

# Tutoría entre pares y avalúo del aprendizaje: una experiencia institucional para mejorar el rendimiento matemático

## Peer Tutoring and Learning Assessment: An Institutional Experience to Improve Mathematical Performance

Héctor A. Aponte-Alequín<sup>1</sup>  , Óscar Y. Castrillón Velandia<sup>1</sup>  

### RESUMEN

**Introducción:** La deserción en cursos introductorios STEM exige apoyos académicos sostenidos en educación superior.

**Metodología:** Se realizó un estudio mixto cuasiexperimental, con muestreo no probabilístico por conveniencia y asignación por autoselección a tutoría. Participaron 140 estudiantes en dos grupos (con tutoría; sin tutoría) del curso MECU 5051, Métodos Cuantitativos para Administración de Empresas, en una universidad pública en Puerto Rico. Se aplicaron preprueba y posprueba con una rúbrica analítica de cinco criterios (0–20), cuestionarios de satisfacción y registros narrativos.

**Resultados:** En la posprueba, el grupo con tutoría obtuvo  $M = 12,84$  y el grupo sin tutoría  $M = 9,64$ . La prueba  $t$  no indicó diferencia significativa ( $t = -0,071$ ,  $p = ,944$ ;  $d \approx 0,01$ ).

**Discusión:** Aunque la evidencia inferencial fue débil, los perfiles por criterio y los testimonios estudiantiles sugirieron mejoras en comprensión conceptual, motivación y autorregulación. La autoselección introduce un posible sesgo de selección que limita la atribución causal.

**Conclusiones:** La articulación entre tutoría entre pares y avalúo del aprendizaje produjo evidencia útil para decisiones de retención y mejoramiento, y ofrece un modelo replicable en cursos de alta dificultad.

**Palabras clave:** Enseñanza superior, Enseñanza de las Matemáticas, Enseñanza mutua (peer teaching), Tutoría (educación), Evaluación del estudiante.

**Clasificación JEL:** I21; I25; I29.

**Received:** 08-08-2025

**Revised:** 30-10-2025

**Accepted:** 15-12-2025

**Published:** 02-01-2026

**Editor:** Alfredo Javier Pérez Gamboa 

<sup>1</sup>Universidad de Puerto Rico Recinto de Río Piedras. San Juan, Puerto Rico.

**Cite as:** Aponte-Alequín, H. A. y Castrillón Velandia, O. Y. (2026). Tutoría entre pares y avalúo del aprendizaje: una experiencia institucional para mejorar el rendimiento matemático. Región Científica, 5(1), 2026560. <https://doi.org/10.58763/rc2026560>



## INTRODUCTION

Higher education in Latin America faces the persistent challenge of reducing student dropout rates and improving academic performance, especially in Science, Technology, Engineering, and Mathematics (STEM) disciplines, which often include highly complex introductory courses. In this context, mastery of logical-mathematical reasoning is a critical area, as it influences students' academic progression, retention, and career opportunities. Several studies have indicated that introductory mathematics courses act as bottlenecks associated with both cognitive and institutional factors, disproportionately affecting first-generation students and those from vulnerable backgrounds (Dekker et al., 2023; González-Ortiz-de-Zárate et al., 2025; Swail et al., 2003).

In response to this problem, the Recinto de Río Piedras de la Universidad de Puerto Rico implemented the *Proyecto Universitario Estudia[N]til de Trayectoria al Éxito* (PUE[N]TE) during the second semester of the 2024-2025 academic year as a pilot institutional intervention aimed at student retention. The pilot focused on providing individualized peer tutoring to students enrolled in the course with the highest failure and dropout rates at the institution, MECU 3031, Quantitative Methods for Business Administration. Thus, in this precalculus course, the goal was to strengthen students' academic performance and reduce course repetition. Unlike conventional tutoring centers, the model was based on a sustained mentoring relationship and systematic monitoring of progress through learning assessment (Aponte-Alequín, 2020, 2024, 2025).

The implemented model was based on a formative conception of learning, in which the tutor's role is not limited to the technical resolution of exercises, but also acts as a mediator of metacognitive and self-regulatory processes. This structure aligns with the principles of Socratic tutoring and the Elaboration Likelihood Model of Persuasion, which highlights the peer's role as a credible source for fostering motivation, deep information processing, and changes in attitudes toward learning (Gao et al., 2024; Petty & Cacioppo, 1986a, 1986b; Rose et al., 2001). At the same time, academic support is integrated with persistence models that incorporate cognitive, social, and institutional dimensions to explain student success (Baek & Doleck, 2021; Banihashem et al., 2022; Swail et al., 2003).

From a learning assessment perspective, the project aligns with the trend of using direct performance evidence to document outcomes and guide decisions regarding curriculum improvement or support services. The intervention incorporated an analytical rubric aligned with the institutional learning domains defined by the División de Investigación Institucional y Avalúo (DIIA) of the Recinto, particularly logical-mathematical reasoning. This rubric allowed for the operationalization of specific competencies across five criteria and the evaluation of progress through a pre-test and a post-test, supplemented by satisfaction questionnaires and open-ended verbalizations. This approach responds to arguments about the importance of authentic, formative, and useful assessment practices for decision-making (Banihashem et al., 2022; Hardt et al., 2023; Hutchings et al., 2015; Jankowski et al., 2018; Mintz, 2020).

