MODESTUM

EURASIA Journal of Mathematics, Science and Technology Education, 2026, 22(9), em2901
ISSN:1305-8223 (online)

OPEN ACCESS

Research Paper

https://doi.org/10.29333/ejmste/19125

Effects of a flipped learning – Project-based learning supported by GeoGebra on calculus achievement in engineering

Hernan Javier Lara-Saavedra [1*]

[1] Universidad Técnica Particular de Loja, Loja, ECUADOR

Received 02 April 2026 ▪ Accepted 31 July 2026

Abstract

Pre-class learning resources, collaborative classroom activities, project-based applications, and dynamic visualization afforded by GeoGebra collectively promote deeper engagement. Using descriptive statistics, normality assessment, and robust inferential procedures appropriate for non-normal distributions, Test #1 assessed performance on five differential calculus tasks. Of the 151 participants, 70.9% were assigned to the control group and 29.1% to the experimental group. Test #2 was designed across three item formats. Of the 143 participants, 69.2% were assigned to the control group and 30.8% to the experimental group. Levene's test indicated that the homogeneity-of-variance assumption was not met, F (1, 149) = 24.49, p < .001. U statistic reduction from 1,071.0 (test #1) to 815.5 (test #2) reflects a more pronounced rank-order separation between groups on the second assessment. Aiming cognitive hierarchical level 3, students were encouraged to move beyond algorithm competency toward conversion from one semiotic representation system to another.

Keywords: flipped learning, project-based learning, GeoGebra, undergraduate calculus, semiotic representation

INTRODUCTION

Mathematical software such as GeoGebra enables students to identify patterns and make generalizations associated with those patterns (Addae, 2025; Aytekin & Kiymaz, 2019; Campos Nava et al., 2021; Córdova-Esparza & Jimenez Piñon, 2025; Gallo et al., 2019; Havelková, 2013; Stewart et al., 2019; Trigueros & Wawro, 2020; Ziatdinov & Valles, 2022). Students frequently encounter significant difficulties with abstraction, spatial visualization, and the transfer of mathematical concepts to applied problem contexts (Quintilla Castán & Fernández Morales, 2021; Scott et al., 2016; Ting et al., 2019).

While flipped learning (FL) has demonstrated positive effects in mathematics classrooms (Davies et al., 2013; Hew & Lo, 2018; Lai & Hwang, 2016; Paiva et al., 2025; Pinnelli & Fiorucci, 2015; Sohrabi & Iraj, 2016), its implementation in advanced mathematical courses at undergraduate level remains insufficiently studied due to the need to operationalize key variables such as performance, motivation, and cognitive development using validated instruments. Similarly, project-based learning (PBL) has been shown to deepen mathematical reasoning and foster collaborative problem-solving (Cirillo et al., 2022; Paleenud et al., 2023; Rahman et al., 2016; Sholahuddin et al., 2023), yet its integration with FL in multi-course mathematics programs has received little systematic attention.

In Ecuador, the predominance of lecture-based instruction in science, technology, engineering, and mathematics (STEM) courses reflects both deeply ingrained pedagogical traditions and systemic constraints on access to educational technology (Aguiar & Calabrese, 2025; Dancy et al., 2024; Nagamalla et al., 2025; Nguyen et al., 2021). International literature on active learning has been generated predominantly in North American and European institutional settings (Allen & Tanner, 2005; Armbruster et al., 2009; Doolittle et al., 2022; Dzaiy & Abdullah, 2024; Klein et al., 2022; Saad et al., 2025; Sadia et al., 2025; Steen-Utheim & Foldnes, 2018; Walker et al., 2008). This transferability concern is compounded by underrepresentation of

✉ hjlara@utpl.edu.ec **(*Correspondence)**

Contribution to the literature

- This study provides empirical evidence from a quasi-experimental study in undergraduate engineering calculus, comparing students' achievement across two assessment formats using robust statistical analyses.
- A post hoc power analysis was conducted based on the observed effect sizes obtained from the Welch's t tests. For Test #1 (n = 151), the observed effect size was large (Cohen's *d* = 1.26), yielding an achieved statistical power of 0.9999 (α = .05, two-tailed). Likewise, for Test #2 (n = 143), the observed effect size was also large (Cohen's d = 1.28), with an achieved statistical power greater than 0.9999.
- It uses descriptive statistics, normality assessment, and robust inferential procedures appropriate for non-normal distributions and performance on each test subscale.

developing-country contexts (Idsardi, 2019; Indorf et al., 2021; Ting et al., 2022).

The study examined the development of algorithmic thinking and generalization skills through the use of dynamic mathematics software GeoGebra. Research design involved 15 secondary school students in the Netherlands and employed a qualitative approach based on the analysis of students' workbooks, semi-structured interviews, and teachers' reflective logbooks (van Borkulo et al., 2021). Findings indicated positive perceptions from both teachers and students regarding the software's contribution to mathematical learning. However, participants encountered several challenges, particularly in becoming familiar with the GeoGebra interface, understanding its syntax, and using commands effectively. These difficulties were closely associated with the teaching and learning process, highlighting the importance of adequate instructional support during software integration. Based on these findings, authors recommend incorporating pedagogical strategies that explicitly foster computational thinking within mathematics instruction, complementing the development of algorithmic thinking and enhancing students' problem-solving abilities.

Mathematical competence was examined across four levels of cognitive demand through students' engagement with the GeoGebra environment (Williner, 2024). Using a teaching experiment conducted in an undergraduate mathematical analysis course, researchers analyzed how varying levels of task complexity influenced competency development. Results indicated that tasks characterized by moderate cognitive demand and scaffolded problem-solving effectively promoted the acquisition of specific mathematical competencies. Moreover, students' interaction with GeoGebra was positively associated with enhanced conceptual understanding and mathematical reasoning, underscoring the pedagogical value of dynamic mathematics software in supporting higher-order cognitive processes.

To examine GeoGebra effectiveness in integral calculus instruction, researchers administered a 20-item written assessment to 52 students assigned to experimental and control groups (Ortega et al., 2024). Results demonstrated that students exposed to GeoGebra achieved significantly better learning outcomes than those receiving traditional instruction (Dietrich et al., 2022; von Korff et al., 2016). Authors concluded that GeoGebra constitutes an effective pedagogical resource for creating more interactive and student-centered learning environments, thereby fostering deeper conceptual understanding and improving problem-solving performance in integral calculus.

The theory of semiotic representation registers, proposed by philosopher and psychologist Raymond Duval, provides a foundational framework for understanding mathematical cognition (Lizana Chauca & Antezana Iparraguirre, 2021). Inherently abstract mathematical objects (e.g., numbers, functions, and geometric figures) cannot be directly perceived. Their understanding and communication depend on its use across multiple systems of semiotic representation. For instance, meaningful mathematical learning requires students to translate between different representations, such as converting an algebraic equation into its corresponding graphical representation or interpreting a graph to derive its symbolic expression.

Grounded in this theoretical perspective, the present study integrates FL and PBL with GeoGebra to address students' persistent difficulties in connecting algebraic and graphical representations of quadratic and transcendental functions. A common challenge observed in calculus courses is students' inability to derive the parabola equation from its graph using the vertex and intercepts. Subsequently, they apply either the definition of the derivative or standard differentiation rules. These conceptual difficulties are often compounded by weaknesses in prerequisite algebraic skills, including operations with fractions, least common multiples, simplification of algebraic expressions, application of the distributive property, and manipulation of exponential expressions.

GeoGebra was incorporated as a visualization and exploration tool to facilitate the transition between symbolic and graphical representations. Through commands related to tangent lines, derivatives, extrema, and concavity, students were encouraged to explore the graphical behavior of functions while simultaneously interpreting their algebraic properties. Although both the desktop and Android App versions of GeoGebra were introduced, students demonstrated a clear preference for the mobile application because of its accessibility and ease of use during independent study.

Consistent with the FL model, students received instructional materials before each class throughout the 16-week semester. These resources included PDF lecture notes formatted with four slides per page, recorded video lectures, audio explanations, self-assessment quizzes, flashcards, mind maps, and concise study summaries. Face-to-face sessions were then devoted to collaborative problem solving, mathematical discussions, and guided exploration using GeoGebra.

As the culminating PBL activity, students prepared a brief technical report applying differential and integral calculus to a discipline-specific problem selected from a database of authentic case studies. Engineering-oriented projects focused primarily on optimization problems involving derivatives, whereas applications of definite integrals emphasized volume calculations and other contexts.

The study pursued three research objectives:

(1) to examine students' performance using descriptive statistical analysis (Kozanitis & Nenciovici, 2022);

(2) to evaluate score distributions through normality assessment and appropriate non-parametric statistical procedures (Baig & Yadegaridehkordi, 2023; Freeman et al., 2014; Hao, 2016; Prince, 2004;); and

(3) to compare students' performance on test #1 and test #2 across different item-format subscales in order to determine the educational effectiveness of the proposed FL – PBL instructional approach supported by GeoGebra (Darwazeh & Branch, 2015).

