Journal of Curriculum Studies Research

https://curriculumstudies.org

E-ISSN: 2690-2788

Volume: 8 Issue: 2 2026

pp. 290-317

Improving Academic Performance in the Lowest Quartile: A Multilevel Analysis of a Remedial Intervention in Ecuadorian Public Schools

María Eugenia Verdugo Guamán*[a], Freddy Patricio Cabrera Ortiz*[b], Homero Patricio Cabrera Tenecela[a], & Mauricio Esteban Reyes Guaranda[b]

* Corresponding author
Email:
freddy.cabrera@ucuenca.edu.ec
a. Universidad de Cuenca, Cuenca, Ecuador
b. Grupo de Investigación en Políticas Educativas, Universidad de Cuenca, Cuenca, Ecuador

Article Info
Received: August 05, 2025
Accepted: August 02, 2026
Published: September 1, 2026

10.46303/jcsr.2026.34

How to cite
Verdugo Guamán, M. E., Cabrera Ortiz, F. P., Cabrera Tenecela, H. P., & Reyes Guaranda, M. E. (2026). Improving Academic Performance in the Lowest Quartile: A Multilevel Analysis of a Remedial Intervention in Ecuadorian Public Schools. *Journal of Curriculum Studies Research*, *8*(2), 290-317.
https://doi.org/10.46303/jcsr.2026.34

Copyright license
This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

ABSTRACT
This study assessed the academic impact of an educational leveling intervention implemented across various institutions in Ecuador. Using a Difference-in-Differences (DiD) strategy within a multilevel modeling framework, we estimated changes in academic performance (GPA) over time between intervention and control groups. The analysis was conducted in RStudio using linear mixed-effects models (via the lmerTest package) to account for the hierarchical structure of the data (students nested within grades, institutions, and intervention types). Five models were compared, ranging from a full triple-interaction specification to main effects only, with Model 2 (including Time × Intervention and Time × Moment interactions) providing the best fit according to AICc weights and conditional R^2. The findings show a statistically significant relative improvement in GPA for the intervention group compared to the control group, particularly following the second implementation point. Moreover, performance differences were moderated by students' prior academic levels, highlighting the importance of stratified interventions. Diagnostic tests confirmed that assumptions of linearity, homoscedasticity, and independence were reasonably met. However, residuals were not perfectly normal; the literature supports the robustness of mixed models to such deviations, especially in large samples. These results suggest that targeted academic interventions can positively influence student achievement when they account for students' baseline performance levels and institutional diversity.

KEYWORDS
Educational intervention; academic performance; low academic achievement; multilevel analysis; education.

INTRODUCTION

Student academic performance in basic education is a key indicator of equity and social development (OECD, 2018). However, significant inequalities persist, disproportionately affecting students in the lowest performance quartile, who are often from disadvantaged socioeconomic backgrounds. These disparities not only perpetuate cycles of poverty and exclusion but also hinder progress toward the Sustainable Development Goals (SDGs), specifically SDG 4, which seeks to ensure inclusive and equitable quality education, and SDG 10, which aims to reduce inequalities within and between countries (United Nations, 2015).

From a cognitive perspective, Sweller's Cognitive Load Theory (Sweller, 1988; Sweller et al., 2011) posits that effective learning depends on optimizing the presentation of information to avoid overloading working memory. This framework is particularly relevant to addressing acute cases of academic underachievement, as it supports the design of instructional interventions that reduce extraneous cognitive load and promote efficient processing of essential content. Complementing this perspective, John Hattie's research (2009), based on meta-analyses of thousands of educational studies, highlighted feedback as one of the most influential factors in improving academic achievement. According to Hattie, effective feedback not only informs students about their performance but also guides them in improving, closing gaps, and progressing toward learning goals. Together, these two theoretical perspectives are not merely complementary but structurally linked to the equity aims of this study. Cognitive Load Theory addresses the instructional design dimension of the intervention—ensuring that leveling materials are structured to avoid overwhelming students already facing accumulated learning gaps—while Hattie's feedback framework addresses the relational and pedagogical delivery dimension, justifying the tutor-mediated, small-group format through which preservice teachers could calibrate feedback to each student's proximal needs. This dual foundation positions the intervention not as a generic remediation effort, but as one deliberately designed around mechanisms known to be effective specifically for students facing the compounded disadvantages of low academic performance and socioeconomic vulnerability described earlier in this introduction.

