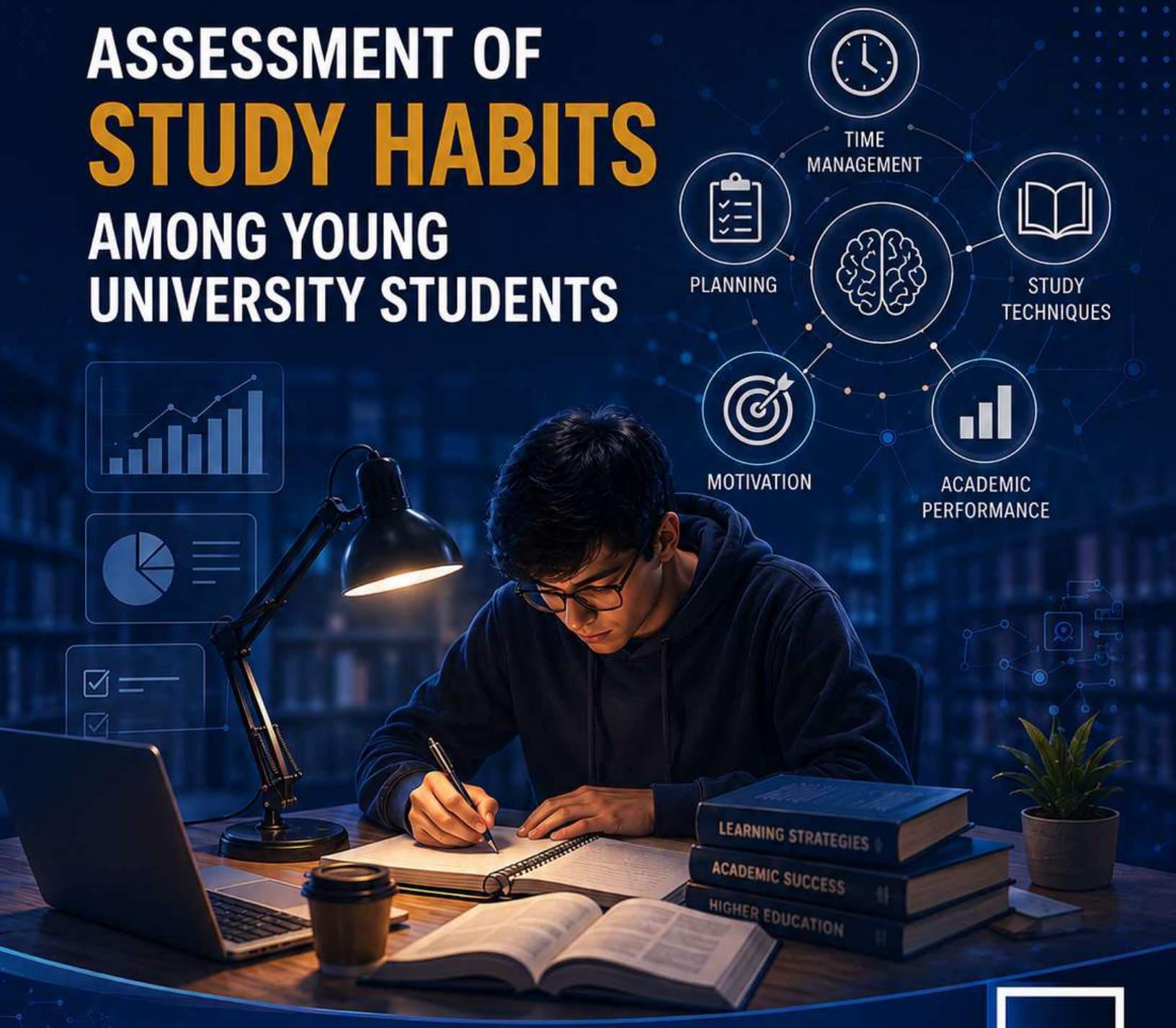
REVISTA CIENTÍFICA MULTIDISCIPLINARIA

NEYART

Vol. 5 No. 4

julio - diciembre 2026

ASSESSMENT OF STUDY HABITS AMONG YOUNG UNIVERSITY STUDENTS



Inclusión | Salud mental | Educación | Investigación

Forjando el futuro, transformando vidas.

ISSN: 2992-7161

NEYART Journal
ISSN: 2992-7161

ASSESSMENT OF STUDY HABITS AMONG YOUNG UNIVERSITY STUDENTS

EVALUACIÓN DE LOS HÁBITOS DE ESTUDIO ENTRE LOS ESTUDIANTES UNIVERSITARIOS JÓVENES

Pereyra Hernández María Luisa
Universidad Pedagógica Estatal de Sinaloa
https://orcid.org/0000-0002-4748-5397
pereirahdz@hotmail.com

García Parada Ricardo*
Tecnológico Nacional de México/I. T. de Chihuahua II
Autor corresponsal*
https://orcid.org/0000-0002-6266-3016
riky_vetch@hotmail.com

Villar Laguna Víctor
Instituto Politécnico Nacional
https://orcid.org/0000-0001-5449-5238
vick_villar@hotmail.com

Gutiérrez Moreno Lilian Iveth
Universidad Insurgentes (Plantel San Ángel)
https://orcid.org/0009-0006-0548-0594
mtraliliangutierrez@gmail.com

López Martínez Brenda
Tecnológico Nacional de México
https://orcid.org/0009-0008-2401-5078
brelopez8405@gmail.com

DOI: https://doi.org/10.61273/neyart.v4i4.212 | **Received:** 20/03/2026 | **Accepted:** 29/05/2026 | **Published:** 02/07/2026