MATERIALS AND METHODS

This study adopted an exploratory, descriptive, and correlational design to examine the effect of a hybrid pedagogical model combining FL and PBL on the development of STEM competencies among university students.

Test #1 and Test #2 Evaluation: Control and Experimental Group Randomization

Study population

The intervention consisted of an FL combined with PBL (FL+PBL) hybrid instructional model implemented in the univariate mathematical analysis (differential and integral calculus) course during the April-August 2025 and April-August 2026 academic term in Loja, Ecuador. At Universidad Técnica Particular de Loja, Univariate mathematical analysis comprises differential calculus until midterm exams and integral calculus until final exams in one academic cycle.

Test #1 assessed performance on five conceptual and procedural differential calculus tasks: Q1, graphical interpretation of functions; Q2, definition of the derivative; Q3, power rule; Q4, chain rule; Q5, identification of critical points. Of the 151 participants included in this analysis, 107 (70.9%) were assigned to the control group and 44 (29.1%) to the experimental group.

Test #2 was designed across three formats:

(a) multiple choice (MC),

(b) true or false (TF),

(c) fill-in-the-blanks.

Of the 143 participants included in this analysis, 99 (69.2%) were assigned to the control group and 44 (30.8%) to the experimental group.

Intervention sample comprised undergraduate students enrolled across two engineering programs: food and agricultural engineering. Control sample comprised students enrolled across biology, agricultural engineering, industrial engineering, and computer science.

During the instructional cycle, all students across all programs additionally attended weekly out-of-class tutoring sessions facilitated by a seventh-term scholarship student, who contextualized differential calculus.

Data collection instruments

Two purpose-built assessments were administered to measure student performance before and after the FL+PBL intervention using the revised Bloom's taxonomy (Darwazeh & Branch, 2015).

Test #1 consisted of five open-ended questions targeting higher cognitive levels of the revised Bloom's taxonomy (apply-evaluate). Test items addressed the following mathematical competencies:

(1) graphical interpretation of derivatives,

(2) differentiation of algebraic and transcendental functions,

(3) tangent line equation at a given point,

(4) chain rule, and

(5) critical points and extrema.

Each item required students to produce extended written solutions. No technological aids were permitted.

Test #2 comprised 13 items targeting the lower-to-intermediate cognitive levels of the revised Bloom's taxonomy (remember-apply) and was structured across three item formats:

Table 1. Descriptive statistics of test #1 item scores by group

Variable	Group	M	SD	Q1	Q3	Skewness	Kurtosis
Q1	Control	7.41	16.75	0.0	2.50	2.945	9.695
	Experimental	13.64	14.72	0.0	25.00	0.497	-0.778
Q2	Control	5.50	13.36	0.0	0.00	2.882	8.23
	Experimental	11.93	23.18	0.0	6.25	1.860	2.949
Q3	Control	41.93	35.49	5.0	66.50	0.110	-1.532
	Experimental	58.81	43.35	12.5	100.00	-0.247	-1.778
Q4	Control	25.52	36.08	0.0	50.00	1.016	-0.580
	Experimental	42.61	39.56	12.5	100.00	0.587	-1.388
Q5	Control	5.75	13.27	0.0	0.00	2.803	8.391
	Experimental	63.35	39.01	25.0	100.00	-0.340	-1.579

(a) five MC items (Q1-Q5) assessing recognition and procedural knowledge;

(b) four true/false items (Q6-Q9) evaluating conceptual understanding; and

(c) four completion items (Q10-Q13) requiring differentiation results.

Formal mathematical content mandated by the Ecuadorian secondary school curriculum for bachillerato general unificado, explicitly requires students to develop three differential calculus competencies prior to university entry. Specifically, the *currículo priorizado con enfasis en competencias comunicacionales, matemáticas, digitales y socioemocionales – nivel de bachillerato general* (Ministerio de Educación, Deporte y Cultura del Ecuador, 2025) mandates theses performance standards: skill M.5.1.33 requires students *to calculate intuitively the derivative of quadratic functions from the incremental quotient*; skill M.5.1.35 requires students *to interpret geometrically and physically the first derivative – including the slope of the tangent line and instantaneous velocity – of quadratic functions with the support of digital technologies*; and skill M.5.1.36 *requires students to interpret physically the second derivative – including mean and instantaneous acceleration – of quadratic functions using graphing calculators, software, and applets*. These three mandated competencies are directly and substantively continuous with higher-order items comprising test #1, which assessed graphical interpretation of derivatives (item 1), determination of the tangent line equation at a given point (item 3), and identification of critical points and classification of extrema (item 5).

Originally derived from a large-scale survey of introductory physics courses using the force concept inventory across more than 6,000 students in 62 courses, normalized gain analysis was applied to a calculus-based course and reported gain distributions consistent with Hake's medium-gain category (Hoellwarth & Moelter, 2011). Hake's framework was applied in nursing education problem-based learning contexts and demonstrated that the low, medium, and high categories retained meaningful pedagogical interpretability when the instructional model involved collaborative and applied problem-solving components analogous to PBL (Morales-Mann & Kaitell, 2001).

Data cleaning and preparation

Both assessments were administered during regularly scheduled class sessions. Students were allocated approximately 25 minutes to complete each instrument under standard classroom conditions.

Raw data from both assessment points were subjected to a structured data-cleaning protocol prior to analysis. Column renaming includes student cell IDs, program, group, timing, score percentages.

Analytical Framework

All statistical analyses were performed in R version 4.5.2 (R Core Team, 2022) using RStudio version 2026.01.1 (Posit, 2026). These R packages were employed: *readxl* (v1.4.x) for import of Microsoft Excel data files; *dplyr* (v1.1.x) and *tidyr* (v1.3.x) for data manipulation and reshaping; *ggplot2* (v3.5.x), *ggpubr* (v0.6.x), and *scales* (v1.3.x) – data visualization; *rstatix* (v0.7.x) for tidy wrappers for t-tests, ANOVA, Shapiro-Wilk, and Friedman tests; *car* (v3.1.x) for Levene's test for homogeneity of variance and type III ANOVA; *flextable* for reporting tables; officer for Word and PowerPoint document manipulation; *effsize* (v0.8.x) for Cohen's d (Cohen, 1988), partial eta-squared (η^2p), and rank-biserial correlation r; *svglite* (v2.1.x) – vector-format figure export; *viridis* for color maps to improve graph readability; *patchwork* for arbitrarily complex composition of plots; *e1071* for simple implementation of machine learning methods.

RESULTS

Descriptive Statistics

Descriptive statistics for test #1 scores are presented in **Table 1** for control and experimental groups. On the total score, the control group obtained a mean of 17.31% (standard deviation [SD] = 13.53, median = 15.80, Q1 = 6.6, Q3 = 26.60, skewness = 0.56, kurtosis = -0.66), whereas the experimental group obtained a mean of 38.07% (SD = 22.08, median = 43.75, Q1 = 15.0, Q3 =

Table 2. Descriptive statistics of test #2 section scores by group

Variable	Group	M	SD	Q1	Q3	Skewness	Kurtosis
S1	Control	47.23	30.15	20	60	0.245	-0.937
	Experimental	71.36	25.66	60	100	-0.535	-0.415
S2	Control	56.40	30.23	25	75	-0.350	-0.897
	Experimental	76.14	28.51	50	100	-0.914	-0.336
S3	Control	19.44	26.86	0	25	1.324	0.918
	Experimental	59.09	37.39	25	100	-0.496	-1.259

Note. S1 represents MC; S2 represents TF; and S3 represents fill-in-the-blanks

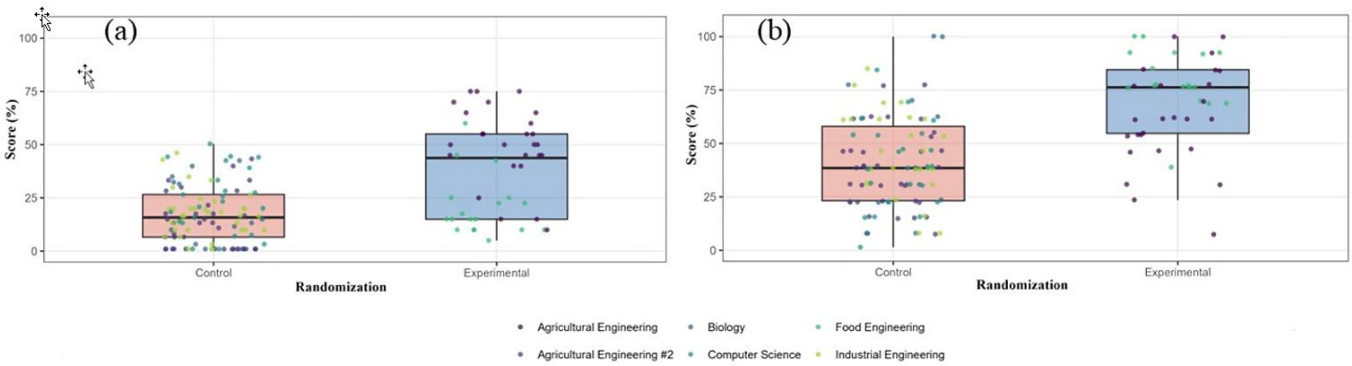

Figure 1. Test #1 paired boxplots of control (T0) and experimental (T1) score distributions (%) with individual data points color-coded by academic program (part a corresponds to test #1 and part b to test #2) (created by the author)

55.00, skewness = 0.08, kurtosis = -1.41). At the item level, mean scores were higher for the experimental group across all five subscales, with differences ranging from 6.2 percentage points on Q1 to 57.6 percentage points on Q5.