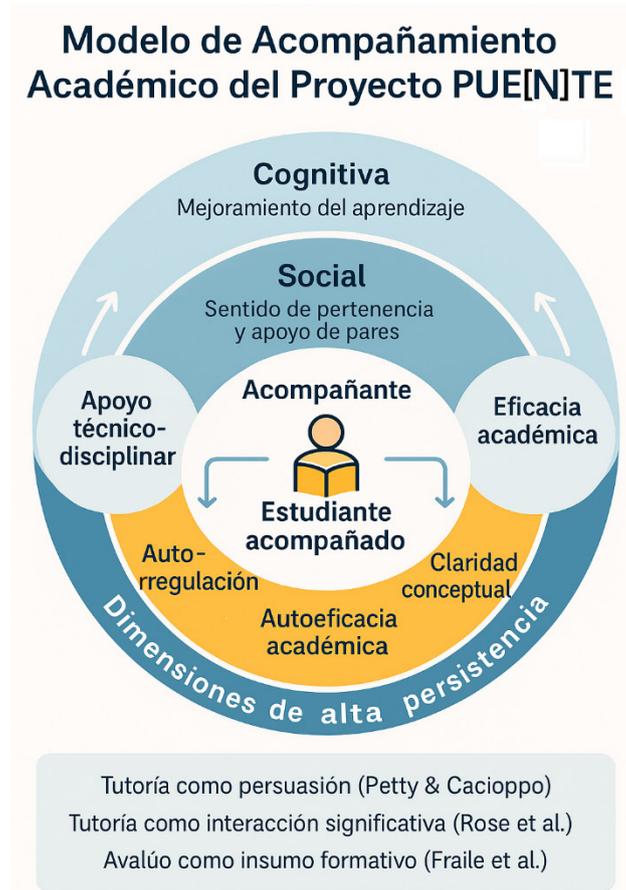
Within this framework, the purpose of this article is to analyze the effect of a peer tutoring model integrated into learning assessment on the logical-mathematical reasoning performance of students in a precalculus course at a Faculty of Business Administration. The study focuses on comparing the performance of students who received individualized support with those who did not, as well as exploring their perceptions of the tutoring experience. In this way, it seeks to provide empirical evidence on an institutional intervention strategy that links academic support, assessment, and retention management, with the potential to be replicated in other higher education contexts in the region.

At the institutional level, these challenges are often most evident in courses identified as "bottlenecks," as they concentrate high rates of failure, withdrawal, and repetition. This pattern not only affects performance but also impacts academic progression and retention, with cumulative effects on students' academic trajectories. In the current context, the literature on student success has reinforced the importance of interventions focused on high-demand and high-risk courses, particularly when the institution seeks to reduce gaps and maintain equity in student retention. These approaches converge with recent findings that highlight the value of evidence-based academic support strategies, paying attention to diverse student profiles and institutional conditions that facilitate or hinder persistence (Gao et al., 2024; Kuh et al., 2014; Oliva-Córdova et al., 2021; Tinto, 2022).

Contemporary discussions have also highlighted that the effectiveness of support services depends on their integration with the learning experience and their ability to promote sustained participation. Consequently, support models with intentional follow-up tend to surpass approaches focused solely on "access to resources" because they allow for monitoring progress, adjusting interventions, and strengthening study habits and self-regulation. In

courses with quantitative reasoning components, peer tutoring has shown particular relevance when combined with formative practices, timely feedback, and measurement tools that make performance visible beyond mere perception (Mullen et al., 2024). According to recent syntheses on interventions in higher education, combining structured academic support with monitoring and measurement strategies helps to define more precise explanations of what works, for whom, and under what conditions. This notion is key when designing pilot programs with the intention of replicability (What Works Clearinghouse, 2022).

**Figure 1.**  
*Academic Support Model of the PUE[N]TE Project*



**Note:** the figure appears in its original language

## METHODOLOGY

The research was framed within the pragmatic paradigm, with a mixed-methods approach that was predominantly quantitative, and was classified as an applied institutional research study with an explanatory-comparative scope. A quasi-experimental design with non-equivalent groups, using a pre-test and post-test, was employed to analyze the effect of a peer tutoring model on performance in logical-mathematical reasoning in a precalculus course. The population consisted of 140 students enrolled in the MECU 3031 course, Quantitative Methods for Business Administration, at a Business Administration School of a public university in Puerto Rico.

The PUE[N]TE project was implemented as a pilot academic intervention. Based on a pre-test administered in the first weeks of the semester, students at academic risk were identified and offered the opportunity to participate in the peer tutoring model. Two comparison groups were formed: students who received individualized peer tutoring and students who did not participate in the peer tutoring program. Each student in the mentored group was assigned a specific tutor, with weekly meetings scheduled. The intervention lasted approximately twelve weeks, from the third to the fourteenth week of the semester, and involved six mentors who supported 51 students within the total population.

Regarding materials and instruments, an analytical rubric was designed, aligned with the course criteria and the institutional framework for logical and mathematical reasoning. The rubric included five criteria: definition of

variables, algebraic formulation, matrix representation, interpretation of results, and final analysis, each evaluated on a scale of 1 to 4 points, for a maximum of 20 points per assessment. A pre-test and a post-test with tasks comparable in structure and level of difficulty were created using this rubric and administered at the beginning and end of the semester. In addition, student satisfaction questionnaires with closed-ended and open-ended questions were used, along with narrative logs from the mentors, in which they recorded observations about the mentoring process.

The validity of the pre- and post-test analytical rubric was supported by a validity argument centered on content and construct evidence, consistent with the use of assessment as an inquiry-based tool for improvement rather than a mere compliance exercise (Hutchings et al., 2015; Pan et al., 2024). First, content validity evidence was based on the collaborative design process by the course faculty and the expert review documented in the project: two specialists in mathematics education and one specialist in educational research. This interaction allowed for the refinement of descriptors, the leveling of expectations, and the assurance of disciplinary relevance. The five criteria operationalized, comprehensively and sequentially, the cognitive demands of the target performance in the domain of logical-mathematical reasoning: modeling of the verbal problem (definition of variables and formulation of the system), matrix representation (augmented matrix), procedural execution supported by technology (Gauss-Jordan on a graphing calculator) and interpretation of the result in the context of the problem (conclusion and analysis).