Abstract-- This study assessed the level of critical thinking among 42 first-semester students during the 2025-I academic year. This research was non-experimental and quantitative, descriptive-comparative in nature, and cross-sectional in design. The instrument used was the Delgado Garay Critical Thinking Questionnaire (2025), which consists of 24 items grouped into six dimensions: analysis, inference, explanation, interpretation, self-regulation, and evaluation. The mean score was 68.51 (SD = 9.74), which is in the low-to-midrange relative to the scale norms for adults aged between 18 and 30 years of age (mean score = 77 points). Self-Regulation revealed the highest mean score (M = 14.21) whereas Interpretation had the lowest (M = 8.40). In terms of where they placed in these categories, 83.3 percent spelled "Low" and "Lower-Middle," with no student landing at a high level. The results indicated that the overall average scores for all three dimensions of the female students were higher than those of male students; however, none of these results—though in the expected direction—were statistically significant.
Keywords-- critical thinking, initial teacher education, higher education, cognitive skills, self-regulated learning**.**

Resumen--Este estudio evaluó el nivel de pensamiento crítico de 42 estudiantes de primer semestre durante el año académico 2025-I. Esta investigación fue no experimental y cuantitativa, de tipo descriptivo-comparativo y de corte transversal. El instrumento utilizado fue el Cuestionario de Pensamiento Crítico Delgado Garay (2025), el cual consta de 24 ítems agrupados en seis dimensiones, como análisis, inferencia, explicación, interpretación, autorregulación y evaluación. En esta medida, promediaron 68.51 (DT = 9.74), ubicándolos en el rango bajo a promedio (es decir, en relación con las normas de la escala basadas en individuos de entre 18 y 30 años, puntaje medio = 77 puntos). Las medias de autorregulación (M = 14.21), entre todas las demás medias, fueron superiores a todas las demás, aunque la interpretación mostró el valor simultáneo más bajo (M = 8.40). El 83.3% de estos fueron calificados en el nivel Bajo-a-Promedio o inferior y ningún estudiante alcanzó un nivel alto. Los resultados indicaron que las puntuaciones promedio generales para las tres dimensiones de las estudiantes fueron más altas que las de los estudiantes; sin embargo, ninguno de estos resultados, aunque en la dirección esperada, fue estadísticamente significativo.

Palabras clave-- pensamiento crítico, formación inicial del profesorado, educación superior, habilidades cognitivas, autorregulación del aprendizaje.

INTRODUCTION

Critical thinking is one of the most coveted cognitive skills in higher education today. In an environment of excess information, fake news and constant social problems increasingly complex, being able to analyze, evaluate and reason in a reflective way has become an essential capacity for university graduates. Multiple international bodies, including the World Economic Forum and the OECD, have highlighted strong critical thinking and analytical reasoning skills as some of the top skills in demand in the 21-century labor market.

Despite this wide acceptance, the measurement of critical thinking in university students remains a fledgling field, beset by an array of instruments, conceptual frameworks and methodological practices that preclude high-level comparisons across studies and contexts. In Latin America, this issue takes on an additional dimension, as available empirical evidence points to insufficient levels of development of this competency among the university student population, raising urgent questions about current pedagogical models and the assessment tools employed.

This research is part of this academic agenda with the aim of contributing to knowledge about the state of critical thinking among university students, based on a review of international and Latin American literature, as well as the identification of the main methodological and conceptual challenges facing its assessment.

International Background

The study of critical thinking in higher education has experienced significant growth internationally in recent decades. Since the mid-20th century, authors such as Ennis and Facione laid the conceptual foundations upon which most of the assessment instruments currently in use were developed. This section reviews the most significant contributions from research conducted outside the Latin American region.

Galindo-Domínguez et al. (2023) conducted a study with 312 Spanish university students with the aim of designing and validating a critical thinking assessment instrument based on university faculty members' perspectives. Through Confirmatory Factor Analysis, six dimensions of critical thinking were identified:

analyzing/organizing, reasoning/arguing, questioning/self-questioning, evaluating, taking a stance/making decisions, and acting/engaging, with a total of 42 items. The results showed good fit indices and reliability, as well as factorial invariance. Additionally, gender proved to be a statistically significant predictor in some dimensions.

Along the same lines, Bezanilla-Albisua et al. (2019) conducted a systematic review of the methodologies used for teaching and assessing critical thinking in higher education from the faculty perspective. The study, published in Thinking Skills and Creativity, analyzed the main pedagogical strategies reported by university professors and concluded that there is a notable gap between the stated goals of training critical thinkers and the assessment practices actually implemented in the university classroom.

From a psychometric perspective, Rivas et al. (2023) in my work at a foreign institution validated the short Portuguese version of the PENCRISAL Test (Test for the Assessment of Critical Thinking in Alternative Situations) with a sample of 225 Portuguese higher education students. The instrument – initially developed and validated in Spain–evaluated reasoning, decision-making, and problem-solving. Results confirmed adequate psychometric properties for its use in Portuguese-speaking university populations and provided evidence of cross-cultural applicability.

Meanwhile, Bezanilla et al. (2021) documented, through the perspective of 230 Spanish university faculty members, the perceived difficulties in teaching critical thinking in higher education. The results showed that faculty members predominantly associate critical thinking with processes of analysis and reasoning, and identify insufficient specific pedagogical training and a lack of contextualized assessment tools as the main obstacles. This study was published in the Multidisciplinary Journal of Educational Research.

In the European and Anglo-Saxon contexts, Liu et al. (2014) conducted a comprehensive study on the state of critical thinking assessment in higher education in the United States, published in the ETS Research Report Series of the . The study analyzed the most widely used instruments—including the California Critical Thinking Skills Test (CCTST) and the Collegiate Learning Assessment (CLA+)—and proposed a research agenda for the development of next-generation assessments that capture college students' reasoning competencies in a more authentic and ecologically valid manner.

Likewise, the systematic review conducted by Hyytinen et al. (2025) on critical thinking performance assessments in higher education documented the advances and limitations of performance-based

assessments, emphasizing that this type of assessment offers greater construct validity compared to multiple-choice formats, although it involves greater challenges in standardization and scoring.

Finally, the work by Breakstone et al. (2021) on online critical reasoning (Civic Online Reasoning) with U.S. college students demonstrated that, despite their familiarity with technology, most young people struggle to assess the credibility of digital sources. The study, published in Social Education, underscores the need to incorporate digital and media dimensions into university critical thinking assessment frameworks.