This theoretical foundation is further strengthened by two complementary perspectives that directly inform the study's second objective and its attention to student motivation. First, Bandura's (1997) self-efficacy theory holds that individuals' beliefs in their own capacity to execute the actions required to attain specific outcomes are among the strongest predictors of sustained effort and academic performance, particularly under conditions of prior underachievement. This construct provides the theoretical grounding for the self-efficacy and motivation dimension referenced earlier in this introduction, linking the pedagogical design of the intervention to students' internal beliefs about their own capability to improve. Second, Bronfenbrenner's (1979) ecological systems theory conceptualizes human development as shaped by nested and interacting environmental systems—ranging from the immediate microsystem of the classroom and family to the broader macrosystem of regional and

institutional conditions. This perspective offers a theoretical rationale for treating region, school shift, educational level, and timing of intervention delivery not as mere statistical covariates, but as expressions of the layered contextual systems in which student learning is embedded, thereby reinforcing the conceptual coherence of the study's second research question.

Social and digital inequality are crucial factors in this context. Students from vulnerable backgrounds often face significant barriers, such as limited access to high-quality technological and pedagogical resources, which deepen learning gaps. Recent research highlighted that digital divides amplify educational inequalities, particularly during and after disruptive events such as the COVID-19 pandemic (World Bank, 2021). These gaps affect not only access to learning but also the acquisition of essential 21st-century life skills, exacerbating structural inequality in academic performance (Mapisa & Makena, 2024) This pattern is not confined to K-12 settings: even at the postsecondary level, socioeconomic disparities in technology access continue to condition students' ability to engage with digital learning resources, underscoring the need for institutional strategies—infrastructure investment, targeted training, and financial support—that address the socioeconomic roots of the divide rather than technology access alone (Ajani et al., 2025; Dyantyi & Mkabile-Masebe, 2025).

In this context, educational intervention programs emerge as key tools for addressing these inequalities. The literature suggests that interventions designed with an inclusive approach and tailored to the specific needs of the most vulnerable students can produce significant positive effects on their academic performance and overall well-being (OECD, 2020). These strategies contribute not only to the development of academic skills but also to the strengthening of self-efficacy and motivation, key determinants of educational success (Jabery & Arabiat, 2024; Mokotjo, 2024).

In this context, the present study evaluates the effectiveness of an educational leveling intervention aimed at improving the academic performance of low-performing students in Zone 6 of Ecuador. Specifically, the study pursues two objectives: (1) to estimate the impact of the intervention on the academic performance (GPA) of students in the lowest performance quartile, compared to students in the upper quartiles who did not receive the intervention; and (2) to examine how contextual, geographic, and temporal factors—namely region, school shift, educational level, and timing of intervention delivery—moderate the effectiveness of the intervention among low-performing students.

Research Questions

RQ1: How does the educational intervention affect the academic performance of students in the lowest quartile compared to those in the higher quartiles?

RQ2: Which contextual, geographic, and demographic factors influence the intervention's effectiveness within the low-performing group?

Zone 6 of Ecuador is a planning and administrative zone composed of the provinces of Azuay, Cañar, and Morona Santiago, located in the southern part of the country. The first two are in the Andean highlands, or mountainous region, of Ecuador, while the province of Morona

Santiago is in the Ecuadorian Amazon region, reflecting the area's marked geographic, ethnic, and cultural diversity.

In the educational field, the landscape is also heterogeneous and influenced by geographic, social, and economic factors. According to data from the Ministerio de Educación (2024), in the 2022-2023 academic year, there were 1,832 educational institutions in this area, serving a student population of 330,085 students, supported by 17,680 teachers. However, the unequal distribution of resources and territorial dispersion, especially in Amazonian regions such as Morona Santiago, hinder equitable access to quality educational services. These conditions generate particular challenges for students with low academic performance, who face not only pedagogical gaps but also structural limitations linked to territory, connectivity, and the availability of teacher support.

One of the key indicators for understanding educational inequalities in Ecuador is the promotion rate. During the 2019–2020 academic year, the country achieved a promotion rate of 97.91%, while the non-promotion and dropout rates stood at 0.36% and 1.73%, respectively (Ministerio de Educación, 2023). Although these figures reflect a positive trend at the aggregate level, school retention and success remain fragile in areas characterized by higher levels of poverty and limited access to educational infrastructure, such as many communities in Zone 6. Furthermore, the average student-teacher ratio in Ecuador was 20.06, exceeding the internationally recommended standard of 15 students per teacher, potentially limiting the individualized attention required by students experiencing academic lag (Ministerio de Educación, 2023). These conditions underscore the importance of implementing targeted pedagogical interventions that specifically address the needs of the lowest-performing quartile.