Second, the evidence for construct validity was based on the internal consistency between the institutional construct being measured and the rubric's progressive architecture, which distinguishes conceptual and procedural stages of quantitative reasoning and allows the total score to be interpreted as an indicator of integrated mastery. This design aligns with the recommendation to prioritize direct and authentic evidence of learning to support inferences useful for academic and institutional decision-making (Jankowski et al., 2018; Mullen et al., 2024).

To document the inter-rater reliability of the analytical rubric (five criteria; ordinal scale 1–4), scores from three evaluators were randomly selected independently from a subsample of 19 pieces of evidence. Before analysis, the scores were normalized to the 1–4 range. One case included at least one N/A value; therefore, the coefficients per criterion were estimated using 18 pieces of evidence with complete data. For the total score (sum of the five criteria), a two-way intraclass correlation coefficient with random effects and absolute agreement criterion was estimated. The results indicated moderate reliability for a single rater,  $ICC(2,1) = 0,629$ , 95 % bootstrap CI [0,290, 0,780], and high reliability for the average of three raters,  $ICC(2,3) = 0,836$ . For each criterion, quadratically weighted Cohen's kappa was calculated and averaged across the three pairs of raters. The highest levels of agreement were observed in "Solving by Gauss-Jordan on the calculator" and "Conclusion and analysis of the result" ( $\kappa_w = 0,629$  in both), while "Defining variables" ( $\kappa_w = 0,556$ ), "Formulating the system of equations" ( $\kappa_w = 0,537$ ), and "Constructing the augmented matrix" ( $\kappa_w = 0,507$ ) showed satisfactory agreement.

The research procedures included prior training of the tutors in mathematical content, academic support strategies, and the use of the rubric. During the semester, support was provided through problem-solving sessions, guided conceptual discussions, and activities focused on self-reflection and metacognition. The tutoring team maintained systematic communication with the course instructor to coordinate thematic emphases and monitor student progress.

For the quantitative analysis, measures of central tendency (averages) and dispersion (standard deviation and range) were calculated, both overall and by rubric criterion. These analyses were performed separately for the tutored group and the non-tutored group. To ensure consistency in the comparisons, only students who completed the pretest and posttest with scores greater than zero were included in the calculations. Based on this, averages expressed on a 20-point scale were used, a methodological decision aimed at facilitating the interpretation of the results in terms of overall achievement. Additionally, an independent samples t-test was applied to compare the overall posttest averages between the two groups.

The decision to prioritize performance-based indicators in the rubric stemmed from an evidence-based approach to learning assessment, which prioritizes observable learning outcomes over isolated perceptions. This perspective aligns with approaches that emphasize the need for meaningful and actionable indicators for academic and management decision-making.

In the qualitative dimension, open-ended responses from satisfaction questionnaires and anecdotal records from tutors were analyzed. These materials underwent an initial thematic coding process, which allowed for the identification of categories related to emotional support, clarity of explanation, mathematical confidence, and

individualized monitoring (Creswell & Creswell, 2018). The interpretation of these patterns relied on content analysis procedures and constant comparison between cases (Creswell & Poth, 2018). Triangulation of quantitative rubric results, attendance data, questionnaires, and narratives allowed for a deeper understanding of the impact of the mentoring on students' academic and attitudinal development (Denzin, 2017).

Finally, the pilot study was designed with replicability criteria in mind. The standardization of the rubric, the detailed description of the intervention, the documented training of the mentors, and the coordination with the faculty are inputs that facilitate the adaptation of the model to other introductory courses with similar challenges in mathematics education and student retention in higher education contexts.

## RESULTS AND DISCUSSION

The results of the PUE[N]TE pilot program reflected measurable improvements in the academic performance of students in the MECU 3031 course and allowed for an analysis of the potential of peer mentoring as an institutional support strategy. In the pre-test, the overall average score of the 140 students was 8,44 out of 20 points. This score increased to 10,89 in the post-test, representing an overall gain of 2,45 points in the assessed domain. This increase aligns with the literature that associates academic support interventions with progressive improvements in performance in highly complex courses in Mathematics and related areas (Dekker et al., 2023; González-Ortiz-de-Zárate et al., 2025).

**Table 1.**  
*Student participation in individualized tutoring*

Participation Category	Number of students	Percentage
Did not attend tutoring sessions	89	63,6 %
Attended 1 to 3 times	32	22,8 %
Attended 4 or more times	19	13,6 %
Total	140	100 %

The analysis by group revealed substantial differences between those who participated in the mentoring program and those who did not. The 51 students who attended individualized tutoring achieved an average score of 12,84 on the post-test, while the 89 students who did not receive mentoring scored 9,64. This difference of 3,2 points favored the mentoring group and reinforced the usefulness of the personalized mentoring model. Table 1 summarizes the levels of participation in the mentoring program, with three categories: students who did not attend, those who attended one to three times, and those who attended four or more times. The distribution showed a subgroup with sustained mentoring, which is consistent with the evidence linking the intensity of the intervention with academic improvements and persistence (Swail et al., 2003).

To analyze the results according to the rubric criteria, table 2 presents the baseline equivalence for each group, established based on the pre-test.