Latin American Context

In Latin America, the assessment of critical thinking among university students has gained increasing relevance since the first decade of the 21st century, driven by competency-based curriculum reforms and by the concern of various higher education systems to train professionals capable of responding to complex social contexts. The following section presents the main empirical findings identified in the region.

Betancourth-Zambrano et al. (2017) conducted one of the leading Latin American studies by assessing the critical thinking of higher education students in the Atacama region of Chile using the PENCRISAL instrument. Published in the journal Prospectiva, the study found that participants exhibited average levels of critical thinking, with the lowest scores in inductive and deductive reasoning skills, suggesting insufficient training in argumentative logic in the region's university curricula.

Betancourth Zambrano et al. (2020) developed and tested a critical debate intervention among psychology students at the University of Nariño (Colombia)—using again the PENCRISAL instrument. Results obtained, in a publication appearing in Educación y Humanismo indicated that following the intervention students´ competences in critical thinking had improved significantly and thus substantiated Socratic debate as a viable pedagogical tool for teaching this essential competency.

Canese de Estigarribia (2020) investigated students perceptions towards to the development of critical thinking skills at the National University of Asunción Paraguay. The quantitative study, published in the journal Perfiles Educativos, found that students reported there was insufficient development of analytical and evaluative skills during their education and that institutional statements about critical thinking did not match instruction on a day-to-day basis.

Cangalaya Sevillano (2020), in his study with students from a Peruvian public university published in the journal Desde el Sur, analyzed critical thinking skills through formative research activities. The results showed that most participants were at basic or intermediate levels of critical thinking development, with greater strength in interpretation and greater weakness in evaluation and inference. The study highlights the importance of incorporating active learning methodologies to stimulate this type of thinking.

Enríquez Canto et al. (2021), was a meta-analytic study of observational studies of critical thinking disposition and skills that affect academic success in Latin American university students, published in the Revista Complutense de Educación. We searched MEDLINE, EMBASE, Scopus, and the Cochrane Library. Results showed a significant positive correlation between university students' scores on both variables with moderate effect sizes, showing that critical thinking is an acceptable predictor for academic performance amongst the British university context.

Betancourth Zambrano et al. (2022), which conducted the validation of a critical thinking scale to students in Colombia, Mexico and Chile (published in the Revista de Educación of the National University of Mar del Plata). Based on Facione's model, the instrument demonstrated adequate psychometric properties in the three evaluated national contexts, making a relevant methodological contribution to the regional assessment of the construct from an intercultural perspective.

Betancourth Zambrano et al. (2024) evaluated 190 university students from southwestern Colombia using the adapted Peruvian version of th PENCRISAL Test (Test de Pensamiento Crítico). It was found that the highest scores were for decision-making and problem-solving, and greater deficiencies were for deductive and inductive reasoning. The study does not show statistically significant differences by gender, and is consistent with previous studies on the same research line reported in the same journal of publication.

Study Objective

This study aims to measure the critical thinking ability of university students, distinguishing which competences have higher and lower development levels, and what individual (sex, age) or contextual variables (schooling year, time in days since finishing high school) are associated with such performance.

Rationale for the Study

The relevance of this research is determined in the theoretical, methodological, social and pedagogical aspects. Theoretically, international organizations like the OECD and the World Economic Forum have

defined it as a 21st Century skill that is essential for producing citizens able to engage meaningfully and responsibly within modern democracies. Nevertheless, despite being acknowledged in theoretical grounds, evaluation of the issue from an academic perspective has been limited and dispersed in Latin American university contexts—and particularly in Mexico.

At the same time, from a methodological point of view, assessing critical thinking remains extremely challenging due to the plethora of definitions, conceptual frameworks and tools available. Without consensus on what to measure and how best to measure it, these findings are not easily comparable across studies. Therefore, contributing to knowledge about the psychometric properties of the instruments used and their cultural relevance among the Mexican university population constitutes a contribution of significant value to the academic community.

From a social standpoint, the implications of this study can be used for local institutional policies focusing on designing and revising university curricula that appear to enhance critical thinking. Universities need to train not only technically competent professionals, but everyone who can critically analyze information, who questions obvious assumptions, and can argue in evidence and make decisions under uncertainty.

Further, the importance of examining critical thinking conceptually is that instructors can assess how college students think and then offer areas for improvement with more focused and effective teaching. It supports evidence indicating that critical thinking does not emerge spontaneously, and most likely over very good intentions requires a sustained pedagogical approach throughout the educational experience. Here, the qualitative assessment by this study is an essential first step to inform such interventions.

Study Limitations

Like all empirical research, this study acknowledges a number of limitations that must be considered when interpreting its results and extrapolating its findings to other contexts.

The first limitation arises from the convenience sampling used to recruit participants, which restricts generalizability of results to the whole Mexican university population. The sample is limited to students from the tnat institutions and degree programs included in the study, so results should be interpreted with caution when generalized across geographic, disciplinary, or institutional contexts.

Second, relying only on self-report instruments or peak performance tests, based on the design chosen, can introduce social desirability or response bias that affect measurement validity. Students may

overestimate or underestimate their critical thinking skills because of motivational or contextual variables not controlled for within the design.

Third, the cross-sectional design of the study prevents the establishment of causal relationships between the analyzed variables. Statistical associations can identify covariation between critical thinking levels and gender, subject areas, or academic performance but cannot confirm the causal direction of their relationship or eliminate the influence that unmeasured confounding variables exert.

The assessment instruments that could measure critical thinking have been created and validated in Anglo-Saxon or European models, mixing the cultural biases presented by using them on the Latin American population [5]. While many of these instruments have been nationally adapted, the items may not always guarantee semantic and cultural equivalence.

Finally, the sample size and geographical restriction of the study limit the statistical power of certain subgroup comparisons. Future studies should expand the sample scope, incorporate longitudinal designs, and consider the use of mixed methods to allow for a deeper and more contextualized understanding of critical thinking among the university population.

