction in expert judgment and pilot testing allows for richer qualitative feedback and the identification of problems not anticipated by an AI model.[28]

In quantitative and structural validation (Phases 6, 7, 8), AI acts as an assistant in the pre-statistical stages, facilitating the identification of redundancies, thematic clustering, and optimization of the instrument structure. This can streamline the initial debugging process before large-scale data collection. AI complements, but does not replace, traditional empirical statistical analyses, such as Exploratory Factor Analysis (EFA) or Confirmatory Factor Analysis (CFA), for assessing construct validity and calculating reliability coefficients, including Cronbach's Alpha or KR-20. IA may suggest a theoretical structure or identify conceptual inconsistencies, but empirical confirmation of these properties requires the application of the instrument and statistical analysis of the resulting data.[25,27]

CONCLUSIONS

The AI-assisted approach demonstrates efficiency, replicability, and accessibility; these strengths make it an alternative for independent researchers or institutions with logistical constraints. AI does not replace the methodological richness of traditional methods, such as human expert judgment or statistical analyses of real data. Instead, it is positioned as a technical complement that optimizes the preliminary phases of the validation process, reducing time and improving the initial quality of the items.

The use of AI in validation processes faces limitations that must be addressed with scientific responsibility, such as the absence of standardized protocols to validate the results generated by models such as ChatGPT and the opacity regarding the sources used. Ethical risks must be considered, including algorithmic biases and the uncritical automation of methodological decisions. It is, therefore, essential that the researcher retains epistemological control and critically evaluates the suggestions made by the AI, rigorously documenting the entire process.

The procedure developed opens up new routes for methodological innovation in the social and educational sciences. In the future, it should focus on applying the validated instruments to empirical samples to explore their construct validity and reliability using classical statistical techniques, such as AFE, AFC, or Cronbach's Alpha. It should also empirically compare the efficacy of this procedure against traditional methods and promote the development of ethical and technical guidelines that standardize the use of AI in educational research.

BIBLIOGRAPHICAL REFERENCES

1. Guillermo Cornetero MC, López-Regalado O. Validez y confiabilidad de instrumentos de investigación en el aprendizaje: una revisión sistemática. Rev Trib [Internet]. 22 de enero de 2025 [citado 14 de mayo de 2025];5(10):653-75. Disponible en: https://revistatribunal.org/index.php/tribunal/article/view/370

2. Contreras Cázarez CR, Campa Álvarez RDLÁ. Diseño instrumental y validación de un cuestionario para la competencia informacional en estudiantes universitarios. Sinéctica [Internet]. 14 de noviembre de 2022 [citado 14 de mayo de 2025];(59). Disponible en: https://sinectica.iteso.mx/index.php/SINECTICA/article/view/1428

3. González Arroyo AP, Bermeo Escalona JR. Metodología para la construcción y validación de instrumentos

en el área de la salud. IBN SINA [Internet]. 5 de abril de 2024 [citado 14 de mayo de 2025];Vol. 15 Núm. 1:84-94 Páginas. Disponible en: https://revistas.uaz.edu.mx/index.php/ibnsina/article/view/2541

4. Hair J, Black W, Babin B, Anderson R. MULTIVARIATE DATA ANALYSIS [Internet]. EIGHTH EDITION; 2019. Disponible en: https://eli.johogo.com/Class/CCU/SEM/_Multivariate%20Data%20Analysis_Hair.pdf

5. Posso Pacheco RJ. Diseño metodológico de sistematización de preguntas abiertas: un esfuerzo para mejorar la investigación cualitativa. MENTOR Rev Investig Educ Deport [Internet]. 28 de septiembre de 2023 [citado 14 de abril de 2024];2(6):919-25. Disponible en: https://revistamentor.ec/index.php/mentor/article/view/6780

6. Lopez Domínguez SI, Sotelo Asef JG, Vargas Hernández M. Instrumento para evaluar la percepción sobre la infraestructura física y virtual en las organizaciones: diseño y validación. RIDE Rev Iberoam Para Investig El Desarro Educ [Internet]. 19 de febrero de 2024 [citado 14 de mayo de 2025];14(28). Disponible en: http://www.ride.org.mx/index.php/RIDE/article/view/1814

