

PSYCHOMETRIC VALIDATION OF A DIAGNOSTIC TRIGONOMETRY TEST APPLIED TO UNIVERSITY STUDENTS

VALIDAÇÃO PSICOMÉTRICA DE UM EXAME DIAGNÓSTICO DE TRIGONOMETRIA APLICADO A ESTUDANTES UNIVERSITÁRIOS

VALIDACIÓN PSICOMÉTRICA DE UN EXAMEN DIAGNÓSTICO DE TRIGONOMETRÍA APLICADO A ESTUDIANTES UNIVERSITARIOS



<https://doi.org/10.56238/sevened2026.022-012>

Dalia Imelda Castillo Márquez¹, Richard Gutiérrez Cuesta², Georgina Elizabeth Partida López³, Alan Roberto Figueroa Flores⁴, Arturo Javier Gómez Dávalos⁵, Juan Luis Hernández Méndez⁶

ABSTRACT

The present study aimed to analyze the psychometric properties of a diagnostic trigonometry test designed for university students. The study design was quantitative, non-experimental, cross-sectional, and instrumental. A total of 50 higher education students participated, selected through non-probabilistic purposive sampling. The instrument consisted of 10 dichotomous items (1 = correct, 0 = incorrect) assessing trigonometric functions, identities, and triangle solving. The difficulty index (p), corrected point-biserial correlation ($r_{p\beta}$), and Cronbach's alpha coefficient (α) were calculated. The results indicated acceptable reliability ($\alpha = 0.748$). Item analysis showed that most items had moderate difficulty and adequate discrimination; however, Item 8 showed virtually no discrimination ($r_{p\beta} = 0.001$), while Items 3 and 7 presented low discrimination. It is concluded that the instrument has satisfactory psychometric properties as a diagnostic tool, although it requires revision of three specific items to improve its technical quality. The findings provide useful empirical evidence for decision-making in introductory university-level trigonometry instruction.

Keywords: Psychometric Validation. Diagnostic Test. Trigonometry. Higher Education. Item Analysis.

¹ Dr. Universidad Autónoma de Nayarit. México. E-mail: dalia.castillo@uan.edu.mx

Orcid: <https://orcid.org/0000-0002-5890-0437>

² Master's degree. Universidad Hispanoamericana. Costa Rica. E-mail: richard.gutierrez0647@uhispano.ac.cr

Orcid: <https://orcid.org/0009-0005-2750-3702>

³ Master's degree. Universidad Autónoma de Nayarit. México. E-mail: gina.partida@uan.edu.mx

Orcid: <https://orcid.org/0000-0001-5183-0846>

⁴ Master's degree. Universidad Autónoma de Nayarit. México. E-mail: alan.figueroa@uan.edu.mx

Orcid: <https://orcid.org/0009-0008-1972-6203>

⁵ Dr. Universidad Autónoma de Nayarit. México. E-mail: agomez7001@uan.edu.mx

Orcid: <https://orcid.org/0000-0002-4523-5423>

⁶ Master's degree. Universidad Autónoma de Nayarit. México. E-mail: juan.l@uan.edu.mx

Orcid: <https://orcid.org/0000-0002-5686-2887>

RESUMO

O presente estudo teve como objetivo analisar as propriedades psicométricas de um exame diagnóstico de trigonometria elaborado para estudantes universitários. O desenho do estudo foi quantitativo, não experimental, transversal e de natureza instrumental. Participaram 50 estudantes do ensino superior selecionados por amostragem não probabilística intencional. O instrumento foi composto por 10 itens dicotômicos (1 = correto, 0 = incorreto) que avaliaram conteúdos de funções trigonométricas, identidades e resolução de triângulos. Foram calculados o índice de dificuldade (p), a correlação ponto-biserial corrigida ($r_{p\beta}$) e o coeficiente alfa de Cronbach (α). Os resultados indicaram uma confiabilidade aceitável ($\alpha = 0,748$). A análise dos itens revelou que a maioria apresentou dificuldade moderada e discriminação adequada; entretanto, o Item 8 apresentou discriminação praticamente nula ($r_{p\beta} = 0,001$), enquanto os Itens 3 e 7 apresentaram baixa discriminação. Conclui-se que o instrumento possui propriedades psicométricas satisfatórias como ferramenta diagnóstica, embora necessite da revisão de três itens específicos para aprimorar sua qualidade técnica. Os achados fornecem evidências empíricas úteis para a tomada de decisão no ensino inicial de trigonometria no ensino superior.