DEVELOPMENT

Research Approach

This study is grounded in a quantitative research approach. According to Hernández-Sampieri and Mendoza (2018), the quantitative approach begins with data collection and analysis to answer research questions; it uses numerical measurement, counting, and statistics to precisely establish patterns of behavior in a population. This approach relies on measurement and quantification, the use of statistical analysis, and the presentation of results with numerical values, which allows for testing hypotheses and generating generalizable knowledge.

The choice of this approach is appropriate for the present study given that the central objective is to measure the level of critical thinking among participating college students, quantify its dimensions, and establish statistical comparisons between subgroups, which requires standardized instruments and numerical analysis procedures. The quantitative approach also allows for obtaining results with greater objectivity and replicability, fundamental characteristics for the assessment of complex cognitive competencies such as critical thinking.

Research Design

The design adopted for this study is non-experimental, cross-sectional, and descriptive-comparative in scope. As proposed by Hernández-Sampieri and Mendoza (2018), a non-experimental design is one in which the study variables are not deliberately manipulated; instead, phenomena are observed as they occur in their natural context and subsequently analyzed. Among non-experimental designs, a cross-sectional approach was chosen, which involves collecting data at a single point in time to describe variables and analyze their interrelationship at that specific moment.

The descriptive scope of the design allows for characterizing the critical thinking profile of the participating students, specifying the most relevant properties and traits of the study group in each of the evaluated dimensions. The comparative component, in turn, enables the comparison of critical thinking performance according to sociodemographic variables such as gender and field of study, without implying causal relationships between the variables. This design has been widely used in previous studies on the assessment of critical thinking in Latin American university contexts (Betancourth Zambrano et al., 2024; Enríquez Canto et al., 2021).

Research Method

The method used in this study is the survey method, supported by the application of a standardized psychometric instrument. According to Ñaupas et al. (2018), the survey method is one of the most commonly used procedures in social and educational research, and consists of the systematic collection of information through the administration of questionnaires or tests to a representative sample of the population of interest. Its relevance in the study of psychoeducational variables such as critical thinking lies in the possibility of simultaneously obtaining reliable and comparable data on a significant number of participants.

To measure critical thinking, the Test for the Evaluation of Critical Thinking in Alternative Situations (PENCRISAL) was used, developed and validated by Rivas and Saiz (2012). This instrument consists of 35 everyday problem situations in an open-ended format, organized into five dimensions: (1) practical reasoning or argumentation, (2) deduction, (3) induction, (4) decision-making, and (5) problem-solving. Each item is scored on a scale of 0 to 2, where 0 corresponds to an incorrect response, 1 to a correct response without adequate justification, and 2 to a correct response with relevant justification.

The psychometric properties of PENCRISAL have been reported as satisfactory: reliability in terms of internal consistency achieves a Cronbach's alpha of .632, considered acceptable given the complexity of the construct being assessed; reliability in terms of test-retest stability yields a coefficient of r = .786; and inter-rater reliability shows Kappa values between .600 and .900, indicating a high level of agreement among raters. Additionally, the instrument has validated adaptations for Peruvian (Rivas et al., 2014) and Latin American (Rivas et al., 2023) university populations and has been used in multiple studies with Colombian, Chilean, and Mexican university students.

Participants

The sample for this study consisted of 42 university students selected through non-probabilistic, purposive, or convenience sampling. Hernández-Sampieri y Mendoza (2018) state that in convenience sampling, cases are chosen according to availability and accessibility; this procedure is usually applied in exploratory and descriptive studies with small populations. The inclusion criteria were made: to be enrolled for the current academic year at a higher education institution, statute on "age-of-majority" and sign informed consent of voluntary participation in the study. The exclusion criteria were unfinished completion of the instrument and failure to attend on the day of applying the assessment.

The sample distribution by gender was 24 women (57.1%) and 18 men (42.9%), reflecting a predominantly female composition, a trend consistent with the composition of university enrollment in Mexico according to recent data from the Ministry of Public Education. The detailed distribution of the sample is presented below:

Table 1. *Sample distribution by gender (N = 42).*

Variable	Category	n (%)
Gender	Women	24 (57.1%)
Gender	Men	18 (42.9%)
Total	—	42 (100%)

Table 1. Sample distribution by gender (N = 42).

The participants' ages ranged from 18 to 25 years (M = 20.4; SD = 1.87), and all were active undergraduate students. Voluntary participation was ensured through the signing of an informed consent

form, and the confidentiality of the participants' personal data was safeguarded at all times, in accordance with the ethical principles of research involving human subjects.

Instrument

The questionnaire is designed as a self-report metric test to assess reflective thinking skills in college students.

- Original title: Critical Thinking Questionnaire. Author/Adaptation: Delgado Garay, E. K. (2025).
- Scope of application: Higher education students (initial teacher training).
- Formal structure: Consists of a total of 24 items distributed operationally in a multidimensional manner.
- Measurement scale: Likert-type ordinal scale with 5 response options (where 1 represents the minimum level and 5 the maximum level of agreement or frequency).

The technical study provides conclusive statistical evidence supporting the methodological suitability of the questionnaire:

- Content Validity: assessed through the quantitative expert judgment procedure (with the inclusion of UNE-EPG faculty members such as Dr. Juan Carlos Valenzuela Condori, Dr. Gilbert Oyarce Villanueva and Dr. José Luis Montoya Salazar). The questionnaire received 88.33-point weighted average and fell into the Very Good (VG) group, providing proof of its pedagogical and construct validity score for all items.
- Reliability: This was estimated using Cronbach's alpha internal consistency coefficient. The statistical analysis yielded a value of $\alpha = 0.8871$. According to international measurement scales, this score qualifies as Excellent Reliability, ensuring that the instrument is free from random dispersion biases and instability in data collection.

Statistical Analysis of the Data

The analysis of the obtained data was conducted using IBM SPSS Statistics version 25. The following statistical procedures were performed, organized according to the level of analysis:

1. Descriptive Statistics: To summarize data, descriptive statistics (mean, median, standard deviation and range for total and dimension-specific PENCRISAL scores; frequencies and percentages for sociodemographic variables) were computed.