Descriptive statistics for test #2 scores are presented in Table 2 for control and experimental groups. On total score, the control group obtained a mean of 41.48% (SD = 21.46, median = 38.50, Q1 = 23.25, Q3 = 58.00, skewness = 0.443, kurtosis = -0.267), whereas the experimental group obtained a mean of 69.11% (SD = 21.84, median = 76.25, Q1 = 54.75, Q3 = 84.5, skewness = -0.695, kurtosis = 0.023). At section level, mean scores were higher for the experimental group across all subscales, with differences ranging from 19.74 percentage points on TF to 39.65 percentage points on fill-in-the-blanks.

In part a in Figure 1, the control group displayed a markedly lower and more compact distribution (median ≈ 15%, IQR roughly 5-25%), with most individual scores concentrated below 25% and no evidence of high-performing outliers. The experimental group showed a substantially higher and more dispersed distribution (median ≈ 43%, IQR roughly 30-55%), with scores spanning the full range from near zero to approximately 75%. Individual data points in both groups reveal considerable variability across academic programs, with no single program appearing to dominate either tail.

In part b in Figure 1, both groups achieved higher scores overall relative to test #1. The control group exhibited a median of approximately 38% (IQR roughly 23-58%), while the experimental group showed a notably elevated median of approximately 75% (IQR roughly 55-85%), with most individual scores clustering above 60%. The experimental group also displayed a tighter interquartile range in test #2 compared to test #1, suggesting more homogeneous performance. A small number of low-scoring outliers were observed in the Experimental group in both panels.

In part a in Figure 2, the control group violin is wide and flat across low score values (0-25%), reflecting a high concentration of students at the lower end of the scale, with a long upper tail extending toward 60%. The embedded boxplot confirms a low median and a narrow IQR. The experimental group violin presents a more symmetrical, elongated shape with the greatest density concentrated between approximately 25% and 55%, indicating a more uniformly distributed performance range compared to the control group.

In part b in Figure 2, both distributions shift markedly upward. The control group violin widens considerably across mid-range (25-60%), indicating greater density in that zone, though its shape remains notably flatter than the experimental group and extends down to near-zero scores. The experimental group violin is strikingly narrow, with density concentrated between approximately 60% and 90%, reflecting a highly consistent performance pattern. Near-zero width of both violins particularly for the experimental group at low scores suggests that very low performance was exceptional.

In part a in Figure 3, the control group distribution is strongly right-skewed, with the highest density mass concentrated between 0% and 20% and a long tapering tail toward 60%. The kernel density curve confirms a

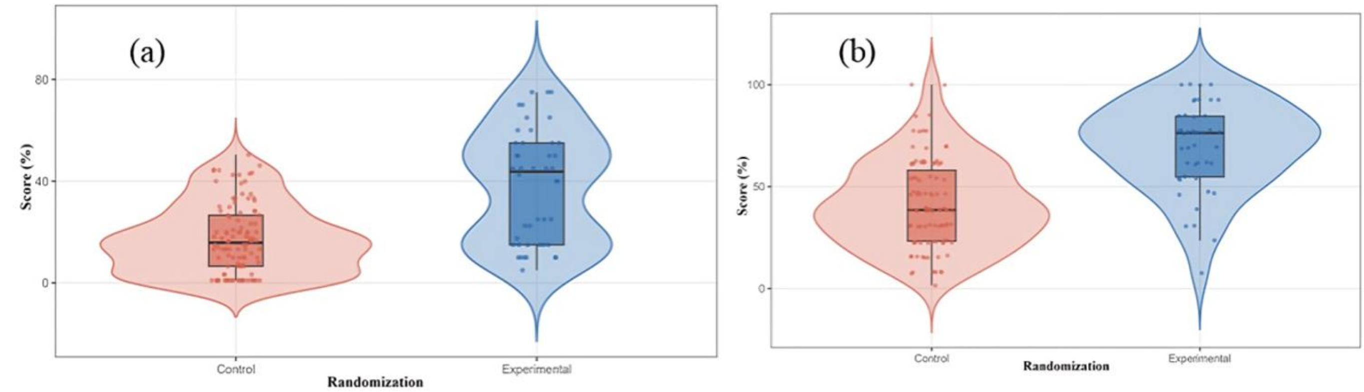

Figure 2. Violin-boxplot display and kernel density of score percentages (individual observations are overlaid as jittered points for the control and experimental groups, providing a detailed view of score density and distributional shape; part a corresponds to test #1 and part b to test #2) (created by the author)

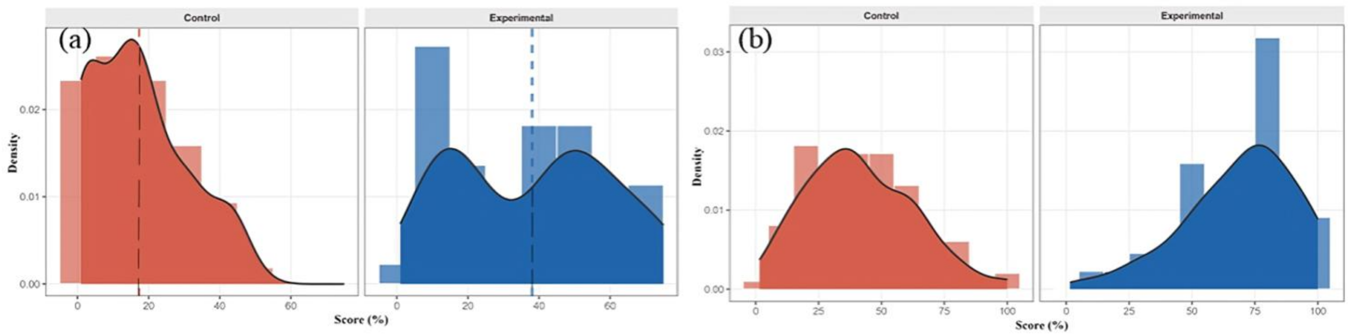

Figure 3. Density histograms of control and experimental scores with an overlaid smoothed density curve (scaled to frequency) (the dashed vertical line marks a mean value) (the author's own elaboration using R)

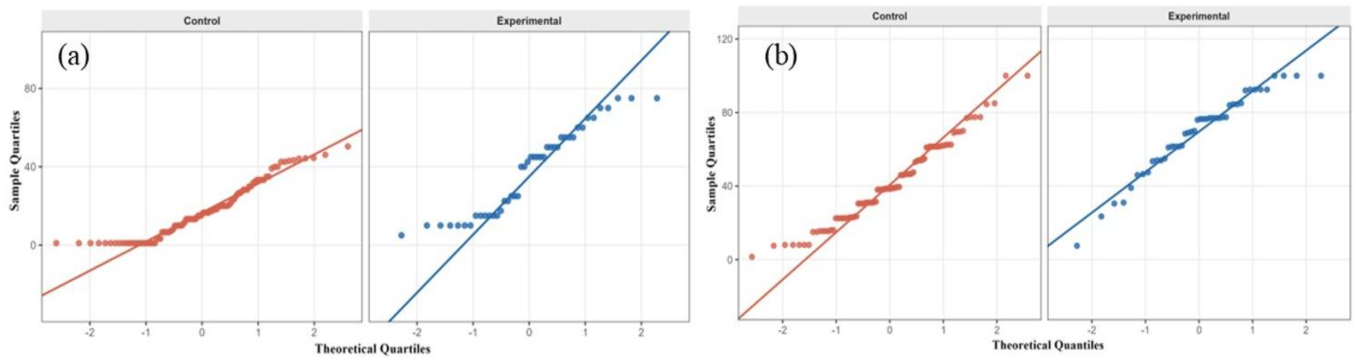

Figure 4. Normal Q-Q plots of control and experimental scores (points on the diagonal line represent normal distribution) (created by the author)

unimodal, positively skewed shape, consistent with skewness values reported in **Table 1**. By contrast, the experimental group displays a bimodal distribution: one mode near 15-20% and a more prominent mode around 40-45% suggesting two distinct performance subgroups.

In part b in **Figure 3**, both distributions shift toward higher scores. The control group retains a unimodal, approximately symmetric shape with its peak around 35%, though with a visible right tail. The experimental group shows a markedly different profile: a pronounced spike at 75, with a long-left tail, yielding a negatively skewed, high-concentration distribution at the upper end of the scale. The kernel density curve for the experimental group confirms this left-skewed, near-ceiling pattern.