7. Harnol Alexander CC, Jehimy Johana SV. Metodología para comprender la interacción entre las organizaciones y el sistema regional de innovación en el desarrollo de capacidades dinámicas. Conoc Glob [Internet]. 10 de junio de 2024 [citado 14 de mayo de 2025];9(2):432-50. Disponible en: https://conocimientoglobal.org/revista/index.php/cglobal/article/view/475

8. Posso Pacheco RJ, Bertheau EL. Validez y confiabilidad del instrumento determinante humano en la implementación del currículo de educación física. Rev Educ - UPEL-IPB - Segunda Nueva Etapa 20 [Internet]. 19 de diciembre de 2020 [citado 25 de noviembre de 2024];24(3):205-23. Disponible en: https://revistas.investigacion-upelipb.com/index.php/educare/article/view/1410

9. Durán Pérez FB, Lara Abad GE. Aplicación del coeficiente de confiabilidad de Kuder Richardson en una escala para la revisión y prevención de los efectos de las rutinas. Bol Científico Esc Super Atotonilco Tula [Internet]. 5 de enero de 2021 [citado 14 de mayo de 2025];8(15):51-5. Disponible en: https://repository.uaeh.edu.mx/revistas/index.php/atotonilco/article/view/6693

10. A Uyanah D, U. I N. The Theoretical and Empirical Equivalence of Cronbach Alpha and Kuder-Richardson Formular-20 Reliability Coefficients. Int Res J Innov Eng Technol [Internet]. 2023 [citado 14 de mayo de 2025];07(05):17-23. Disponible en: https://irjiet.com/common_src/article_file/1683708204_3625019ec7_7_irjiet.pdf

11. 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 Innovació Recer En Educ [Internet]. 29 de junio de 2024 [citado 14 de mayo de 2025];17(2). Disponible en: https://revistes.ub.edu/index.php/REIRE/article/view/46238

12. Ramírez Chávez MA, Ramírez Torres TZ. El método DELPHI como herramienta de investigación. Una revisión: The DELPHI method as a research tool. A review. LATAM Rev Latinoam Cienc Soc Humanidades [Internet]. 11 de marzo de 2024 [citado 14 de mayo de 2025];5(1). Disponible en: https://latam.redilat.org/index.php/lt/article/view/1842

13. Meneses Ayllón B, Loza Rodríguez JC, Durán Gutiérrez LF. Escala de percepción de habilidades investigativas en estudiantes universitarios: desarrollo y validación. Rev Divers Científica [Internet]. 16 de agosto de 2024 [citado 14 de mayo de 2025];4(2):13-22. Disponible en: https://revistadiversidad.com/index.php/revista/article/view/125

14. Lund BD, Wang T. Chatting about ChatGPT: how may AI and GPT impact academia and libraries? Libr Hi Tech News [Internet]. 16 de mayo de 2023 [citado 15 de mayo de 2025];40(3):26-9. Disponible en: https://www.emerald.com/insight/content/doi/10.1108/LHTN-01-2023-0009/full/html

15. Eden CA, Chisom ON, Adeniyi IS. Integrating AI in education: Opportunities, challenges, and ethical considerations. Magna Sci Adv Res Rev [Internet]. 30 de marzo de 2024 [citado 15 de mayo de 2025];10(2):006-13. Disponible en: https://magnascientiapub.com/journals/msarr/content/integrating-ai-education-opportunities-challenges-and-ethical-considerations

16. Matthay EC, Neill DB, Titus AR, Desai S, Troxel AB, Cerdá M, et al. Integrating Artificial Intelligence into Causal Research in Epidemiology. Curr Epidemiol Rep [Internet]. 24 de marzo de 2025 [citado 15 de mayo de 2025];12(1):6. Disponible en: https://link.springer.com/10.1007/s40471-025-00359-5

17. Dos Santos R. Enhancing Physics Learning with ChatGPT, Bing Chat, and Bard as Agents-to-Think-With: A Comparative Case Study. SSRN Electron J [Internet]. 2023 [citado 15 de mayo de 2025]; Disponible en: https://www.ssrn.com/abstract=4478305

18. Duong T, Suppasetseree S. The Effects of an Artificial Intelligence Voice Chatbot on Improving Vietnamese Undergraduate Students' English Speaking Skills. Int J Learn Teach Educ Res [Internet]. 30 de marzo de 2024 [citado 29 de marzo de 2025];23(3):293-321. Disponible en: https://ijlter.org/index.php/ijlter/article/view/9787/pdf