Palavras-chave: Validação Psicométrica. Exame Diagnóstico. Trigonometria. Ensino Superior. Análise de Itens.

RESUMEN

El presente estudio tuvo como objetivo analizar las propiedades psicométricas de un examen diagnóstico de trigonometría diseñado para estudiantes universitarios. El diseño fue cuantitativo, no experimental, transversal y de tipo instrumental. Participaron 50 estudiantes de nivel superior seleccionados mediante muestreo no probabilístico intencional. El instrumento estuvo compuesto por 10 ítems dicotómicos (1 = correcto, 0 = incorrecto) que evaluaron contenidos de funciones trigonométricas, identidades y resolución de triángulos. Se calcularon el índice de dificultad (p), la correlación punto-biserial corregida (r_{p^b}) y el coeficiente alfa de Cronbach (α). Los resultados indicaron una confiabilidad aceptable ($\alpha = 0.748$). El análisis de ítems reveló que la mayoría presentó dificultad moderada y adecuada discriminación; sin embargo, el Ítem 8 mostró discriminación prácticamente nula ($r_{p^b} = 0.001$), mientras que los Ítems 3 y 7 presentaron discriminación baja. Se concluye que el instrumento posee propiedades psicométricas satisfactorias como herramienta diagnóstica, aunque requiere revisión de tres ítems específicos para mejorar su calidad técnica. Los hallazgos ofrecen evidencia empírica útil para la toma de decisiones en la enseñanza inicial de la trigonometría universitaria.

Palabras clave: Validación Psicométrica. Examen Diagnóstico. Trigonometría. Educación Universitaria. Análisis de Ítems.

1 INTRODUCTION

Diagnostic assessment is a fundamental practice in higher education, particularly in disciplines with highly structured and sequential content such as mathematics. In the university environment, the prior knowledge with which students enter exact science courses determines, to a large extent, their academic trajectory and their probability of success (Ausubel et al., 1983). Trigonometry, as an area of mathematical knowledge, represents one of the indispensable conceptual bridges between high school education and university-level studies in engineering, architecture, physics, and computer science, among others.

The difficulty that university students face in learning trigonometry has been widely documented in the specialized literature. Research carried out in Latin American contexts indicates that concepts such as sine, cosine and tangent functions, trigonometric identities and the resolution of triangles represent significant cognitive obstacles for a considerable proportion of new students (Buendía & Montiel, 2011; Gamboa Araya & Ballesterero Alfaro, 2010). These findings reveal that deficiencies in trigonometry at university entrance represent a widespread problem that requires systematic and rigorous diagnosis. These obstacles are not limited to initial academic performance, but also have a negative impact on the achievement of mathematical competencies essential for subsequent professional performance, in such a way that they are projected on performance in subsequent courses in calculus, analytical geometry, and physics (Orhun, 2001).

In this scenario, the design and validation of accurate and technically sound diagnostic assessment instruments acquire strategic relevance. A quality diagnostic instrument allows teachers to accurately identify their students' conceptual gaps, make informed pedagogical decisions, and design focused didactic interventions (Ravela et al., 2017). In this sense, Castillo Márquez et al. (2025a) highlight that the application of diagnostic tests at the beginning of the school year is an essential practice to make knowledge gaps in mathematics visible. Along the same lines, Castillo Márquez et al. (2025b) documented that new students entering the bachelor's degree in mathematics presented diagnostic averages of only 4.2 out of 10 in arithmetic and basic algebra, which shows that mathematical deficiencies at university admission constitute a structural problem that demands valid and reliable diagnostic instruments as a starting point for pedagogical intervention.

However, for a diagnostic test to reliably fulfill this function, it must undergo a rigorous validation process that includes the analysis of its psychometric properties.