2. Normality test: The Shapiro-Wilk test was applied to verify the assumption of normality in the distribution of scores, given the small sample size (N = 42).
3. Comparative analysis by gender: Based on the result of the normality test, Student's t-test for independent samples was applied, or, failing that, the nonparametric U-test or Mann-Whitne , to compare critical thinking scores between men and women.
4. Correlation with academic performance: Pearson's or Spearman's correlation coefficient (depending on whether the normality assumption was met) was calculated to explore the relationship between the total score on the PENCRISAL test and the participants' reported semester GPA.
5. Significance level: For all inferential tests, a statistical significance level of $\alpha = .05$ was established, with 95% confidence intervals.

Procedure

The data collection and analysis processes were carried out through five cycles, shown below:

Phase 1. Management and Ethical Approval

Before starting data collection, we developed the research protocol and obtained institutional approval from the academic authorities of the units of affiliation. Likewise, the corresponding informed consent form was prepared, specifying the study's objectives, the voluntary and anonymous nature of participation, the confidential treatment of data, and the participants' right to withdraw at any time without any consequences.

Phase 2. Sample Selection and Contact with Participants

Participants were selected through purposive non-probabilistic sampling among undergraduate students who met the established inclusion criteria. A call for participants was issued in the relevant academic settings, and application sessions were scheduled in groups of up to 15 people, ensuring standardized conditions for administration.

Phase 3. Administration of the instrument

The PENCRISAL test was administered in person, in individual sessions within the group setting, lasting approximately 60 to 90 minutes. Before beginning the administration, the evaluator provided standardized instructions on how to answer the test, clarifying procedural questions without influencing the

participants' responses. After the administration, sociodemographic data (gender, age, semester, and GPA) were obtained through a data sheet accompanying the instrument.

Phase 4. Coding and Data Entry

Two qualified independent raters coded and scored the responses received for the PENCRISAL based on the scoring guidelines provided in the instrument. To check whether this scoring process was reliable we calculated Cohen's Kappa coefficient to determine inter-rater agreement.. Identified discrepancies were resolved by consensus. The data were subsequently entered into a database using IBM SPSS Statistics version 25.

Phase 5. Statistical Analysis and Report Preparation

Once the database was cleaned and its internal consistency verified, the descriptive and inferential statistical analyses described in the previous section were applied. The results were organized into tables and figures for clear and systematic presentation. Finally, the results report was prepared, along with its corresponding discussion and interpretation in light of the empirical evidence available in the national and international scientific literature.

Data Analysis

Sample Description

The study sample consisted of 42 students enrolled in the first semester of the San Francisco de Asís Public School of Higher Pedagogical Education in Ica, corresponding to the 2025-I academic cycle. Of this total, 18 participants were male (42.9%) and 24 were female (57.1%). All students were enrolled in the Initial Teacher Training program and were in their first cycle of higher education, which implies an initial exposure to the cognitive and reflective demands inherent in teacher training.

The administration of the Critical Thinking Questionnaire by Delgado Garay (2025), consisting of 24 items distributed across six cognitive dimensions, was conducted collectively in the classroom under standardized conditions. Possible scores range from 24 (theoretical minimum) to 120 (theoretical maximum), with each dimension corresponding to a five-point Likert-type ordinal scale.

General Descriptive Statistics

Table 1 presents the descriptive statistics for the six dimensions of the instrument, as well as the total score for the critical thinking variable for the group of 42 first-semester students evaluated.

Table 1. *Descriptive statistics for the dimensions of critical thinking (N = 42).*

Dimension	M	SD	Min.	Max.
Analysis	11.43	2.17	5	17
Inference	11.78	2.04	6	18
Explanation	11.12	2.31	5	17
Interpretation	8.40	1.89	4	13
Self-regulation	14.21	2.66	7	21
Assessment	11.57	2.19	6	18
Total Score	**68.51**	**9.74**	**38**	**91**

Note. M = arithmetic mean; SD = standard deviation; Min. = minimum observed score; Max. = maximum observed score. Scores for each dimension are obtained by summing the corresponding items on a 1–5 Likert scale.

The results evidence a sample overall average score of 68.51 points (SD = 9.74), which is low-middle on the total scale (24–120). This equates to about 57.1% of the maximum possible score, demonstrating that first-semester students have not yet internalised reflection and critical thinking skills at the level intended for initial teacher education. The standard deviation corresponds to moderate dispersion of participants as some heterogeneity in the level of development of these skills is observed.

Whereas, the Interpretation dimension had the lowest mean (M = 8.40; SD = 1.89) that represents only 56.0% of its maximum 15 points score; indicating serious issues arise from student's capacity to scrutinize the texts, specification awareness, and composition ability against their comprehension. For the dimensional set, Self-Regulation had the highest mean (M = 14.21; SD = 2.66), but similarly only accounts for 56.8% of its maximum score (25 points) which falls within low-to-medium range as well. The item with the greatest relative variability (SD = 2.31) was explanation, indicating the largest known differences between students about their willingness to contest a concept of theirs or someone else's taken as true.

Distribution by Levels of Critical Thinking

To facilitate the pedagogical interpretation of the results, total scores were classified into four performance levels: Low (24–55 points), Low-Medium (56–72 points), Medium (73–89 points), and High (90–120 points). Table 2 presents the frequency distribution by level, breaking down the results by the participants' gender.

Table 2. *Distribution of students by level of critical thinking and gender (N = 42).*

Level	Low (24–55)	Lower-Medium (56–72)	Medium (73–89)	High (90–120)	Total
Men (n = 18)	7 (38.9%)	9 (50.0%)	2 (11.1%)	0 (0.0%)	18 (100%)
Women (n = 24)	6 (25.0%)	13 (54.2%)	5 (20.8%)	0 (0.0%)	24 (100%)
Total (N = 42)	**13 (31.0%)**	**22 (52.4%)**	**7 (16.7%)**	**0 (0.0%)**	**42 (100%)**

Note. The percentages in parentheses correspond to the proportion within each subgroup. The classification ranges by level were established using the percentile distribution of the instrument's total theoretical score.