**Table 2.**  
*Baseline equivalence by group (total pre-test, 0-20)*

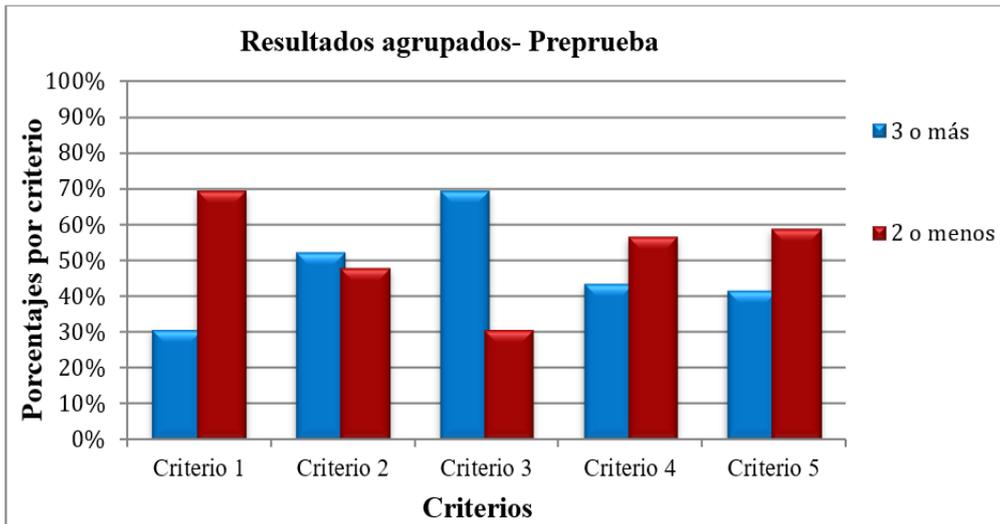
Group	n	Total pre-test (0-20), M	DE
With tutoring	42	15,02	4,64
Without tutoring	57	13,67	5,93

The breakdown by rubric criteria showed that the greatest progress occurred in the areas of conceptual interpretation and outcome analysis. In criterion 1 (definition of variables), the mentored group obtained an average of 2,81, compared to 2,19 in the unmentored group. In criterion 5 (final analysis of the problem), the averages were 2,75 for the mentored group and 2,04 for the unmentored group. These data suggest that mentoring fostered the development of logical and reflective thinking, beyond mere algorithmic repetition, and align with Socratic mentoring approaches that emphasize the active construction of knowledge through dialogue and questioning (Komakis, 2023; Rose et al., 2001).

On the other hand, criteria 2 and 3—related to algebraic formulation and matrix representation—showed the

lowest averages in both groups, even though the students who received support obtained slightly higher results. These findings point to the need to deliberately strengthen these dimensions in course planning and support strategies. The increasing difficulty of the post-test and the students' limited exposure to similar exercises in the practice lab were identified as factors that may have affected their familiarity with the content, which coincides with studies that document the importance of guided and structured practice in cognitively demanding tasks in mathematics (Fraile et al., 2023).

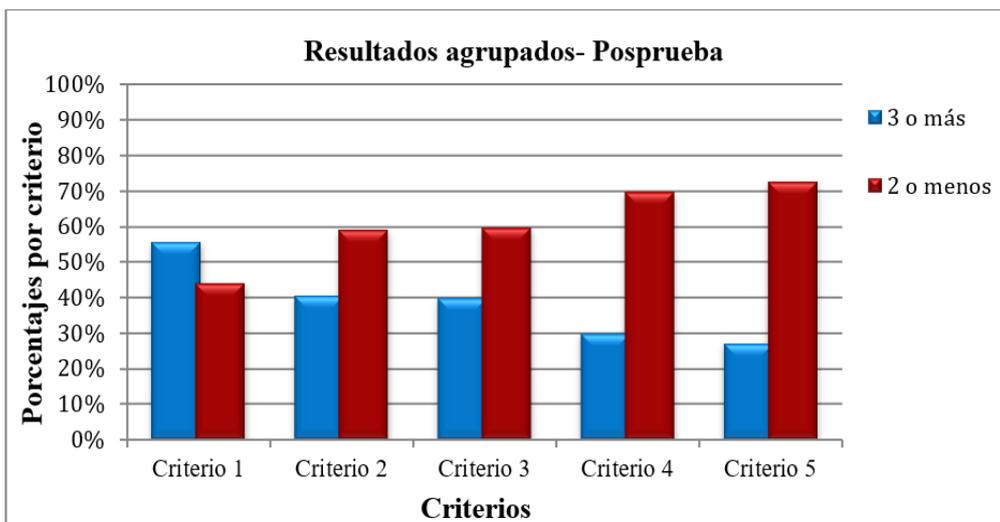
**Figure 2.**  
*Grouped results – Pre-test*



Note: the figure appears in its original language.

Figure 2 presents, for the pretest, the proportion of performances classified as “3 or more” versus “2 or less” for each criterion of the rubric. The pattern shows a heterogeneous profile, with one criterion having a higher proportion of “3 or more” and another having a higher proportion of “2 or less,” suggesting differentiated strengths and weaknesses from the outset.

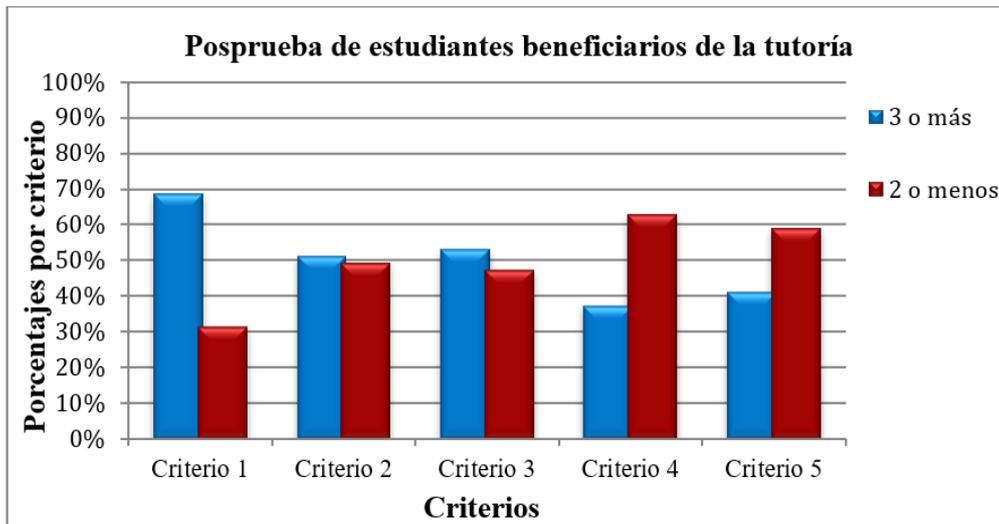
**Figure 3.**  
*Grouped results - Post-test*



Note: the figure appears in its original language.