Classical psychometrics provides robust tools to evaluate the technical quality of educational instruments. Among the most commonly used indicators in the analysis of

dichotomous response exams are the difficulty index (p), which expresses the proportion of examinees who correctly answer an item, and the discrimination index, usually operationalized by the corrected point-biserial correlation (r_{pb}), which reflects the ability of the item to differentiate between students with greater and lesser mastery of the construct evaluated (Martínez Arias et al., 2006; Muñiz, 2010). Likewise, the reliability of the instrument, understood as the consistency of the scores across the items, is usually estimated using Cronbach's alpha coefficient (Cronbach, 1951).

Cronbach's alpha coefficient is the most widely used internal consistency statistic in educational and psychological research (Frías-Navarro & Pascual Soler, 2012). According to the criteria proposed by George and Mallery (2003), values of $\alpha \geq 0.70$ are considered acceptable for diagnostic and research instruments. However, it is important to contextualize this parameter according to the purpose of the instrument: in high-impact evaluations, such as professional certification exams, values above 0.90 are required, while in training and diagnostic instruments, more moderate thresholds are technically justifiable (Morales Vallejo, 2007).

Item analysis, on the other hand, allows the identification of items that do not contribute to the measurement of the construct of interest or that have deficient technical characteristics. An item with r_{pb} close to zero indicates that there is no relationship between the response to that item and the total score of the test, which suggests that the item measures something different from the rest of the instrument or that it presents problems of ambiguity in its wording (Martínez Arias et al., 2006). Similarly, items with an extreme difficulty index – very easy or difficult – offer little diagnostic information and reduce the variance of the scores, negatively affecting the internal consistency of the instrument (Muñiz, 2010).

Despite the recognized importance of psychometric validation, in university teaching practice it is common for diagnostic tests to be developed and applied without having been subjected to this process, which compromises their usefulness and the interpretation of their results. This gap between everyday evaluation practice and recommended technical standards is a relevant methodological problem, whose attention contributes to improving the quality of the data that guide pedagogical decisions (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014).

In this context, the present study aimed to analyze the psychometric properties of a trigonometry diagnostic test designed for university students, by calculating the difficulty index, the discrimination index and Cronbach's alpha reliability coefficient. The findings reported seek

to provide empirical evidence that guides decision-making on the use, revision or improvement of the instrument in university mathematics teaching contexts.

2 METHOD

2.1 RESEARCH DESIGN

The study followed a quantitative approach, with a non-experimental design, cross-sectional scope and instrumental type, in accordance with the classification proposed by Montero and León (2007) for studies aimed at the development and validation of measurement instruments in psychology and education. This type of design is appropriate when the central purpose of the study is to analyze the psychometric properties of a test or scale, without manipulation of variables or random assignment of participants.

2.2 PARTICIPANTS

The sample was made up of 50 first-semester students in the area of Basic Sciences and Engineering. The inclusion criteria were as follows: to be officially enrolled in the institution during the school year in which the study was carried out, to study for the first time the level or subject of trigonometry or related area, and to express voluntary availability to participate. Students with previous trigonometry courses at the higher level were excluded, as well as those who did not complete the instrument in at least 80% of its items. The sampling was non-probabilistic of an intentional type, since accessible academic groups that met the established criteria were selected.

2.3 INSTRUMENT

A diagnostic trigonometry test was applied consisting of 10 dichotomous items, with answer options coded as 1 (correct) and 0 (incorrect). The items evaluated the following contents: angles and their measurements in degrees and radians, trigonometric functions sine, cosine and tangent, fundamental trigonometric identities, and resolution of right and oblique triangles. The total score of the instrument ranges from 0 to 10 points, with one point corresponding to each correct answer. The format of the items was multiple choice, designed to be answered in approximately 10 to 15 minutes.

2.4 PROCEDURE

The instrument was applied in person during the first semester, August-September 2025. Prior to the administration of the review, participants were informed about the objectives of the

study and the voluntary and confidential nature of their participation, ensuring the anonymity of the responses. All participants signed their consent at the time of answering the diagnostic test, in accordance with the guidelines of the Code of Ethics of the American Psychological Association (2017). The data were kept in a protected file and used exclusively for academic and research purposes.

2.5 DATA ANALYSIS

Descriptive statistics of the total score (mean, standard deviation, minimum, maximum and median) were calculated. For the item analysis, the following were obtained: the difficulty index (p), defined as the proportion of participants who answered each item correctly with respect to the number of those who answered it; and the corrected point-biserial correlation (r_{pb}), which expresses the relationship between the score on the item and the total score of the test excluding that item. Reliability was estimated using Cronbach's alpha coefficient (α). The items with omissions were analyzed based on the valid cases available for each item. All analyses were performed with the jamovi statistical package, version 2.4 (The jamovi Project, 2023).