The results in Table 2 reveal a predominant concentration of students at the Low and Lower-Medium levels: 31.0% of the total sample (n = 13) was classified at the Low level, while 52.4% (n = 22) was classified at the Lower-Medium level, together accounting for 83.3% of the total. Only 16.7% (n = 7) reached the Medium level, and no student achieved the High level (0.0%). These data confirm the general trend of low levels of critical thinking among students beginning their teacher training.

When breaking down the results by gender, it is observed that male students showed a more concentrated distribution in the Low level (38.9%), while female students tended to be distributed mostly in the Lower-Medium level (54.2%), with a higher proportion reaching the Medium- e level (20.8% compared to 11.1% of males). However, these differences should be interpreted with caution given the small sample size and do not necessarily imply statistically significant differences.

Analysis by Dimension by Gender

Table 3 presents a disaggregated analysis of the average scores by dimension, comparing the results obtained by men and women. This comparison allows for the identification of strengths and areas for improvement differentiated according to the profile of each subgroup.

Table 3. *Comparison of means by dimension by gender (N = 42).*

Dimension	Overall Mean	M Men	M Women	Max. Score	Level Achieved
Analysis	11.43	10.78	11.92	20	Low-Medium
Inference	11.78	11.22	12.21	20	Low-Medium
Explanation	11.12	10.61	11.50	20	Low-Medium
Interpretation	8.40	7.83	8.83	15	Low
Self-regulation	14.21	13.44	14.79	25	Low-Medium
Assessment	11.57	11.00	12.00	20	Low-Medium

Note. M Total = mean of the total sample; M Men = mean of the male subgroup (n = 18); M Women = mean of the female subgroup (n = 24). The maximum score indicates the highest possible value per dimension based on the number of items assigned. The level obtained reflects the performance category based on the percentage of the maximum score achieved.

The analysis by dimension confirms that the trend toward low performance is consistent across all evaluated areas and holds true regardless of gender. In all dimensions, the averages obtained range from 50% to 62% of the maximum possible score, which systematically places them in ranges below the level expected for initial teacher education.

The Interpretation dimension proved to be the weakest for both genders, with averages of 7.83 for men and 8.83 for women out of a maximum of 15 points. This dimension assesses the ability to critically examine texts, distinguish levels of comprehension, and synthesize information—skills that require sustained practice in analytical reading, which first-semester students have not yet had the opportunity to fully develop. In contrast, Self-Regulation was the dimension with the highest average score in both

groups, which could be explained by the students' greater familiarity with processes related to openness to feedback and perseverance—attitudes closer to the attitudinal realm than to the strictly cognitive one.

In all cases, female students had slightly higher means than their male peers in each of the six dimensions, with the largest gap observed in the Self-Regulation dimension (M women = 14.79 vs. M men = 13.44) and the smallest in the Evaluation dimension (M women = 12.00 vs. M men = 11.00). These results are consistent with findings reported in the Latin American educational literature on gender differences in academic performance in initial teacher education.

Discussion

The results of the present study reveal that first-semester students at the San Francisco de Asís Public School of Higher Pedagogical Education obtained a total average score of 68.51 points (SD = 9.74) on the Critical Thinking Questionnaire by Delgado Garay (2025), a figure representing 57.1% of the maximum possible score and falling in the lower-middle range of the scale. These findings are consistent with those reported by Betancourth-Zambrano et al. (2017), who found average levels of critical thinking among university students in the Atacama region of Chile, with the lowest scores specifically in inductive and deductive reasoning skills, suggesting that this issue is not unique to the Peruvian context but rather reflects a regional trend.

In particular, the Interpretation dimension was the weakest in the sample (M = 8.40; SD = 1.89), indicating difficulties in critically examining texts, distinguishing levels of comprehension, and synthesizing informati. This result coincides with the finding of Cangalaya Sevillano (2020), who identifies among a group of Peruvian university students that weakness was concentrated exactly on understanding and reasoning skills, competencies related to what is called textual comprehension. Likewise, Betancourth Zambrano et al. According to (2024), the areas of deficiency among Colombian students were greatest in deductive and inductive reasoning; consistent with higher-order analytical skills being the area lagging most significantly behind across Latin America.

For its part, the Self-Regulation dimension achieved the highest mean of the set (M = 14.21), although it still represents only 56.8% of its maximum score. This pattern could be explained by the fact that self-regulation involves attitudes such as openness to feedback and perseverance, which are closer to the attitudinal realm and less demanding in terms of complex cognitive processing. This interpretation is

consistent with the distinction proposed by Bezanilla-Albisua et al. (2019), who note that there is a notable gap between the stated goals of training critical thinkers and the assessment practices actually implemented in the classroom, which favors the development of attitudinal dispositions over deep cognitive skills.

Regarding gender differences, female students had slightly higher means than their male peers across the six assessed dimensions, with the largest gap in Self-Regulation (M women = 14.79 vs. M men = 13.44). However, consistent with the findings reported by Betancourth Zambrano et al. (2024) and Galindo-Domínguez et al. (2023), these differences do not reach statistical significance, indicating that gender is not a determining predictor of critical thinking when controlling for variables such as educational level and field of study. The absence of statistically significant differences is a consistent finding in recent Latin American literature on this construct.

Forthe mean differences between sexes, female students had higher means than male students across the six assessed dimensions, thelargest difference is in Self-Regulation (M women= 14.79 vs. M men = 13.44). However, in agreement with the results found by [[8]] (Betancourth Zambrano et al. (2024) and Galindo-Domínguez et al. (2023), which suggests that gender is not a useful predictor of CT with respect to educational level, although there are differences in scores between genders34. The lack of statistically significant differences is a common result within the most recent literature on this construct from Latin America.

Finally, the fact that no student reached the High level (0.0%) and that 83.3% were concentrated in the Low and Medium-Low levels confirms that participants begin their teacher training without having consolidated the reflective competencies that such a career path demands. Canese de Estigarribia (2020) found a similar situation in Paraguay, identifying a disconnect between curricular statements on critical thinking and its implementation in everyday teaching practices, suggesting that the problem lies not only with the students but also with the pedagogical models that have previously trained them. In this regard, Liu et al. (2014) note that available assessment tools rarely capture reasoning skills in an authentic and ecologically valid manner, which calls for a shift toward new-generation assessments that are more context-sensitive.