Normality and Non-Parametric Statistics

In part a in **Figure 4**, both Q-Q plots reveal systematic departures from normality. The control group plot shows a pronounced clustering of points near zero at the lower tail consistent with the floor effect. It is followed by a gradual upward deviation from the reference line at higher quantiles, indicating a right-skewed distribution. The experimental group plot exhibits a characteristic S-shaped pattern, with points falling below the reference

Table 3. Shapiro-Wilk normality test results for total scores by group and assessment

Group	Test	n	W	*p*	Normal
Control	Test #1	107	0.9236	< .001	No
Experimental		44	0.9143	0.0031	No
Control	Test #2	99	0.9703	0.0245	No
Experimental		44	0.9468	0.0418	No

line at low quantiles and above it at high quantiles, reflecting a platykurtic distribution.

In part b in **Figure 4**, the control group points follow the reference line reasonably closely through the middle range of the distribution, suggesting approximate normality in the central portion. However, notable deviations are visible at both tails: points fall below the reference line at the lower quantiles and curve above it at the upper quantiles, indicating a distribution with a heavier right tail and some degree of positive skewness. In the experimental group, the departure from normality is more pronounced and systematic. Points at the lower quantiles fall substantially below the reference line, while the upper quantile points align more closely with it, tracing a concave pattern characteristic of a left-skewed distribution. This is consistent with the ceiling concentration of experimental group scores documented in test #2, where most students clustered in the 75-100% range.

These patterns are consistent with the Shapiro-Wilk results reported in **Table 3**, which confirmed non-normality in both groups, thereby justifying the use of non-parametric and robust inferential tests. The Shapiro-Wilk test indicated statistically significant departures from normality in all four group-by-test combinations. For test #1, non-normality was confirmed in both the control group and the experimental group. For test #2, the assumption of normality was likewise rejected for the control group and the experimental group, though W statistics for test #2 were closer to 1.0 in both groups, suggesting that deviations from normality were less severe than those observed in test #1.

Given that the normality assumption was violated across all conditions, Welch's t-test was retained as the primary inferential procedure, and the Mann-Whitney U test was employed as a non-parametric corroboration for all between-group comparisons. Levene's test further indicated that the homogeneity-of-variance assumption was not met, F (1, 149) = 24.49, $p < .001$.

Table 4. Welch's t-test and Mann-Whitney U test (also called the Wilcoxon rank-sum test) results for total scores by assessment

Instrument	Test	Statistic	gl_W	*p*_value
Test #1	Welch's t	-5.805	56.8	< .001
	Mann-Whitney U	1,071.000		< .001
Test #2	Welch's t	-7.020	81.3	< .001
	Mann-Whitney U	815.500		< .001

Note. gl_W represents Welch-Satterthwaite degrees of freedom

In **Table 4**, Welch's t-test for test #1 revealed a statistically significant difference between the control and experimental groups. The Mann-Whitney U test corroborated this result, confirming the between-group difference through a non-parametric approach that does not assume normality. For test #2, the between-group difference was likewise statistically significant and of greater magnitude. Welch's t-test yielded -7.020, and the Mann-Whitney U test again confirmed the result, U = 815.500, $p < .001$. Reduction in the U statistic from test #1 to test #2 from 1,071.0 to 815.5 reflects a more pronounced rank-order separation between groups on the second assessment.

Performance on Test #1 and Test #2 Item-Format Subscales

In part a in **Figure 5**, the experimental group achieved a mean score of 38.1%, more than double that of the control group. The non-overlapping confidence intervals (CIs) provide visual confirmation of the statistically significant between-group difference, consistent with the Welch's t-test result. The error bars also reveal that the experimental group carried greater uncertainty around its mean estimate, reflecting its larger SD and smaller sample size relative to the control group.

In part b in **Figure 5**, the between-group gap widens further in absolute terms. The experimental mean

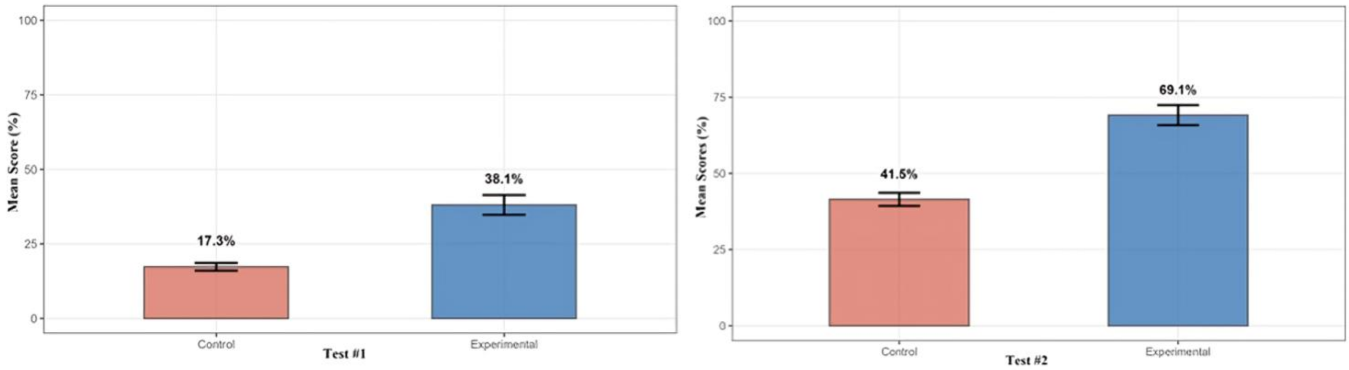

Figure 5. Bar charts of mean total scores (%) with 95% CI error bars for the control and experimental groups (created by the author)

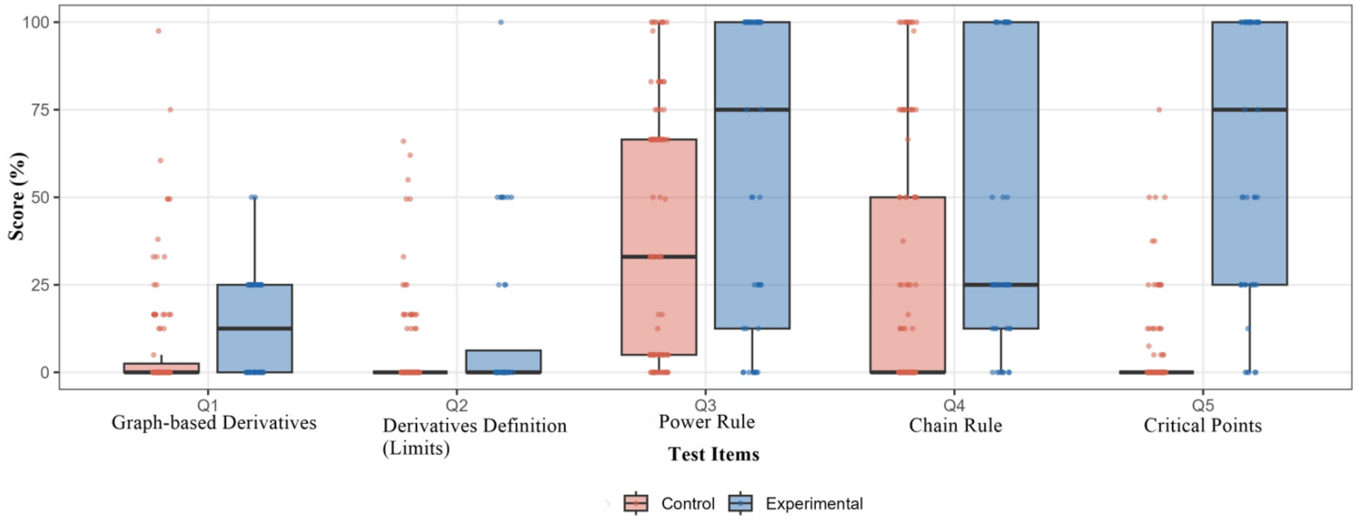

Figure 6. Side-by-side boxplots of item-level scores (%) for the control and experimental groups across five test #1 subscales (Q1-Q5), allowing direct visual comparison of group performance at task level (created by the author)

exceeded the control mean by approximately 27.6 percentage points. Both groups show narrower CIs than test #1, particularly the experimental group, indicating more precise mean estimation. Both tests indicate that the experimental group consistently outperformed the control group across both assessment points, with the magnitude of the advantage remaining substantial in test #2.

On Q1, both groups performed poorly overall, though the experimental group showed a somewhat higher median and a wider IQR compared to the control group (**Figure 6**). Several high-scoring outliers were observed in both groups, indicating that correct responses were exceptional.

On Q2, performance was markedly low in both groups. The control group median was effectively zero with minimal IQR, and the experimental group, while slightly higher, showed a median near zero and an IQR extending only to approximately 6%, with a single outlier at 50%. This test item proved to be the hardest for both groups.

On Q3, the largest absolute scores were recorded. The control group displayed an IQR spanning roughly 5-65%, while the experimental group achieved an IQR from approximately 12% to 100%, reflecting substantially higher and more variable performance. Both groups showed scores at the full 0-100% range.