3 RESULTS

3.1 ITEM ANALYSIS

Table 1 presents the psychometric analysis of each item of the diagnostic test. The results show that Item 1 turned out to be the easiest item in the test ($p = 0.820$), classified as an easy item according to the criteria of Muñiz (2010), although with good discrimination ($r_{pb} = 0.533$). Item 2 presented the highest level of difficulty ($p = 0.360$), placing it in the difficult category, and obtained the highest point-biserial correlation of the instrument ($r_{pb} = 0.709$), which indicates an excellent discriminative capacity. Items 3 and 7 showed low discrimination ($r_{pb} = 0.290$ and 0.257 , respectively), which suggests that these items have limitations in differentiating between students with greater and lesser trigonometric proficiency. Item 8 represented the most critical case of the instrument, with a practically zero point-biserial correlation ($r_{pb} = 0.001$), which indicates that this item does not discriminate between the performance levels of the examinees and, therefore, should be reviewed in its wording, answer options or correction key, or eliminated from the instrument. Items 4, 7, 9 and 10 presented omitted responses from some students, with between one and four participants who did not respond to each of these items.

Table 1

Psychometric analysis of the items of the trigonometry diagnostic test

Item	p	Difficulty Level	rpb	Level discrimination of	Decision
1	0.820	Very easy	0.533	Good	Store
2	0.360	Difficult	0.709	Excellent	Store
3	0.540	Moderate	0.290	Low	Review
4a	0.413	Moderate	0.656	Very good	Store
5	0.680	Moderate	0.440	Good	Store
6	0.460	Moderate	0.478	Good	Store
7a	0.500	Moderate	0.257	Low	Review
8	0.640	Moderate	0.001	Does not discriminate	Delete
9a	0.438	Moderate	0.462	Good	Store
10a	0.429	Moderate	0.464	Good	Store

Grade. p = difficulty index; rpb = point-biserial correlation corrected. Criteria: p < 0.30 = very difficult; 0.30–0.50 = difficult; 0.51–0.70 = moderate; 0.71–0.90 = easy; > 0.90 = very easy. Discrimination: rpb < 0.20 = unacceptable; 0.20–0.29 = low; 0.30–0.39 = acceptable; ≥ 0.40 = good to excellent (Muñiz, 2010). a Items with omitted answers.

3.2 INSTRUMENT RELIABILITY

The reliability of the diagnostic test, estimated using Cronbach's alpha coefficient, was $\alpha = 0.748$. According to the criteria established by George and Mallery (2003), this value corresponds to acceptable reliability. This result indicates that the items of the instrument show an adequate level of internal consistency, that is, that together they tend to measure the same underlying construct: the mastery of basic contents of university trigonometry. In the context of diagnostic and training instruments whose purpose is not to make high-impact decisions on individuals, but rather to identify strengths and areas of opportunity as a group, a coefficient of this magnitude is considered technically sufficient (Morales Vallejo, 2007).

3.3 DESCRIPTIVE STATISTICS OF THE TOTAL SCORE

The descriptive statistics of the total score of the 50 participants are presented in Table 2. The mean scores were $M = 5.20$ ($SD = 2.73$), with a range ranging from 1 to 10 points, and a median of 5.0. These results indicate that, on average, students answered half of the items of the diagnostic test correctly, with a considerable dispersion between subjects. The breadth of the range (9 points) and the value of the standard deviation suggest a remarkable heterogeneity in the level of trigonometric proficiency of the sample, which is consistent with the diagnostic

nature of the instrument and with the diversity of previous academic trajectories of first-year university students.

Table 2

Descriptive statistics of total score on the trigonometry diagnostic test (N = 50)

M	OF	Min.	Max.	Medium
5.20	2.73	1	10	5.0

Note. M = mean; SD = standard deviation; Min. = minimum score; Max. = maximum score. Maximum possible score = 10. N = 50.