Figure 3 presents, for the post-test, the proportion of “3 or more” and “2 or less” across the five criteria, using the same classification logic. The pattern shows that the post-test demanded more consistent performance at various stages of the process, which is reflected in the relative predominance of “2 or less” across several criteria.

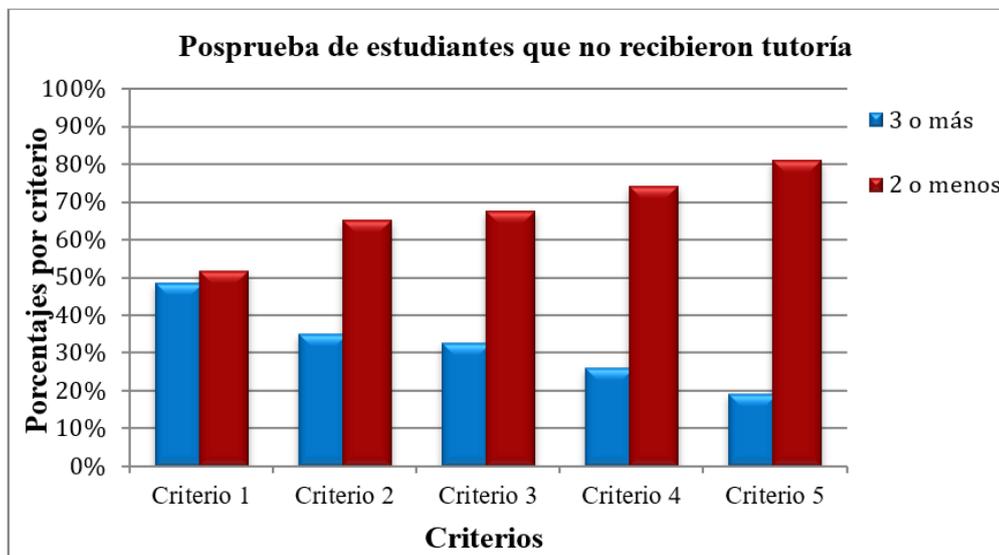
**Figure 4.**  
Post-test of students benefiting from tutoring



Note: the figure appears in its original language.

Figure 4 shows the distribution of “3 or more” versus “2 or less” in the post-test for students benefiting from peer tutoring. Overall, the profile of the mentored group exhibits a higher relative concentration of “3 or more” across several criteria, although areas with a predominance of “2 or less” persist, indicating areas that require reinforcement.

**Figure 5.**  
Post-test of students who did not receive tutoring



Note: the figure appears in its original language.

Figure 5 presents the distribution of “3 or more” and “2 or fewer” scores on the posttest for students who did not receive tutoring. The pattern tends to be concentrated in “2 or fewer” across the criteria, which contrasts with the profile of the tutored group and suggests a practical difference in the assessed domain.

Regarding the inferential analysis, the independent samples t-test applied to the overall posttest averages did not reveal a statistically significant difference between the groups,  $t(97) = 1,68$ ,  $p = ,096$ . On the posttest, the tutored group obtained  $M = 13,79$ ,  $SD = 5,29$  ( $n = 42$ ), while the untutored group obtained  $M = 11,89$ ,  $SD = 5,71$  ( $n = 57$ ); the effect size was small to moderate (Cohen’s  $d \approx 0,34$ ). This result requires careful interpretation. The sample size, variability between sections, and the quasi-experimental nature of the design are factors that can limit the

statistical power of the comparison. However, as Baek and Doleck (2021) point out, learning assessment should be based on concrete evidence of performance and not solely on statistical significance. In this study, the combination of descriptive analyses, results by rubric criterion, and qualitative data provided a more complete picture of the impact of the mentoring.

This evidence, however, does not invalidate the practical interpretation of the pilot study, because the study was conceived as an applied institutional intervention with a quasi-experimental design, unequal group sizes, and academic heterogeneity, which can reduce statistical power and increase variability. Within this framework, the practical relevance is supported by the descriptive magnitude of the differences, the profile by criteria, and the triangulation with trajectory indicators, which allows for inferences aimed at improvement even when the inferential comparison does not reach conventional significance. Consequently, the interpretation prioritizes the educational value of the observed patterns and their usefulness for institutional decision-making, rather than a dichotomous reading based solely on p-values.

Similarly, the estimated effects of the mentoring should be interpreted with caution, since the variability attributable to the evaluation in the inter-rater reliability coefficient may have attenuated or distorted real differences between groups. Even so, the good performance of the indices at the aggregate level supports the usefulness of the total score as an integrated indicator of logical-mathematical reasoning skills, while also identifying a priority area for calibration in future implementations. This calibration should involve brief calibration sessions, response anchors, and consensus-based examples for the criterion with the least agreement.

The discussion also requires acknowledging plausible biases associated with the implementation. Group assignments were not randomized, and participation in tutoring was based on student invitation and decision, introducing selection bias: those who accepted mentoring may have differed in motivation, time availability, external support, or predisposition to use institutional resources. This bias can inflate or attenuate the observed differences and limit the causal scope of the conclusions. For this reason, the findings are interpreted as preliminary evidence of effectiveness in a real-world context, with a call to strengthen the design in future replications through strategies such as pretest matching, models with covariates, and dose-response analysis based on the number of sessions, as well as rubric calibration criteria to reduce criterion-related measurement error.

The benefits of the PUE[N]TE model were also reflected in attitudinal and academic trajectory dimensions. None of the students who received tutoring received a grade of F, in contrast to the unaccompanied group, which accounted for the majority of failures. Furthermore, improvements were documented in attitude toward the course, participation in sessions, and confidence in tackling mathematical tasks—aspects that have been identified as key variables in student persistence models (Swail et al., 2003). These observations suggest that peer tutoring functioned as a protective mechanism against dropout and repetition, consistent with the institutional goal of strengthening retention.