Beyond the theoretical framework outlined above, the following section reviews the empirical literature on educational leveling and remediation programs, with particular attention to studies conducted in basic education and in Latin American contexts. Among the studies referring to basic education, the work of Gallegos de Dios (2022), for example, sought to analyze the need for extraordinary educational programs, such as tutoring, private face-to-face or virtual classes, among others, to reduce school absenteeism and educational lag during the COVID-19 pandemic, emphasizing that this problem would generate a significant socio-educational gap, especially among those with limited financial means.

Sanz and López's (2021) research analyzed the impact of COVID-19 on elementary school students, highlighting how the pandemic intensified learning, digital, and academic gaps, particularly affecting the most vulnerable students. The identified consequences include learning loss, lack of basic skills, academic disengagement, increased dropout rates, and emotional difficulties. To address these issues, the authors propose strategies such as curricular flexibility, motivational methodologies, personalized teaching, strengthening family-school relationships, and emotional education programs. They also underscore the key role of teachers and the need to draw lessons from this crisis in order to face future educational challenges.

Along these lines, Giler-Medina (2023) highlighted the need to establish educational leveling strategies and instructional planning that incorporate instruments such as diagnostic assessment, the use of technology as a pedagogical tool, and processes focused on skill acquisition in specific areas such as numerical reasoning in mathematics, as key elements for improving academic performance. Likewise, the research by Contreras et al. (2023) concluded that learning loss among basic education students is evidenced by poor academic results, gaps in foundational knowledge, and socio-emotional problems such as depression, anxiety, and academic disengagement. Factors such as lost instructional time, pedagogical improvisation, socioeconomic inequalities, and a lack of family support networks have exacerbated these problems, particularly since the COVID-19 pandemic. To address these challenges, strategies such as academic remediation, updated educational models, teacher training in technological and socio-emotional skills, and increased investment in education have been proposed to ensure a comprehensive response to students' needs.

In Ecuador, Suárez and Cevallos (2023) studied the results of a pedagogical leveling and acceleration program for upper basic education students with academic lag, identifying that the economic situation, family context, and the effects of the pandemic, among other factors, acted as key determinants of such lag, and concluded that the leveling program achieved the reintegration of students into regular and remedial educational programs, highlighting the importance of promoting an equitable educational environment. Rosales-Reyes et al. (2024), on the other hand, highlight the importance of formative leveling in light of the variability in students' learning rhythms and styles, noting that strategies employed in these programs should be tailored to individual needs, particularly given the lasting effects of the pandemic on education.

The study by Mendoza-Castro and Arroba-Cárdenas (2021) analyzed the relationship between academic reinforcement and academic achievement in tenth-grade Basic Education students at the Unidad Educativa del Milenio "Ciudad de Pedernales." The results revealed a strong association between the two variables, with an 87% coefficient of determination, indicating that the appropriate application of academic reinforcement significantly reduces gaps in academic achievement. The authors underscored the importance of implementing reinforcement strategies in a timely and sustained manner to address educational gaps and improve student performance. Finally, with specific reference to strategies for educational leveling, Ávila (2024) found that the use of audiovisual resources as a teaching tool yielded favorable results in students' comprehension of the content, and proposed an approach encompassing the diagnosis, planning, implementation, and evaluation of academic leveling programs.

Complementarily, Paunesku et al. (2015) provide empirical evidence that brief online interventions focused on growth mindset and sense of purpose significantly improved academic performance among low-performing high school students. Their study demonstrated that scalable psychological interventions could effectively reduce gaps and improve academic

engagement, especially among students at risk, underscoring the potential of broadly implemented brief interventions. Similarly, the work by Sánchez-Almeida et al. (2021) in Ecuador reported positive outcomes from socio-academic interventions for vulnerable groups, finding that students who participated in preparatory leveling courses performed significantly better in mathematics and geometry than those without intervention. Their findings emphasize the importance of socio-academic accompaniment beyond initial access to education.

On the other hand, Arnaiz-Sánchez et al. (2020) demonstrated, using a quasi-experimental design, that collaborative methodologies and project-based learning positively affected students' academic performance, particularly in linguistic and mathematical competencies. This research supports the notion that inclusive and active pedagogical strategies substantially contribute to narrowing academic performance gaps, especially in disadvantaged contexts.