On Q4, the control group median was near zero with an IQR of approximately 0-50%, whereas the experimental group median was around 25% with a wide IQR extending to 100%. Both groups exhibited considerable within-group variability and heterogenous mastery.

On Q5, the most striking between-group contrast was observed. The control group median was near zero with a very narrow IQR. By contrast, the experimental group achieved an IQR from roughly 25% to 100%, with many students scoring at the ceiling.

On the MC subscale, the control group exhibited an IQR spanning roughly 20-60%, and a full range from near zero to 90% (**Figure 7**). The experimental group achieved a substantially an IQR from roughly 60-90% and most individual scores concentrated in the upper half of the scale. A small number of low-scoring outliers were observed in the experimental group.

On the TF subscale, the between-group gap was more moderate. The control group showed an IQR spanning roughly 30-75%, while the experimental group had an IQR from approximately 50-90%. Both groups displayed wide individual variability and overlapping score ranges, though the experimental group consistently occupied the upper portion. This subscale produced the smallest between-group difference among three formats.

On the fill-in-the-blanks subscale, the most pronounced between-group contrast was observed. The control group median was near zero, with an IQR of approximately 0-25% and numerous individual scores at or close to zero. By contrast, the experimental group achieved an IQR from roughly 25-90%, demonstrating that most students attained meaningful performance on this cognitively demanding format. The Mann-Whitney test reported in **Table 4** confirms that this difference was statistically significant.

In **Figure 8**, bar charts show mean distribution with CIs. On the MC subscale, the experimental group achieved a mean of approximately 71%, compared to approximately 47% for the control group, a difference of roughly 24 percentage points. Non-overlapping CIs confirm a statistically significant between-group gap, consistent with Mann-Whitney. Notably, the control group error bar on this subscale is relatively narrow, reflecting the comparatively large sample size and moderate variability.

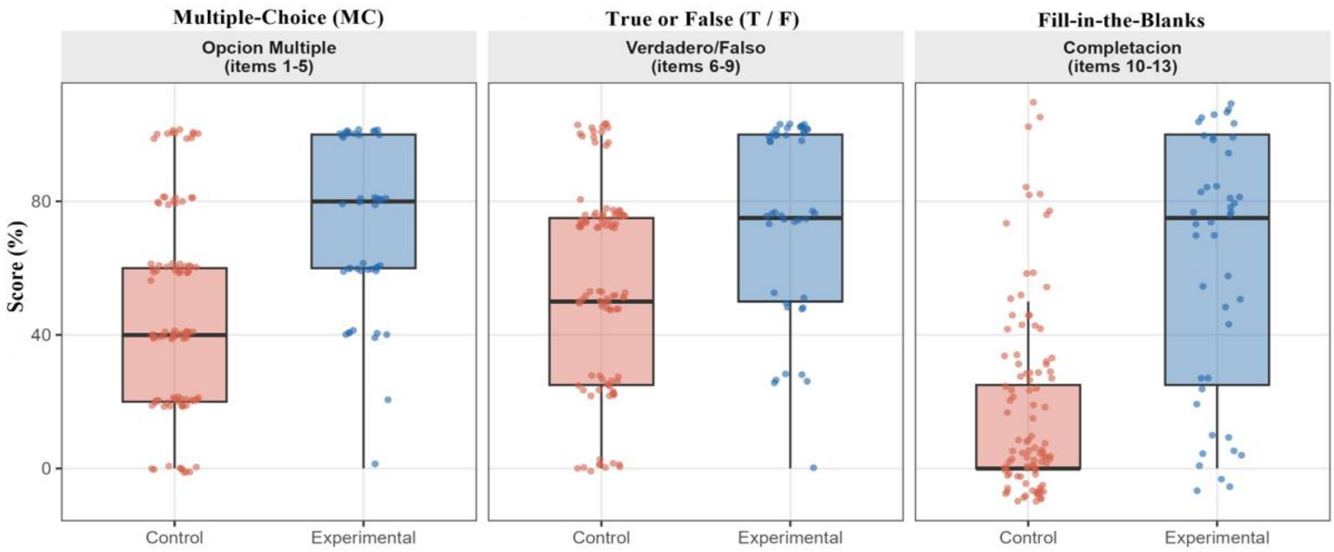

Figure 7. Boxplots of post-test percentage scores by item-format subscale (MC, TF, and completion) and engineering program (individual observations are overlaid as jittered points) (created by the author)

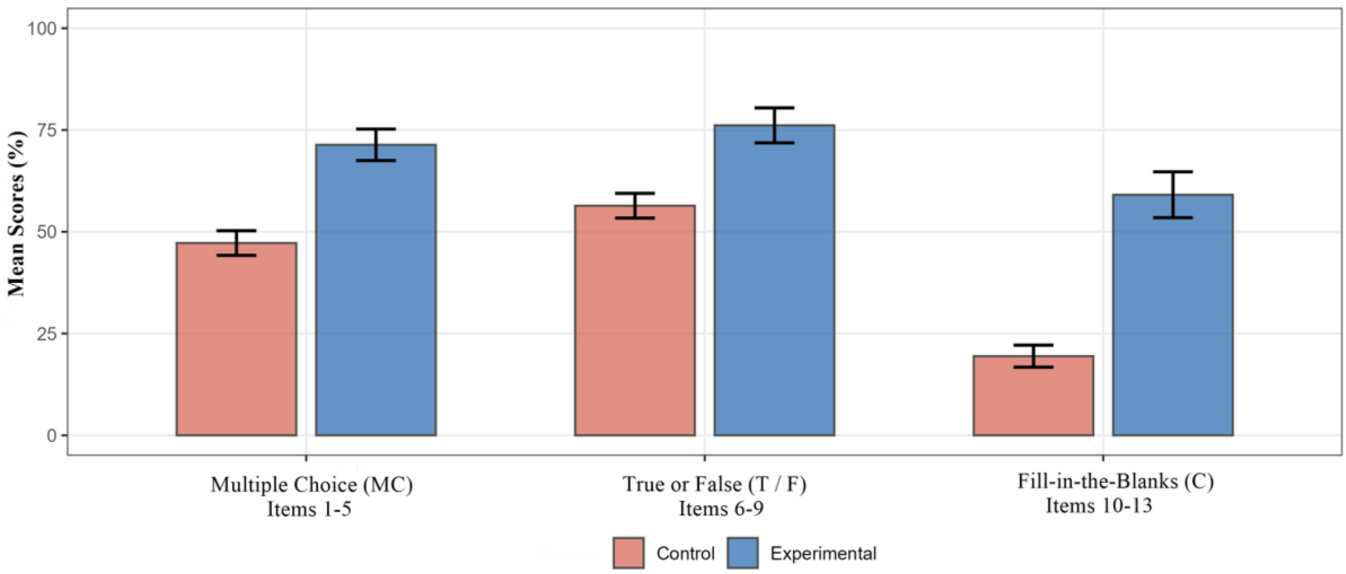

Figure 8. Bar charts of mean subscale scores (%) with 95% CI error bars for the control and experimental groups across the three test #2 item formats: MC(items 1-5), TF (items 6-9), and fill-in-the-blanks (C, items 10-13) (created by the author)

On the TF subscale, the control group recorded the highest mean of the three formats, while the experimental group reached approximately 76%. CIs do not overlap, indicating a significant difference, though the absolute gap of 19 percentage points is the smallest across the three subscales. This subscale was the strongest-performing format for the control group, suggesting that dichotomous-response items were more accessible.

On the fill-in-the-blanks subscale, the most striking contrast emerges. The control group mean was approximately 19%, by far the lowest of any format in either group, a gap of roughly 43 percentage points. Both CIs are wider here relative to the other subscales, reflecting greater within-group variability in open-ended procedural performance, consistent with SDs.

DISCUSSION

The present study investigated the effectiveness of an instructional model that integrates FL, PBL, and GeoGebra in enhancing students' achievement in differential and integral calculus. Overall, findings indicate that students exposed to the proposed methodology consistently outperformed those receiving conventional instruction across both assessments and all item formats. Rather than reflecting isolated improvements in procedural performance, these results suggest that combining active learning strategies with dynamic mathematical visualization promotes deeper conceptual understanding.

From a socio-constructivist perspective, results are consistent with Vygotsky's theory of mediated learning, which emphasizes that knowledge is constructed through social interaction and the use of cultural tools (Vygotsky & Cole, 1981). GeoGebra functioned as a cognitive mediator by allowing students to explore mathematical concepts through multiple representations, while collaborative activities embedded within PBL environment encouraged dialogue, reflection, and shared problem solving. Authentic tasks require students to actively construct knowledge instead of merely reproducing algorithms (Jonassen & Rohrer-Murphy, 1999).