The qualitative analysis of open-ended responses in the satisfaction questionnaires and tutor logs allowed for the identification of four main categories: emotional support, clarity of explanation, mathematical confidence, and individualized monitoring. Table 3 presents examples of verbalizations for each category. In the emotional support dimension, students highlighted a highly cooperative environment, the tutors' constant willingness to help, and a welcoming attitude. These elements are linked to the creation of socio-affective climates favorable to learning and to the sense of institutional belonging described by Swail et al. (2003).

**Table 3.**  
*Categories identified and selection of verbalizations*

Category	Emotional support	Clarity in the explanation	Mathematical confidence	Individualized monitoring
Verbalizations	<ul style="list-style-type: none"> <li>• «They were always willing to help».</li> <li>• «Friendly atmosphere and very cooperative»</li> <li>• «They would quickly approach me to ask what I needed help with, and they always did so with a smile».</li> </ul>	<ul style="list-style-type: none"> <li>• « They explained the problems in detail ».</li> <li>• « to understand the exercises and class topics more quickly and effectively »</li> <li>• « the explanation only to me »</li> </ul>	<ul style="list-style-type: none"> <li>• «They prepared me for my final exam».</li> <li>• «They helped me understand and remember older topics».</li> <li>• «Sometimes I didn't know which formulas to use or how to start, but with the tutors' help and the practice they gave me, everything became clearer. Gradually, I got the hang of it and was able to do the exercises faster».</li> </ul>	<ul style="list-style-type: none"> <li>• «The tutors were willing to explain things until you understood».</li> <li>• «I was able to clarify my doubts; they explained a lot to me, and I received individualized help».</li> <li>• «My tutor's availability, commitment, and help in understanding were invaluable»</li> </ul>

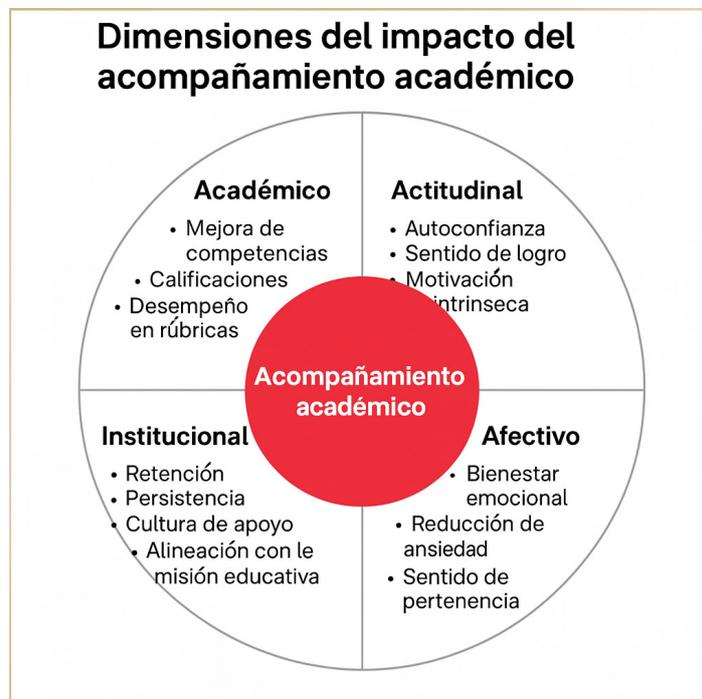
In the clarity of explanation category, the students highlighted the opportunity to receive detailed and personalized explanations, which facilitated a “faster and more effective” understanding of the exercises. This experience relates to the tutor’s role as a cognitive mediator who breaks down complex tasks and directs the student’s attention toward the key steps in solving them, in line with guided and self-regulated learning approaches (Fraile et al., 2023). In the mathematical confidence dimension, the students reported feeling better prepared for exams and recalling past topics, which points to a positive reconstruction of their academic self-efficacy.

The individualized monitoring category reflected the perception of close and systematic follow-up: the students valued the opportunity to clarify doubts, receive help “until they understood,” and have the constant availability of their tutors. This experience aligns with the concept of academic support as a continuous and personalized process that integrates cognitive support and strategic guidance in the use of resources, as conceptualized by Dekker et al. (2023), González-Ortiz-de-Zárate et al. (2025), and Pérez-Burriel et al. (2024).

From an institutional perspective, the PUE[N]TE model generated relevant learning about the use of assessment as a tool for pedagogical transformation. The analytical rubric allowed for monitoring individual and group progress, while integrating the results into the institution’s assessment management system strengthened the data culture and informed decision-making. The future standardization of this rubric with the formative scales of the OLAS system, with levels that distinguish between students “in progress” and “beginning,” is emerging as a key step in aligning learning assessment with a logic of continuous improvement and formative feedback (Hutchings et al., 2015; Jankowski et al., 2018).

Figure 6 summarizes the dimensions of the impact of academic support identified in the study: cognitive (improvement in performance and in interpretation and analysis skills), affective (increased motivation, confidence, and perception of support), and institutional (strengthened retention, strategic use of assessment, and data culture). This triad aligns with the proposal of Swail et al. (2003), who argue that student success depends on the interaction between cognitive, social, and institutional factors, and position peer tutoring as an integrating axis of these dimensions.

**Figure 6.**  
*Dimensions of the impact of academic mentoring*



**Note:** the figure appears in its original language.

The pilot program demonstrated that academic support should not be viewed as a peripheral remedial measure, but rather as a structural academic policy strategy. The individualized tutoring model, centered on peer-to-peer trust and the development of metacognitive skills, is projected as a key component of courses related to the domains of logical-mathematical reasoning, effective communication, and critical thinking. The initial investment of \$30 997,15 allowed

for verification of the proposal's viability and effectiveness. In future iterations, costs can be reduced by reallocating existing funds (such as salaries and funds from online programs) and utilizing internal faculty compensation. Thus, the model emerges as a sustainable and replicable alternative for expanding access to student-centered support strategies in higher education contexts in the region.