19. Sánchez-Bolívar L, Escalante-González S, Martínez-Martínez A. La ética de la inteligencia artificial en educación: ¿Amenaza u oportunidad? Rev Electrónica Educ [Internet]. 21 de diciembre de 2024 [citado 18 de mayo de 2025];28(S):1-20. Disponible en: https://www.revistas.una.ac.cr/index.php/EDUCARE/article/view/20541

20. European Commission. Directorate General for Education, Youth, Sport and Culture. Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators. [Internet]. LU: Publications Office; 2022 [citado 18 de mayo de 2025]. Disponible en: https://data.europa.eu/doi/10.2766/153756

21. Yilmaz FGK, Yilmaz R, Ceylan M. Generative Artificial Intelligence Acceptance Scale: A Validity and Reliability Study. Int J Human–Computer Interact [Internet]. 16 de diciembre de 2024 [citado 18 de mayo de 2025];40(24):8703-15. Disponible en: https://www.tandfonline.com/doi/full/10.1080/10447318.2023.2288730

22. Lund BD, Wang T, Mannuru NR, Nie B, Shimray S, Wang Z. CHATGPT and a new academic reality: ARTIFICIAL INTELLIGENCE-WRITTEN research papers and the ethics of the large language models in scholarly publishing. J Assoc Inf Sci Technol [Internet]. mayo de 2023 [citado 18 de mayo de 2025];74(5):570-81. Disponible en: https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.24750

23. Hernández R, Fernández C, Baptista P. Metodología de la investigación. McGraw-Hill; 2014.

24. Ato M, López-García JJ, Benavente A. Un sistema de clasificación de los diseños de investigación en psicología. An Psicol [Internet]. 1 de octubre de 2013 [citado 18 de mayo de 2025];29(3):1038-59. Disponible en: http://revistas.um.es/analesps/article/view/178511

25. Ge L, Agrawal R, Singer M, Kannapiran P, De Castro Molina JA, Teow KL, et al. Leveraging artificial intelligence to enhance systematic reviews in health research: advanced tools and challenges. Syst Rev [Internet]. 25 de octubre de 2024 [citado 1 de junio de 2025];13(1):269. Disponible en: https://systematicreviewsjournal.biomedcentral.com/articles/10.1186/s13643-024-02682-2

26. Desai A, Abdelhamid M, Padalkar NR. What is Reproducibility in Artificial Intelligence and Machine Learning Research? [Internet]. arXiv; 2024 [citado 1 de junio de 2025]. Disponible en: https://arxiv.org/abs/2407.10239

27. Fan G, Liu D, Zhang R, Pan L. The impact of AI-assisted pair programming on student motivation, programming anxiety, collaborative learning, and programming performance: a comparative study with traditional pair programming and individual approaches. Int J STEM Educ [Internet]. 4 de marzo de 2025 [citado 1 de junio de 2025];12(1):16. Disponible en: https://stemeducationjournal.springeropen.com/articles/10.1186/s40594-025-00537-3

28. Lund BD, Wang T, Mannuru NR, Nie B, Shimray S, Wang Z. CHATGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. J Assoc Inf Sci Technol [Internet]. mayo de 2023 [citado 1 de junio de 2025];74(5):570-81. Disponible en: https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.24750

FUNDING

The authors received no funding for this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Richar Jacobo Posso.
Data curation: Rosangela Caicedo.
Formal analysis: Rosangela Caicedo.
Research: Richar Jacobo Posso, Jorge Ricardo Amancha.
Methodology: Rosangela Caicedo.
Project management: Giceya Maqueira.
Resources: Giceya Maqueira, Jorge Ricardo Amancha.
Software: Richar Jacobo Posso, Jorge Ricardo Amancha.
Supervision: Richar Jacobo Posso, Laura Cristina Barba.
Validation: Julio Barzola.
Visualisation: Julio Barzola, Laura Cristina Barba.
Writing - original draft: Richar Jacobo Posso.
Writing - revision and editing: Julio Barzola, Laura Cristina Barba.

https://doi.org/10.56294/dm20251103 ISSN: 2953-4917