GeoGebra integration appears to have played a central role in facilitating this transition between semiotic representations. According to Duval's theory of semiotic representation registers, meaningful mathematical understanding depends on students' ability to translate among algebraic, graphical, and numerical representations rather than manipulating a single representation in isolation. Many students initially experienced difficulties in interpreting the graphical representation of quadratic functions, deriving equations from geometric information, and connecting these representations with differentiation procedures. Furthermore, weaknesses in prerequisite algebraic skills including fraction operations, simplification of algebraic expressions, exponent manipulation, and distributive property, limited their ability to solve calculus problems efficiently. By enabling students to dynamically visualize tangent lines, extrema, concavity, and function behavior, GeoGebra supported the coordination of symbolic and graphical representations, thereby reducing cognitive load during concept formation. Widespread preference for the mobile version also suggests that accessibility and continuous interaction may have contributed to students' autonomous learning.

By providing learning materials before class including lecture notes, instructional videos, audio explanations, quizzes, flashcards, mind maps, and concise summaries, students were able to acquire foundational knowledge independently, allowing classroom time to be devoted to discussion, collaborative problem solving, and instructor feedback (Bergmann & Sams, 2023). Flipped classrooms improve academic performance, student engagement, and conceptual understanding (Freeman et al., 2014; Kay et al., 2019; Prince, 2004; Stieha et al., 2024) and proves to be an effective pedagogy across higher education contexts (Olivan-Blázquez et al., 2022; Plecis & Cheung, 2025).

PBL appears to have strengthened students' motivation by situating calculus concepts within authentic disciplinary contexts. Students developed projects involving optimization using derivatives and volume calculations using integrals, adapted to their respective engineering programs. Meaningful learning occurs when students actively solve realistic problems that require conceptual knowledge and reflect not only greater procedural competence but also increased cognitive engagement resulting from contextualized mathematical practice (Bell, 2010; Biggs et al., 2001; Hmelo-Silver, 2004).

Findings are also compatible with Self-Determination Theory, which proposes that learning environments supporting autonomy, competence, and relatedness enhance intrinsic motivation (Ryan & Deci, 2000). Flexibility enabled students to regulate their own pace of learning, while collaborative project work and GeoGebra-based exploration promoted both competence and peer interaction. Technology-supported active learning and student motivation become pedagogically effective when integrated within coherent instructional models rather than functioning merely as visualization tools (Baig & Yadegaridehkordi, 2023; Hao, 2016; Lai & Hwang, 2016).

The largest performance differences emerged in completion items, whereas MC and true/false formats produced comparatively smaller differences. Open-ended questions demand students to organize reasoning, establish relationships among concepts, and construct complete mathematical solutions, thereby providing a more sensitive measure of conceptual understanding than objective-response formats. Consistent with recent educational research (Ortega et al., 2024; van Borkulo et al., 2021; Williner, 2024), this pedagogical model represents a promising approach for improving mathematics instruction in engineering education (Darwazeh & Branch, 2015; Ortega et al., 2024; van Borkulo et al., 2021; Williner, 2024).

CONCLUSIONS

The present study aimed to evaluate the educational effectiveness of an instructional model integrating FL, PBL, and GeoGebra in undergraduate differential and integral calculus by examining students' academic performance through descriptive, inferential, and item-format analyses. Instructional design, which combined pre-class autonomous preparation with collaborative in-class problem-solving and authentic project-based applications, created deeper cognitive engagement. Superior performance of the experimental group was observed on the most cognitively demanding assessment components, especially open-ended tasks. These items required students not only to recall procedures but also to generate mathematical expressions and justify intermediate steps (Oliveira-Groenwald & Homa, 2021; Stein & Smith, 1998). Within this framework, mathematical tasks can be classified into four hierarchical levels. Level 1 tasks involve the reproduction of previously learned formulas, rules, and theorems with little cognitive effort. Level 2 tasks require familiar algorithms in situations where the appropriate mechanical procedure is immediately apparent. In contrast, level 3 tasks promote conceptual understanding by requiring students to translate across multiple mathematical representations, such as symbolic, graphical, and numerical forms. Finally, level 4 tasks demand complex, non-algorithmic reasoning, requiring students to formulate conjectures, justify mathematical arguments, verify solutions, and communicate their reasoning effectively. The instructional design implemented in the present study deliberately emphasized level 3 through GeoGebra, collaborative inquiry, and project-based activities. Consequently, students were encouraged to move beyond procedural fluency toward conceptual reasoning (Ayala-Chauvin & Luna-Romero, 2025; Calderón-Sarango, 2025; Guerrero-Chirinos et al., 2025).

The study was conducted at a single higher education institution, which may limit the generalizability of the results to other educational contexts or disciplines. Although the assessments captured students' academic performance, the study did not directly measure changes in normalized student gains (Coletta & Phillips, 2005), self-regulated learning, or motivation through validated psychometric instruments (Marx & Cummings, 2007). Future research should incorporate longitudinal designs, multiple institutions, qualitative evidence, and structural equation modeling.

A longitudinal design tracking the same students across multiple courses and academic cycles would allow researchers to determine whether the learning gains observed in the present study are sustained over time (Hmelo-Silver, 2004; Johnson et al., 1994; Morris et al., 1977). The extension of the present framework to the other courses or subdisciplines would enable a comparative analysis (Gómez Vargas, 2016; Stewart et al., 2019; Ziatdinov & Valles, 2022).

Funding: This study was funded by Universidad Técnica Particular de Loja, grant number POA VIN-56.

Acknowledgments: The author would like to thank Deisy Cecilia Calderón-Sarango, who graduated from the master's program in education with a specialization in mathematics teaching at the Universidad Técnica Particular de Loja. The author would also like to thank Prof. Marco Antonio Ayala-Chauvin, who developed the open-format Test #1 used in the evaluation process, while Deisy Cecilia Calderón-Sarango designed Test #2.

Ethical statement: This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki and adhered to the ethical guidelines of Universidad Técnica Particular de Loja (UTPL). The research involved the analysis of academic performance data administered during regularly scheduled class sessions. Formal ethical review approval was not required for this study under the applicable institutional regulations, given that the data were collected exclusively through routine educational evaluation activities and did not involve experimental procedures, medical interventions, sensitive personal data, or recruitment beyond the enrolled student population. Participation was limited to students enrolled in STEM-related courses at UTPL who experienced the Flipped Learning combined with Project-Based Learning instructional model. Student responses were fully anonymized prior to analysis, and no personally identifiable information was disclosed. All procedures were conducted with full respect for students' academic integrity.

AI statement: During the preparation of this manuscript, the author used ChatGPT-3.5 (OpenAI) to improve the linguistic clarity of the introduction and discussion sections. No AI tool was used for abstract, conclusion, data analysis, or generation of original scientific content. Additionally, Claude 3.5 Sonnet (Anthropic) was utilized to review and correct certain segments of the R code script. The author has carefully reviewed and edited the output from these tools and take full responsibility for the content and accuracy of this publication.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data presented here are openly available on Zenodo at https://doi.org/10.5281/zenodo.19057381.

REFERENCES

Addae, R. K. (2025). Application of GeoGebra to the teaching of vector algebra. *International Journal of Research Publication and Reviews, 6*(4), 15468-15483. https://doi.org/10.55248/gengpi.6.0425.1688

Aguiar, A., & Calabrese, J. (2025). Broadening the perspective of traditional lecturing vs. active learning in introductory college biology. *Discover Education, 4*, Article 278. https://doi.org/10.1007/S44217-025-00729-7

Allen, D., & Tanner, K. D. (2005). Infusing active learning into the large-enrollment biology class: Seven strategies, from the simple to complex. *Cell Biology Education, 4*(4), 262-268. https://doi.org/10.1187/CBE.05-08-0113

Armbruster, P., Patel, M., Johnson, E., & Weiss, M. R. (2009). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE – Life Sciences Education, 8*(3), 203–213. https://doi.org/10.1187/CBE.09-03-0025

Ayala-Chauvin, M. A., & Luna-Romero, R. L. (2025). El uso de GeoGebra para la comprensión de la derivada implícita en ingeniería [The use of GeoGebra for understanding implicit derivatives in engineering]. *Revista Científica Multidisciplinaria SAPIENTIAE, 8*(16), 376-386. https://doi.org/10.56124/sapientiae.v8i16.021

Aytekin, C., & Kiymaz, Y. (2019). Teaching linear algebra supported by GeoGebra visualization environment. *Acta Didactica Napocensia, 12*(2), 75-96. https://doi.org/10.24193/adn.12.2.7

Baig, M. I., & Yadegaridehkordi, E. (2023). Flipped classroom in higher education: A systematic literature review and research challenges. *International Journal of Educational Technology in Higher Education, 20*, Article 61. https://doi.org/10.1186/S41239-023-00430-5

Bell, S. (2010). Project-based learning for the 21st century: Skills for the future. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 83*(2), 39-43. https://doi.org/10.1080/00098650903505415

Bergmann, J., & Sams, A. (2023). *Flip your classroom: Reaching every student in every class every day*. International Society for Technology in Education.

Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology, 71*(1), 133-149. https://doi.org/10.1348/000709901158433

Calderón-Sarango, D. C. (2025). *Influencia del uso del GeoGebra en la enseñanza de la derivada en estudiantes ingeniería química y de alimentos de la UTPL* [Influence of the use of GeoGebra in the teaching of the derivative in chemical and food engineering students at UTPL] [Master's thesis, Universidad Técnica Particular de Loja].

Campos Nava, M., Torres Rodríguez, A., & Morales Maure, L. (2021). GeoGebra como medio para identificar patrones en la clase de álgebra lineal: Una propuesta concreta [GeoGebra as a means to identify patterns in the linear algebra class: A concrete proposal]. *Revista Universidad y Sociedad, 13*(2), 528-537. https://rus.ucf.edu.cu/index.php/rus/article/view/1994

Cirillo, M., Berk, D., LaRochelle, R., Bieda, K. N., & Arbaugh, F. (2022). Undergraduate students' perceptions of features of active learning models for teaching and learning to teach mathematics. *International Journal of Research in Undergraduate Mathematics Education, 10*(1), 172-200. https://doi.org/10.1007/S40753-022-00191-Y

Cohen, J. (1988). Set correlation and contingency tables. *Applied Psychological Measurement, 12*(4), 425-434. https://doi.org/10.1177/014662168801200410

Coletta, V. P., & Phillips, J. A. (2005). Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability. *American Journal of Physics, 73*(12), 1172-1182. https://doi.org/10.1119/1.2117109

Córdova-Esparza, D.-M., & Jimenez Piñon, A. (2025). Analysis of teaching strategies and learning challenges in linear algebra. *CLEI Electronic Journal, 28*(2). https://doi.org/10.19153/cleiej.28.2.11

Dancy, M., Henderson, C., Apkarian, N., Johnson, E., Stains, M., Raker, J. R., & Lau, A. (2024). Physics instructors' knowledge and use of active learning has increased over the last decade but most still lecture too much. *Physical Review Physics Education Research, 20*, Article 010119. https://doi.org/10.1103/PHYSREVPHYSEDUCRES.20.010119

Darwazeh, A. N., & Branch, R. M. (2015). A revision to the revised Bloom's taxonomy. In M. Simonson (Ed.), *Proceedings of the 39th Annual Convention, 2* (pp. 220-225). The Association for Educational Communications and Technology.

Davies, R., Dean, D. L., & Ball, N. (2013). Flipping the classroom and instructional technology integration in a college-level information systems spreadsheet course. *Educational Technology Research and Development, 61*(4), 563-580. https://doi.org/10.1007/S11423-013-9305-6

Dietrich, H., Evans, T., & Tatto, M. T. (2022). Traditional lectures versus active learning – A false dichotomy? *STEME, 2*(4), 275-292. https://doi.org/10.3934/STEME.2022017

Doolittle, P., Wojdak, K., & Walters, A. (2022). Defining active learning: A restricted systematic review. *Teaching and Learning Inquiry, 11*. https://doi.org/10.20343/teachlearninqu.11.25

Dzaiy, A. H. S., & Abdullah, S. A. (2024). The use of active learning strategies to Foster effective teaching in higher education institutions. *Zanco Journal of Humanity Sciences, 28*(4), 328-351. https://doi.org/10.21271/ZJHS.28.4.18

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *PNAS, 111*(23), 8410-8415. https://doi.org/10.1073/pnas.1319030111

Gallo, H. G., Verón, C., & Herrera, C. G. (2019). Interpretación de transformaciones lineales en el plano utilizando GeoGebra [Interpretation of linear transformations in the plane using GeoGebra]. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, (24), 32-37. https://doi.org/10.24215/18509959.24.e04

Gómez Vargas, J. C. (2016). Análisis de los contenidos y el método didáctico de la asignatura de geometría descriptiva desde su perspectiva histórica [Analysis of the content and teaching method of the subject of descriptive geometry from its historical perspective]. *Universidad de Granada. Departamento de Expresión Gráfica, Arquitectónica y en la Ingeniería.* https://digibug.ugr.es/handle/10481/41912

Guerrero-Chirinos, R. A., Pérez-Benítez, H. A., Delgado-Vega, R. J., León-Loaiza, M. A., & Santos-Torres, R. N. (2025). GeoGebra en el rendimiento académico y la comprensión de la derivada en estudiantes de ingeniería: Un estudio cuasiexperimental [GeoGebra in academic performance and understanding of the derivative in engineering students: A quasi-experimental study]. *Ciencia Latina Revista Científica Multidisciplinar, 9*(6), 7214-7229. https://doi.org/10.37811/cl_rcm.v9i6.21873

Hao, Y. (2016). Exploring undergraduates' perspectives and flipped learning readiness in their flipped classrooms. *Computers in Human Behavior, 59*, 82-92. https://doi.org/10.1016/j.chb.2016.01.032

Havelková, V. (2013). GeoGebra in teaching linear algebra. In M. Baptista Nunes, & P. Isaias (Eds.), *Proceedings of the International Conference on E-Learning* (pp. 581-589). IADIS Press.

Hew, K. F., & Lo, C. K. (2018). Flipped classroom improves student learning in health professions education: A meta-analysis. *BMC Medical Education, 18*, Article 38. https://doi.org/10.1186/s12909-018-1144-z

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*(3), 235-266. https://doi.org/10.1023/B:EDPR.0000034022.16470.f3

Hoellwarth, C., & Moelter, M. J. (2011). The implications of a robust curriculum in introductory mechanics. *American Journal of Physics, 79*(5), 540-545. https://doi.org/10.1119/1.3557069

Idsardi, R. (2019). Evidence-based practices for the active learning classroom. In J. J. Mintzes, & E. M. Walters (Eds.), *Active learning in college science* (pp. 13-25). Springer. https://doi.org/10.1007/978-3-030-33600-4_2

Indorf, J. L., Benabentos, R., Daubenmire, P. L., Murasko, D. M., Hazari, Z., Potvin, G., Kramer, L. H., Marsteller, P., Thompson, K. V., Cassone, V. M., & Stanford, J. S. (2021). Distinct factors predict use of active learning techniques by pre-tenure and tenured STEM faculty. *Journal of Geoscience Education, 69*(4), 357-372. https://doi.org/10.1080/10899995.2021.1927461

Johnson, D. W., Johnson, R. T., & Johnson, E. H. (1994). The new circles of learning: Cooperation in the classroom and school. *Association for Supervision and Curriculum Development*. https://files.eric.ed.gov/fulltext/ED241516.pdf

Jonassen, D. H., & Rohrer-Murphy, L. (1999). Activity theory as a framework for designing constructivist learning environments. *Educational Technology Research and Development, 47*(1), 61-79. https://doi.org/10.1007/BF02299477

Kay, R., MacDonald, T., & DiGiuseppe, M. (2019). A comparison of lecture-based, active, and flipped classroom teaching approaches in higher education. *Journal of Computing in Higher Education, 31*(3), 449-471. https://doi.org/10.1007/S12528-018-9197-X

Klein, K., Calabrese, J., Aguiar, A., Sunny, M., Ajani, K., Almajid, R., & Jennifer, A. (2022). Evaluating active lecture and traditional lecture in higher education. *Journal on Empowering Teaching Excellence, 7*(2), Article 6. https://doi.org/10.26077/ba42-a5cc

Kozanitis, A., & Nenciovici, L. (2022). Effect of active learning versus traditional lecturing on the learning achievement of college students in humanities and social sciences: A meta-analysis. *Higher Education, 86*(6), 1377-1394. https://doi.org/10.1007/S10734-022-00977-8

Lai, C., & Hwang, G. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education, 100*, 126-140. https://doi.org/10.1016/j.compedu.2016.05.006

Lizana Chauca, D., & Antezana Iparraguirre, R. P. (2021). Representación semiótica en el aprendizaje de conceptos básicos de la estructura algebraica de grupo [Semiotic representation in the learning of basic concepts of the algebraic structure of a group]. *Horizonte de La Ciencia, 11*(21), 177-188. https://doi.org/10.26490/uncp.horizonteciencia.2021.21.904

Marx, J. D., & Cummings, K. (2007). Normalized change. *American Journal of Physics, 75*(1), 87-91. https://doi.org/10.1119/1.2372468

Ministerio de Educación, Deporte y Cultura del Ecuador. (2025). Currículo priorizado con énfasis en competencias comunicacionales, matemáticas, digitales y socioemocionales. Nivel de bachillerato general. Contiene inserciones curriculares [Prioritized curriculum with emphasis on communication, mathematics, digital, and socio-emotional skills. General high school level. Includes curricular insertions]. *Ministerio de Educación*. https://educacion.gob.ec/wp-content/uploads/downloads/2025/08/Curriculo-Priorizado-Bachillerato.pdf

Morales-Mann, E. T., & Kaitell, C. A. (2001). Problem-based learning in a new Canadian curriculum. *Journal of Advanced Nursing, 33*(1), 13-19. https://doi.org/10.1046/j.1365-2648.2001.01633.x

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*(5), 519-533. https://doi.org/10.1016/S0022-5371(77)80016-9