## CONCLUSIONS

The findings of the PUE[N]TE project showed that peer tutoring, when conceived as personalized and sustained support, had a positive effect on the academic performance of students in a highly challenging mathematics course. The difference in post-test averages between tutored and unaccompanied students, as well as the absence of failing grades among those who participated in the tutoring program, pointed to a concrete impact on learning achievement and course retention. Qualitative analysis reinforced this interpretation by demonstrating improvements in perceived clarity, mathematical confidence, and self-regulated study skills.

The model developed differed from traditional tutoring approaches by assigning a specific mentor to each student from the first weeks of the semester. This ongoing relationship allowed for the simultaneous addressing of cognitive difficulties and emotional barriers, resulting in a holistic intervention. The support was not limited to “explaining exercises” but also promoted the development of strategies for tackling the subject, the organization of study time, and the more strategic use of available institutional resources.

From an institutional perspective, the project demonstrated that investing in peer tutoring can be considered a strategic decision with high potential for academic return. With a moderate initial investment, it was possible to reduce the failure rate in a critical course, strengthen fundamental quantitative skills, and foster a culture of active and collaborative learning. In light of these results, it makes sense for faculties to explore the redistribution of existing budget items—such as stipends and funds generated by online programs—to support these types of initiatives without relying exclusively on extraordinary allocations.

The results also highlighted the value of learning assessment as a central element of the project. The use of an analytical rubric allowed for the documentation of specific progress in criteria such as the interpretation of results and the final analysis of the problem, while the integration of the data into institutional monitoring systems opened up possibilities for the continuous improvement of the course. The experience revealed the need to standardize rubrics at institutional levels for formative purposes and to strengthen the use of learning data to inform the design of tasks, activities, and support strategies.

Like any pilot study, PUE[N]TE presented limitations that warrant critical reflection. The quasi-experimental design, the inability to randomly assign students, and the group sizes introduced restrictions when interpreting statistical significance. However, the triangulation of quantitative results, attendance data, and qualitative evidence allowed for a more nuanced understanding of the impact of the support. This balance between methodological rigor and the recognition of consistent patterns constitutes an important input for academic decision-making.

Regarding future trends, the study opens several lines of research. It is pertinent to delve deeper into the analysis of the differential effect of various tutoring styles—for example, a more directive approach versus a more Socratic one—and to explore which combinations of strategies are most effective for different student profiles. Longitudinal follow-up studies are also needed to examine the impact of this support on subsequent courses, institutional retention, and graduation rates. Thus, experience suggests the potential to transfer this model to other introductory courses with high dropout rates in STEM disciplines and in diverse regional contexts, adapting the design to the specific characteristics of each institution and higher education system.

## REFERENCES

- Aponte-Alequín, H. A. (2020). *Informe de acompañamiento académico: Resultados preliminares en los dominios de comunicación efectiva y pensamiento crítico*. Facultad de Comunicación e Información, Universidad de Puerto Rico, Recinto de Río Piedras. <https://drive.google.com/file/d/1SNZi752YOYfnr9e1OgRtZbdkfGsupnm6/view>
- Aponte-Alequín, H. A. (2024). *Evaluación preliminar del avalúo del aprendizaje en el Recinto de Río Piedras 2024–2025 (EPAAP)*. División de Investigación Institucional y Avalúo, Universidad de Puerto Rico, Recinto de Río Piedras. [https://academicos.uprrp.edu/diia/wp-content/uploads/sites/5/2025/05/INFORME\\_EvalPrelim-AvaluoAprendizajeEst-2024-25\\_Rev-2024-12.pdf](https://academicos.uprrp.edu/diia/wp-content/uploads/sites/5/2025/05/INFORME_EvalPrelim-AvaluoAprendizajeEst-2024-25_Rev-2024-12.pdf)

- Aponte-Alequín, H. A. (2025). Enhancing the writing process: Integrating applied linguistics learning assessment in the classroom. *European Scientific Journal*, 21(5), 35. <https://doi.org/10.19044/esj.2025.v21n5p35>
- Baek, C., & Doleck, T. (2021). Educational data mining versus learning analytics: A review of publications from 2015 to 2019. *Interactive Learning Environments*, 31(6), 3828–3850. <https://doi.org/10.1080/10494820.2021.1943689>
- Banihashem, S. K., Noroozi, O., van Ginkel, S., Macfadyen, L. P., & Biemans, H. J. A. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review*, 37, 100489. <https://doi.org/10.1016/j.edurev.2022.100489>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.
- Dekker, I., Luberti, M., & Stam, J. (2023). Effects of supplemental instruction on grades, mental well-being, and belonging: A field experiment. *Learning and Instruction*, 87, 101805. <https://doi.org/10.1016/j.learninstruc.2023.101805>
- Denzin, N. K. (2017). *The research act: A theoretical introduction to sociological methods*. Routledge. <https://doi.org/10.4324/9781315134543>
- Fraile, J., Gil-Izquierdo, M., & Medina-Moral, E. (2023). The impact of rubrics and scripts on self-regulation, self-efficacy and performance in collaborative problem-solving tasks. *Assessment & Evaluation in Higher Education*, 48(8), 1223–1239. <https://doi.org/10.1080/02602938.2023.2236335>
- Gao, X., Noroozi, O., Gulikers, J., Biemans, H. J. A., & Banihashem, S. K. (2024). A systematic review of the key components of online peer feedback practices in higher education. *Educational Research Review*, 42, 100588. <https://doi.org/10.1016/j.edurev.2023.100588>
- González-Ortiz-de-Zárate, A., Alonso-García, M. A., Gómez-Flechoso, M. de los Á., & Aliagas, I. (2025). Peer mentoring, university dropout and academic performance before, during, and after the pandemic in Spain. *Evaluation and Program Planning*, 113, 102676. <https://doi.org/10.1016/j.evalprogplan.2025.102676>
- Hardt, D., Nagler, M., & Rincke, J. (2023). Tutoring in (online) higher education: Experimental evidence. *Economics of Education Review*, 92, 102350. <https://doi.org/10.1016/j.econedurev.2022.102350>
- Hutchings, P., Kinzie, J., & Kuh, G. D. (2015). Evidence of student learning: What counts and what matters for improvement. En *Using evidence of student learning to improve higher education* (pp. 27–47). Jossey-Bass.
- Jankowski, N. A., Timmer, J. D., Kinzie, J., & Kuh, G. D. (2018). *Assessment that matters: Trending toward practices that document authentic student learning*. National Institute for Learning Outcomes Assessment. <https://files.eric.ed.gov/fulltext/ED590514.pdf>
- μ, G. (2023). The maieutic method combined with dialectics as the highest pedagogical principle according to Plato. *Education Sciences*, 118–136. <https://doi.org/10.26248/edusci.v2023i1.1645>
- Kuh, G. D., Ikenberry, S. O., Jankowski, N. A., Cain, T. R., Ewell, P. T., Hutchings, P., & Kinzie, J. (2014). *Knowing what students know and can do: The current state of student learning outcomes assessment in U.S. colleges and universities*. National Institute for Learning Outcomes Assessment. [https://www.minotstateu.edu/Academic/\\_documents/assessment/Knowing-What-Students-Know-and-Can-Do.pdf](https://www.minotstateu.edu/Academic/_documents/assessment/Knowing-What-Students-Know-and-Can-Do.pdf)
- Mintz, S. (2020). New approaches to assessing institutional effectiveness: How to ensure that institutions improve instructional quality and effectiveness and enhance equity and academic postgraduation outcomes. *Inside Higher Ed*. <https://www.insidehighered.com/blogs/higher-ed-gamma/new-approaches-assessing-institutional-effectiveness>