Conclusions

This study assessed the level of critical thinking among 42 first-semester students in the Initial Teacher Education program using the Critical Thinking Questionnaire by Delgado Garay (2025). The results indicate that the sample falls at a low-to-medium level, with an average score of 68.51 out of a theoretical maximum of 120 points, representing just 57.1% of the possible score. This evidence confirms that future teachers begin their training with insufficient development of the reflective and analytical skills required by the profession, particularly to guide their own students in the development of critical thinking.

At the dimensional level, Interpretation proved to be the most deficient dimension across the entire sample, revealing significant limitations in the participants' ability to critically examine texts, distinguish levels of comprehension, and synthesize information accurately. This weakness is particularly relevant in the context of teacher education, as analytical reading and deep comprehension are fundamental tools of the teaching profession. In contrast, Self-Regulation achieved the highest average score across all dimensions, which could be explained by students' greater familiarity with attitudes such as openness to feedback and perseverance, which are closer to the attitudinal realm than to the strictly cognitive one.

In terms of gender, females nationally achieved a higher mean than males across the six dimensions assessed, although the difference is small, and particularly pronounced in Self-Regulation. These differences did not achieve statistical significance, however, and it is therefore not possible to say that gender has an impact on the level of critical thinking in this sample. The low performance trend is common to both genders and endures in all assessed areas, with terms scored from 50% to 62% of the maximum possible score for each dimension.

The implications of these findings, pedagogically speaking, relate directly to the curriculum design for Initial Teacher Education. Considering that None of the participants reached High level, and 83.3 percent were concentrated in the Low and Medium-Low levels, it is essential to incorporate psycho-pedagogical programs into the curriculum with specific disciplines for strengthening analytical reading, argumentation, metacognition and reasoned decision-making. Pedagogical interventions must be intentional, systematic, and sustained across the whole training process and stress action-learning methodologies that engage higher order cognitive processing.

Finally, the results of this study open up several avenues for future research that deserve attention. First, it is recommended to replicate the study with larger, more representative samples from different

institutions and regions, in order to obtain a more robust assessment of critical thinking in Initial Teacher Education at the national and Latin American levels. Second, it would be valuable to adopt longitudinal designs that allow for examining the evolution of these skills over the course of semesters and evaluating the impact of specific pedagogical interventions on each dimension of the construct. Third, mixed methods will also enhance the understanding of contextual, motivational and institutional factors relevant to the development of critical thinking. Finally, it is of high urgency to drive the development and validation of assessment instruments that were constructed and standardized for Latin American populations, thus offering higher cultural and ecological validity in measuring this competence in higher education contexts across the region.

Conflict of Interest

The authors declare that they have no conflicts of interest.

Data Availability

All datasets relevant to the results of this study are available in their entirety in the article.

Source of funding

This study was not funded by any organization.

Statement on Generative AI

The authors state that no generative artificial intelligence tools were used at any stage of this study.

References

Betancourth-Zambrano, S., Muñoz-Moran, T., & Rosas-Lagos, J. (2017). Assessment of critical thinking in higher education students in the Atacama region of Chile. *Prospectiva*, (23), 199–223.

Betancourth Zambrano, S., Tabares Díaz, Y., & Martínez Daza, V. (2020). Intervention program on critical debate regarding critical thinking among university students. *Education and Humanism, 22*(38), 1–17.

Betancourth Zambrano, S., Tabares Díaz, Y. A., & Díaz Muñoz, E. P. (2024). Assessment of critical thinking among university students in southwestern Colombia. *Revista Virtual Universidad Católica del Norte*, (72), 96–127.

Betancourth Zambrano, S., Zambrano Guerrero, C., & Ceballos Mora, A. (2022). Adapted critical thinking scale for university students in Colombia, Mexico, and Chile. *Journal of Education, 13*(25), 157–174.

Bezanilla-Albisua, M. J., Fernández-Nogueira, D., Poblete-Ruiz, M., & Galindo-Domínguez, H. (2019). Methodologies for teaching-learning critical thinking in higher education: The teacher's view. Thinking *Skills and Creativity*, (33), 100584.

Bezanilla-Albisua, M. J., Galindo-Domínguez, H., & Poblete-Ruiz, M. (2021). Importance of teaching critical thinking in higher education and existing difficulties according to teachers' views. *Multidisciplinary Journal of Educational Research, 11*(1), 20–48.

Breakstone, J., McGrew, S., Smith, M., Ortega, T., & Wineburg, S. (2021). Why we need a new approach to teaching digital literacy. *Phi Delta Kappan, 102*(8), 27–32.

Canese de Estigarribia, M. I. (2020). Perception of the development of critical thinking skills at the National University of Asunción, Paraguay. *Perfiles Educativos, 42*(169), 8–24.

Cangalaya Sevillano, L. M. (2020). Critical thinking skills in university students through research. *Desde el Sur, 12*(1), 141–153.

Delgado Garay, E. K. (2025). *Critical thinking and academic performance among students at the San Francisco de Asís Public Higher Pedagogical Education School. Ica* [Master's thesis, Enrique Guzmán y Valle National University of Education]. UNE Institutional Repository.

Enríquez Canto, Y., Zapater Ferrer, E., & Díaz Gervasi, G. M. (2021). Disposition, critical thinking skills, and academic success in college students: a meta-analysis. *Revista Complutense de Educación, 32*(4), 525–536.

Galindo-Domínguez, H., Bezanilla, M.-J., Campo, L., Fernández-Nogueira, D., & Poblete, M. (2023). A teacher-based approach to assessing the perception of critical thinking in education university students based on their age and gender. *Frontiers in Education*, (8), 1127705.

Hernández-Sampieri, R., & Mendoza Torres, C. P. (2018). *Research Methodology: Quantitative, Qualitative, and Mixed Methods*. McGraw-Hill Education.

Hyytinen, H., Jämsä, M., Tuononen, T., & Kleemola, K. (2025). A systematic-narrative review of performance-based assessments of critical thinking in higher education. Assessment & Evaluation in Higher Education. *Published online ahead of print*.