4 DISCUSSION

The present study analyzed the psychometric properties of a trigonometry diagnostic test applied to 50 university students. The results obtained offer empirical evidence on the technical quality of the instrument and allow the formulation of precise recommendations for its improvement and pedagogical use.

In relation to reliability, the Cronbach's alpha coefficient obtained ($\alpha = 0.748$) is within the acceptable range according to the criteria of George and Mallery (2003), and is consistent with the values reported in similar studies of validation of mathematics tests at a higher level. It should be noted that the interpretation of the alpha must consider the purpose of the instrument. Unlike professional certification exams, which require reliability coefficients equal to or greater than 0.90 due to their high-impact nature (AERA et al., 2014), a diagnostic instrument whose objective is to offer guidance to the teacher on the group level of their students can be considered technically adequate with values of $\alpha > 0.70$ (Morales Vallejo, 2007). In this sense, the results suggest that the diagnostic test analyzed has sufficient internal consistency for the use for which it is intended.

The analysis of items revealed a differentiated picture regarding the technical quality of the items. Most of the items presented moderate difficulty indices (p between 0.413 and 0.640) and point-biserial correlations greater than 0.40, which indicates an adequate discriminative capacity. These results are favorable from the perspective of the classical theory of tests, since items with moderate difficulty and good discrimination are those that maximize diagnostic information and the variance of scores (Muñiz, 2010). Item 2 was particularly informative, combining a high level of difficulty ($p = 0.360$) with the highest discriminative correlation of the instrument ($r_{pb} = 0.709$), which makes it a high technical quality reagent.

However, three items deserve special attention. Items 3 and 7 showed low discrimination ($r_{pb} = 0.290$ and 0.257 , respectively), which may be due to various causes: ambiguity in the

wording, non-functional response options, or that the content evaluated in these items does not sufficiently correspond to the general construct of the instrument (Martínez Arias et al., 2006). It is recommended to review the formulation of these items, submit the statements and alternatives to expert judgment and, if possible, pilot them again after their reformulation. The most critical case is that of Item 8, whose practically zero point-biserial correlation ($r_{pb} = 0.001$) indicates that the performance in this item is independent of the general performance in the test. This finding suggests that the item could be measuring a concept outside the central construct, present errors in the correction key, contain ambiguous wording, or be evaluating content that is not consistent with the rest of the instrument. Its elimination or substantive reformulation is recommended before a new application.

Regarding the items with omissions (Items 4, 7, 9 and 10), although the number of missing responses was small (between one and four participants per item), this phenomenon constitutes a methodological limitation that deserves consideration. The absence of an answer may reflect perceived difficulty, lack of time, ambiguity of the statement or simply carelessness of the examinee (Muñiz, 2010). In future applications, it would be advisable to implement strategies to reduce lost data, such as explicitly indicating that each item is answered, even if you are not completely sure of the answer.

The study has some aspects that should be considered when interpreting its results. First, the sample size ($N = 50$), although it allowed the analysis to be carried out, could be expanded in future research to strengthen the stability of the estimates of the parameters of the items, as suggested by standards of classical psychometric analysis (Martínez Arias et al., 2006). Second, the use of non-probability sampling offers a valuable insight into the institutional context analyzed, although it would be advisable to complement it with other types of sampling to favor a greater generalization of the findings. Finally, the instrument was examined from the classical theory of tests, which provides a solid basis, and opens the possibility of incorporating more advanced approaches, such as item response theory, in subsequent studies to enrich the analysis.

Based on these considerations, the following lines of future research are proposed: to expand the sample to a minimum of 200 participants from different institutions to obtain more robust and generalizable psychometric estimates; Perform a confirmatory factor analysis to evaluate the dimensionality of the instrument and verify if the items are grouped into factors that correspond to the evaluated contents.

From the pedagogical point of view, the findings of this study have direct implications for university teaching practice. The heterogeneity observed in the scores ($M = 5.20$; $SD = 2.73$;

range: 1–10) reflects substantial differences in students' level of trigonometric readiness upon entering college. This result is consistent with what was reported by Castillo Márquez et al. (2025), who documented that new university students in engineering at the Autonomous University of Nayarit obtained only 48.6% of correct answers in a trigonometry diagnostic instrument, evidencing generalized deficiencies that compromise performance in exact science courses. This variability justifies the design of differentiated teaching strategies that cater to both students with deficient mastery —particularly in the contents associated with the most difficult items— and those with advanced mastery. The diagnostic exam, once the problematic items have been refined, can become a valid and reliable tool for making informed pedagogical decisions at the beginning of the course, contributing to reducing learning gaps from the first weeks of the school year.