- Mullen, C., Howard, E., & Cronin, A. (2024). A scoping literature review of research on the evaluation of higher education mathematics and statistics support. *Educational Studies in Mathematics*, *117*, 1–22. <https://doi.org/10.1007/s10649-024-10332-6>
- Oliva-Córdova, L. M., García-Cabot, A., & Amado-Salvatierra, H. R. (2021). Learning analytics to support teaching skills: A systematic literature review. *IEEE Access*, *9*, 58351–58363. <https://doi.org/10.1109/ACCESS.2021.3070294>
- Pan, Z., Biegley, L., Taylor, A., & Zheng, H. (2024). A systematic review of learning analytics-incorporated instructional interventions on learning management systems. *Journal of Learning Analytics*, *11*(2), 52–72. <https://doi.org/10.18608/jla.2023.8093>
- Pérez-Burriel, M., Serra, L., & Fernández-Peña, R. (2024). How can we ensure limited individual tutorial time is reflective? The reflective individual tutoring model for higher education. *Reflective Practice*, *25*(4), 467–483. <https://doi.org/10.1080/14623943.2024.2325413>
- Petty, R. E., & Cacioppo, J. T. (1986a). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer. <https://doi.org/10.1007/978-1-4612-4964-1>
- Petty, R. E., & Cacioppo, J. T. (1986b). The elaboration likelihood model of persuasion. En *Advances in Experimental Social Psychology* (Vol. 19, pp. 123–205). Elsevier. [https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)
- Rose, C. P., Moore, J. D., VanLehn, K., & Allbritton, D. (2001). A comparative evaluation of Socratic versus didactic tutoring. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *23*. <https://escholarship.org/uc/item/98j4479r>
- Swail, W. S., Redd, K. E., & Perna, L. W. (2003). *Retaining minority students in higher education: A framework for success*. Jossey-Bass. <https://doi.org/10.1002/aehe.3002>
- Tinto, V. (2022). Exploring the character of student persistence in higher education: The impact of perception, motivation, and engagement. En A. L. Reschly & S. L. Christenson (Eds.), *Handbook of research on student engagement* (pp. 357–379). Springer. [https://doi.org/10.1007/978-3-031-07853-8\\_17](https://doi.org/10.1007/978-3-031-07853-8_17)
- What Works Clearinghouse. (2022). *What works clearinghouse procedures and standards handbook* (Version 5.0). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. <https://ies.ed.gov/ncee/wwc/Handbooks>

## FINANCING

For the research described in this paper, \$30 997,15 was received from the Junta de Gobierno de la Universidad de Puerto Rico.

## CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

## ACKNOWLEDGMENTS

We thank the Junta de Gobierno de la Universidad de Puerto Rico, and, of the Recinto de Río Piedras, the College of Business Administration and the División de Investigación Institucional y Avalúo. We also extend our special thanks to Dr. Mayra Chárriez Cordero, who at the time of data collection for this research was serving as Vice President for Student Affairs, without whose support this project would not have been possible. Likewise, we acknowledge the tutors for their dedication, and the course instructors for their commitment to a culture of assessment and accountability.

## AUTHORSHIP CONTRIBUTION

Conceptualization: Héctor A. Aponte-Alequín.

Data Curation: Oscar Y. Castrillón Velandia.

Formal Analysis: Héctor A. Aponte-Alequín.

Fundraising: Héctor A. Aponte-Alequín.

Research: Héctor A. Aponte-Alequín, Oscar Y. Castrillón Velandia.

Methodology: Héctor A. Aponte-Alequín, Oscar Y. Castrillón Velandia.

Project Management: Oscar Y. Castrillón Velandia.

Resources: Oscar Y. Castrillón Velandia.

Software: Héctor A. Aponte-Alequín.

Supervision: Héctor A. Aponte-Alequín.

Validation: Héctor A. Aponte-Alequín.

Visualization: Héctor A. Aponte-Alequín.

Writing – Original Draft: Héctor A. Aponte-Alequín.

Writing – proofreading and editing: Héctor A. Aponte-Alequín, Oscar Y. Castrillón Velandia.