Nagamalla, L., Readdy, V., Kumar, S., Kolagani, L., & Suryadevara, G. (2025). Impact of active learning methods on project-based learning (PBL): Enhancing student engagement and outcomes. *Journal of Engineering Education Transformations, 38*, 30-37. https://doi.org/10.16920/JEET/2025/V38IS2/25005

Nguyen, K., Borrego, M., Finelli, C., DeMonbrun, M., Crockett, C., Tharayil, S., Shekhar, P., Waters, C., & Rosenberg, R. S. (2021). Instructor strategies to aid implementation of active learning: A systematic literature review. *International Journal of STEM Education, 8*(9). https://doi.org/10.1186/S40594-021-00270-7

Olivan-Blázquez, B., Aguilar-Latorre, A., Gascón, S., Gómez-Poyato, M. J., Valero, D., Magallón-Botaya, R., Heah, R., & Porroche-Escudero, A. (2022). Comparing the use of flipped classroom in combination with problem-based learning or with case-based learning for improving academic performance and satisfaction. *Active Learning in Higher Education, 24*(3), 373-388. https://doi.org/10.1177/14697874221081550

Oliveira Groenwald, C. L., & Ryokiti Homa, A. I. (2021). Educación matemáticas y tecnologías: Planificación de tareas de investigación centradas en el aprendizaje de los estudiantes [Mathematics and technology education: Planning research tasks focused on student learning]. *Unión – Revista Iberoamericana de Educación Matemática, 17*(63).

Ortega, J., Castañeda, C., Rivera-Casavilca, R., Mencia-Sanchez, N., Simon, M., & Navarro, I. (2024). Software GeoGebra y aprendizaje del cálculo integral en estudiantes de ingeniería civil – Universidad Nacional de Huancavelica [GeoGebra software and the learning of integral calculus in civil engineering students – National University of Huancavelica]. *Ciencia Latina Revista Científica Multidisciplinar, 7*(6), 8019-8031. https://doi.org/10.37811/cl_rcm.v7i6.9329

Paiva, H. M., Hayashi, V. T., Santoro, F. M., Paiva, J. P. L. M., & Garcia, M. (2025). Active learning in STEM education: An approach combining flipped classroom and dialogic teaching. In *Proceedings of the IEEE Global Engineering Education Conference* (pp. 1-10). IEEE. https://doi.org/10.1109/EDUCON62633.2025.11016558

Paleenud, I., Tanprasert, K., & Waleeittipat, S. (2023). Lecture-based and project-based approaches to instruction, classroom learning environment, and deep learning. *European Journal of Educational Research, 13*(2), 531-539. https://doi.org/10.12973/EU-JER.13.2.531

Pinnelli, S., & Fiorucci, A. (2015). University and flipped learning TIC & DIL project: Framework and design. *International Association for Development of the Information Society*. https://eric.ed.gov/?id=ED562195

Plecis, P., & Cheung, M. (2025). Effectiveness of flipped classroom – Project-based learning (FC-PjBL) hybrid approach on student performance in general biology. *Journal of Interdisciplinary Perspectives, 3*(12), 378-398. https://doi.org/10.69569/jip.2025.421

Posit (2026). RStudio: Integrated development environment for R. *Posit Software*. http://www.posit.co/

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223-231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Quintilla Castán, M., & Fernández Morales, A. (2021). GeoGebra para la enseñanza de la geometría descriptiva: Aplicación para la docencia online [GeoGebra for teaching descriptive geometry: An application for online teaching]. In *Proceedings of the 9th Conference on Innovation in Teaching in Architecture* (pp. 285-295). https://doi.org/10.5821/jida.2021.10545

R Core Team. (2022). A language and environment for statistical computing. *R Project*. https://www.r-project.org/

Rahman, A. A., Zaid, N. M., Aris, B., Abdullah, Z., Mohamed, H., & Meijden, H. van der. (2016). Implementation strategy of project based learning through flipped classroom method. In *Proceedings of the 2016 IEEE Conference on E-Learning, e-Management and e-Services* (pp. 1-5). IEEE. https://doi.org/10.1109/IC3E.2016.8009030

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. https://doi.org/10.1037/0003-066X.55.1.68

Saad, R. E. B., García, A. L., & García, J. M. A. (2025). Mapping constructivist active learning for STEM: Toward sustainable education. *Sustainability, 17*(13), Article 6225. https://doi.org/10.3390/SU17136225

Sadia, S., Ghassan, A., Mirza, T. I., Nazir, M., & Abrar, S. (2025). Beyond the lecture: A comparative analysis of active learning interventions in foundational science modules and their impact on student engagement and performance. *European Journal of Clinical Pharmacy, 8*(1), 566-573.

Scott, C., Green, L. E., & Etheridge, D. L. (2016). A comparison between flipped and lecture-based instruction in the calculus classroom. *Journal of Applied Research in Higher Education, 8*(2), 252-264. https://doi.org/10.1108/JARHE-04-2015-0024

Sholahuddin, A., Anjuni, N., Leny, L., & Faikhamta, C. (2023). Project-based and flipped learning in the classroom: A strategy for enhancing students' scientific literacy. *European Journal of Educational Research, 12*(1), 239-251. https://doi.org/10.12973/EU-JER.12.1.239

Sohrabi, B., & Iraj, H. (2016). Implementing flipped classroom using digital media: A comparison of two demographically different groups perceptions. *Computers in Human Behavior, 60*, 514-524. https://doi.org/10.1016/J.CHB.2016.02.056

Steen-Utheim, A. T., & Foldnes, N. (2018). A qualitative investigation of student engagement in a flipped classroom. *Teaching in Higher Education, 23*(3), 307-324. https://doi.org/10.1080/13562517.2017.1379481

Stein, M. K., & Smith, M. S. (1998). Mathematical tasks as a framework for reflection: From research to practice. *Mathematics Teaching in the Middle School, 3*(4), 268-275. https://doi.org/10.5951/MTMS.3.4.0268

Stewart, S., Andrews-Larson, C., & Zandieh, M. (2019). Linear algebra teaching and learning: Themes from recent research and evolving research priorities. *ZDM Mathematics Education, 51*, 1017-1030. https://doi.org/10.1007/s11858-019-01104-1

Stieha, V., Earl, B., Hagens, H., Haynes, M., Ulappa, A., Bond, L., & Oxford, J. T. (2024). An exploration of the relationship between active learning and student motivation in STEM: A mixed methods study. *Advances in Physiology Education, 48*(3), 621-638. https://doi.org/10.1152/ADVAN.00247.2022

Ting, F. S. T., Lam, W. H., & Shroff, R. H. (2019). Active learning via problem-based collaborative games in a large mathematics university course in Hong Kong. *Education Sciences, 9*(3), Article 172. https://doi.org/10.3390/EDUCSCI9030172

Ting, F. S. T., Shroff, R. H., Lam, W. H., Garcia, R. C. C., Chan, C. L., Tsang, W. K., & Ezeamuzie, N. O. (2022). A meta-analysis of studies on the effects of active learning on Asian students' performance in science, technology, engineering and mathematics (STEM) subjects. *The Asia-Pacific Education Researcher, 32*(3), 379-400. https://doi.org/10.1007/S40299-022-00661-6

Trigueros, M., & Wawro, M. (2020). Linear algebra teaching and learning. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 474-478). Springer. https://doi.org/10.1007/978-3-030-15789-0_100021

van Borkulo, S., Chytas, C., Drijvers, P., Barendsen, E., & Tolboom, J. (2021). Computational thinking in the mathematics classroom: Fostering algorithmic thinking and generalization skills using dynamic mathematics software. In *Proceedings of the 16th Workshop in Primary and Secondary Computing Education* (pp. 1-9). https://doi.org/10.1145/3481312.3481319

Von Korff, J., Archibeque, B., Gomez, K. A., Heckendorf, T., McKagan, S. B., Sayre, E. C., Schenk, E. W., Shepherd, C., & Sorell, L. (2016). Secondary analysis of teaching methods in introductory physics: A 50 k-student study. *American Journal of Physics, 84*(12), 969-974. https://doi.org/10.1119/1.4964354

Vygotsky, L. S., & Cole, M. (1981). *Mind in society: The development of higher psychological processes*. Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Walker, J. D., Cotner, S., Baepler, P., & Decker, M. D. (2008). A delicate balance: Integrating active learning into a large lecture course. *CBE – Life Sciences Education, 7*(4), 361-367. https://doi.org/10.1187/CBE.08-02-0004

Williner, B. (2024). Influencia de tareas con *software* GeoGebra en el desarrollo de la competencia matemática en estudiantes de ingeniería [Influence of tasks with GeoGebra software on the development of mathematical competence in engineering students]. *Bolema: Boletim de Educação Matemática, 38*, Article e230215. https://doi.org/10.1590/1980-4415v38a230215

Ziatdinov, R., & Valles, J. R. (2022). Synthesis of modeling, visualization, and programming in GeoGebra as an effective approach for teaching and learning STEM topics. *Mathematics, 10*(3), Article 398. https://doi.org/10.3390/math10030398