REFERENCES

- Abad, F. J., Olea, J., Ponsoda, V., & García, C. (2011). *Medición en ciencias sociales y de la salud*. Síntesis.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct (2002, amended effective June 1, 2010, and January 1, 2017)*. <https://www.apa.org/ethics/code>
- American Psychological Association. (2020). *Publication manual of the American Psychological Association (7th ed.)*. <https://doi.org/10.1037/0000165-000>
- Ausubel, D. P., Novak, J. D., & Hanesian, H. (1983). *Psicología educativa: Un punto de vista cognoscitivo (2.ª ed.)*. Trillas.
- Backhoff, E., Larrazolo, N., & Rosas, M. (2000). Nivel de dificultad y poder discriminativo del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). *Revista Electrónica de Investigación Educativa*, 2(1), 1–16. <https://redie.uabc.mx/redie/article/view/15>
- Buendía, G., & Montiel, G. (2011). La trigonometría en el bachillerato: Un estudio socioepistemológico. *Educación Matemática*, 23(1), 5–28.
- Castillo Márquez, D. I. C., López, G. E. P., & Flores, A. R. F. (2025). Entre senos y cosenos: una radiografía de las admisiones universitarias. *Lumen et Virtus*, 16(53), e9171. <https://doi.org/10.56238/levv16n53-090>
- Castillo Márquez, D. I., Jaime Ramírez, B., Mendoza Reyes, S. M., & Gómez Dávalos, A. J. (2025). El conocimiento matemático pre-universitario en estudiantes que ingresan a la

licenciatura en matemáticas: dificultades y acciones en aritmética y álgebra. *Ciencia Y Reflexión*, 4(2), 853–866. <https://doi.org/10.70747/cr.v4i2.294>

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>

Frías-Navarro, D., & Pascual Soler, M. (2012). Prácticas del análisis factorial confirmatorio con AMOS, FACTOR y LISREL. *Acción Psicológica*, 9(1), 71–89. <https://doi.org/10.5944/ap.9.1.11217>

Gamboa Araya, R., & Ballester Alfaro, E. (2010). La enseñanza y el aprendizaje de la trigonometría en el bachillerato. *Cuadernos de Investigación y Formación en Educación Matemática*, 5(6), 99–122.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Allyn & Bacon.

Martínez Arias, R., Hernández Lloreda, M. J., & Hernández Lloreda, M. V. (2006). *Psicometría*. Alianza Editorial.

Montero, I., & León, O. G. (2007). A guide for naming research studies in psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847–862.

Montiel, G., & Buendía, G. (2012). Articulación escolar del estudio de la trigonometría: Un análisis desde la socioepistemología. *Educación Matemática*, 24(3), 107–130.

Morales Vallejo, P. (2007). La fiabilidad de los tests y escalas. Universidad Pontificia Comillas. <https://matcris5.files.wordpress.com/2014/04/fiabilidad-tests-y-escalas-morales-2007.pdf>

Muñiz, J. (2010). Las teorías de los tests: Teoría clásica y teoría de respuesta a los ítems. *Papeles del Psicólogo*, 31(1), 57–66.

Orhun, N. (2001). Errores y concepciones erróneas en la enseñanza de la trigonometría. En D. Coray et al. (Eds.), *One hundred years of L'Enseignement Mathématique* (pp. 208–211). L'Enseignement Mathématique.

Ravela, P., Picaroni, B., & Loureiro, G. (2017). ¿Cómo mejorar la evaluación en el aula? Reflexiones y propuestas de trabajo para docentes. Grupo Magro Editores.

The jamovi Project. (2023). *Jamovi (Version 2.4) [Software]*. <https://www.jamovi.org>

Vargas Porras, P. A., & Montoya Vargas, J. (2012). La evaluación diagnóstica, formativa y sumativa en la enseñanza de la ciencia. *Revista Electrónica Educare*, 16(1), 113–124. <https://doi.org/10.15359/ree.16-1.8>