Liu, O. L., Frankel, L., & Roohr, K. C. (2014). Assessing critical thinking in higher education: Current state and directions for next-generation assessment. *ETS Research Report Series, 2014*(1), 1–23.

Ñaupas Paitán, H., Valdivia Dueñas, M. R., Palacios Vilela, J. J., & Romero Delgado, H. E. (2018). *Quantitative-Qualitative Research Methodology and Thesis Writing* (5th ed.). U Press.

Rivas, S. F., & Saiz, C. (2012). Validation and psychometric properties of the PENCRISAL critical thinking test. *Electronic Journal of Applied Methodology, 17*(1), 18–34.

Rivas, S. F., Franco, A., Vieira, R. M., Almeida, L. S., & Saiz, C. (2023). Assessing higher education students' critical thinking with the PENCRISAL test – Portuguese short version: A psychometric study. *Frontiers in Psychology*, (14), 1196794.

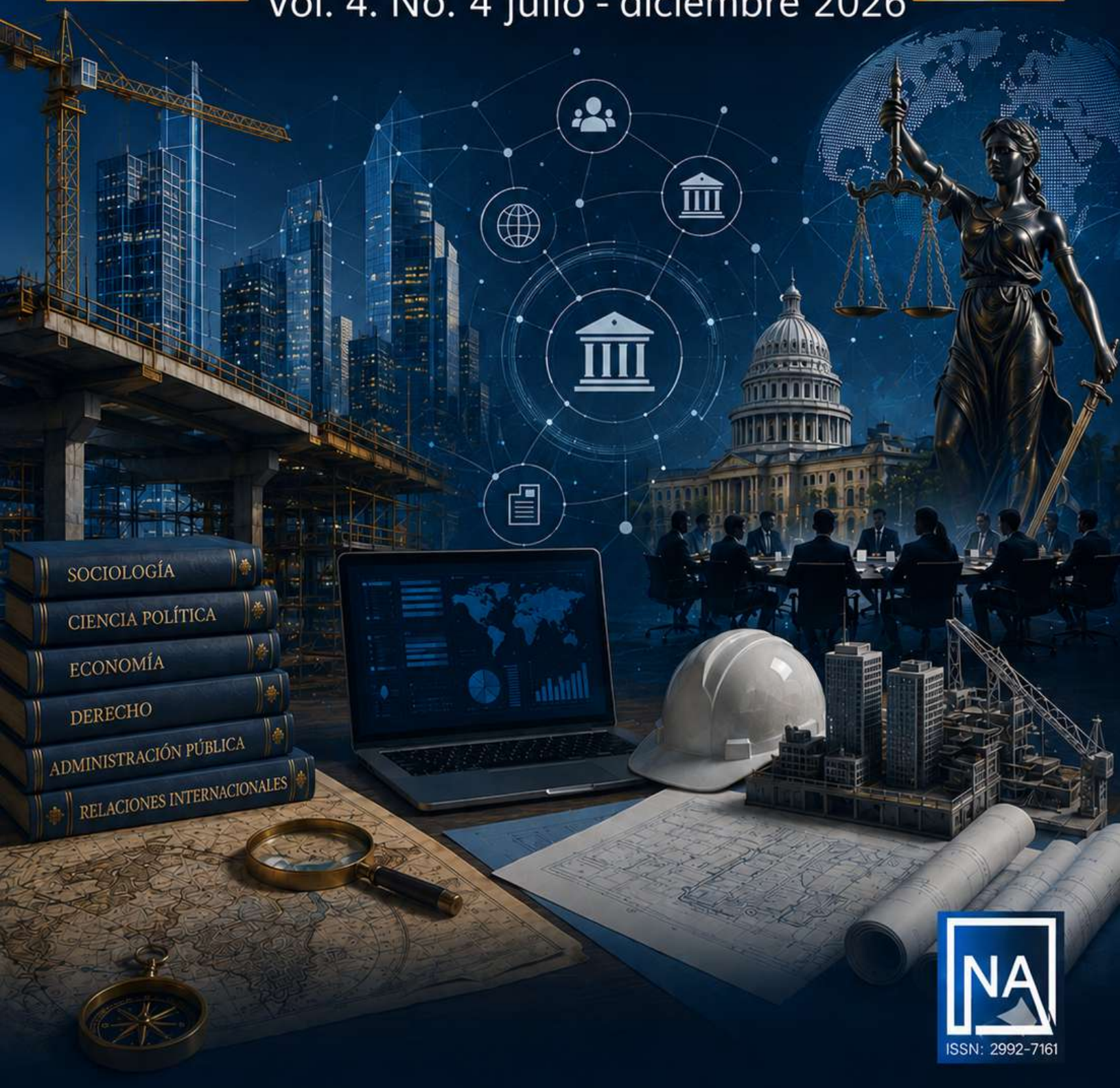
Collaborative Work Table

Role	Author(s)
Conceptualization	María Luisa Pereyra Hernández
Methodology	Ricardo García Parada, Gutiérrez Moreno Lilian Iveth
Software	Villar Laguna Víctor, López Martínez Brenda
Validation	María Luisa Pereyra Hernández
Formal Analysis	Ricardo García Parada, Gutiérrez Moreno Lilian Iveth
Research	Villar Laguna Víctor, López Martínez Brenda
Resources	María Luisa Pereyra Hernández
Data Curation	Ricardo García Parada, Gutiérrez Moreno Lilian Iveth
Writing - Preparation of the original draft	Villar Laguna Víctor, López Martínez Brenda
Writing - Review and editing	María Luisa Pereyra Hernández
Visualization	Ricardo García Parada, Gutiérrez Moreno Lilian Iveth
Supervision	Villar Laguna Víctor, López Martínez Brenda
Project Management	María Luisa Pereyra Hernández
Fundraising	Ricardo García Parada, Gutiérrez Moreno Lilian Iveth

REVISTA CIENTÍFICA MULTIDISCIPLINARIA

NEYART

Vol. 4. No. 4 julio - diciembre 2026



ISSN: 2992-7161