McEwan's (2015) meta-analysis also contributes to understanding effective interventions in basic education, highlighting that pedagogically oriented interventions, such as tutoring programs and targeted instructional improvements, yielded stronger results than resource-only interventions, such as school feeding programs or monetary transfers. Finally, Loyalka et al. (2015) provided evidence from rural schools in China, showing that well-structured tutoring programs significantly enhanced students' academic outcomes. Their findings suggest that targeted, cost-effective tutoring strategies can be especially impactful in low-resource educational contexts.

The study by de Bruijn and Meeter (2023) investigated the effectiveness of post-COVID-19 educational recovery programs implemented in Dutch secondary schools. Using a large sample of 16,675 students—of whom 1,336 participated in targeted support programs—the research evaluated the impact of various intervention types on students' academic performance. The findings revealed that specific, structured tutoring programs led to significant improvements in student grades, particularly in core subjects such as mathematics and language. In contrast, general support programs without individualized or subject-focused strategies showed limited or no measurable effect. These findings highlight the importance of precision and alignment between student needs and intervention design, suggesting that catch-up initiatives are most effective when they are tailored, intensive, and closely linked to curriculum goals. The study underscores the value of evidence-based practices in mitigating learning losses and provides strong empirical support for targeted tutoring as a post-pandemic recovery strategy in education systems.

Despite this growing body of evidence, the effectiveness of leveling programs in basic education remains underexplored, particularly in Latin American contexts and even more so in territories marked by deep geographic, ethnic, and cultural heterogeneity. Most available studies have been conducted in higher education, in urban settings, or in countries with comparatively well-resourced school systems, which limits their direct transferability to contexts such as Zone 6 of Ecuador. Furthermore, the post-pandemic landscape has deepened

existing social and educational gaps—evidenced in academic lag, learning loss, and students' socio-emotional difficulties—rendering the evaluation of targeted leveling interventions especially urgent. The present study contributes to closing this gap by providing empirical evidence on the impact of a structured leveling intervention implemented across diverse educational institutions in the Andean and Amazonian regions of southern Ecuador, using a quasi-experimental design and a multilevel difference-in-differences approach to account for the hierarchical nature of educational data.

METHODOLOGY

Design

This study used a quasi-experimental design with a non-equivalent control group because it is not possible to assign students to a control group at random in an educational setting (Shadish et al., 2002). We employed a longitudinal structure using grade point average as a measure of academic performance across two time points during the 2022–2023 and 2023–2024 academic years. Students in the lowest quartile formed the intervention group, while those in the remaining quartiles formed the control group. Given the nested grouping structure of the data, we analyzed the data using a multilevel hierarchical model.

Population and study sample

Education Zone 6 of Ecuador comprises 1,695 educational establishments serving a total student population of 89,854. From this population, the study included a final sample of 2,011 students enrolled in 4th, 7th, and 10th grades of Basic General Education and the 3rd year of Unified General Baccalaureate across 12 participating schools. The institutions were selected through a sampling strategy that we describe in the following section. The participating schools were selected purposively based on practical feasibility, and the 2,011 students were classified according to their prior academic performance into an intervention group which comprises the lowest quartile (n = 457) and a comparison group which comprises the three higher quartiles (n = 1,554).

Sampling

Due to practical and budgetary considerations, the educational intervention was implemented exclusively for students identified as low-performing academically. This decision was based on the pressing need for targeted support and the impracticability of forming a matched control group at the same performance level, as such a group would have been ethically problematic and disruptive to students' regular schooling. Furthermore, this strategy was consistent with the guidelines of the Zonal Coordination (Zone 6) of the Ecuadorian Ministry of Education, which called for resources to be directed toward the students most in need.

The study sample was structured into two main groups: the intervention group and the control group. The intervention group comprised students identified as being in the lowest quartile of academic performance. These students were selected because they were at high risk

of continued underachievement without targeted support, and the intervention was specifically designed to address their academic needs.

In contrast, the control group consisted of students from the other three academic performance quartiles, who did not receive the specific intervention. These students served as a reference for assessing the impact of the remedial program on the intervention group. Comparisons between groups allowed for an evaluation of the intervention's effectiveness without withholding support from the lowest-performing students, which would have been ethically problematic.

Sample

To determine the study sample of 2,011 students, Cochran's sampling formula was applied. This approach ensures statistical representativeness and allows sample results to be extrapolated to the total population with a high level of confidence. Based on Cochran's formula with a 98% confidence level and a 3% margin of error, a sample of approximately 1,484 students would be sufficient to obtain results representative of the population of 89,854 students. The use of 2,011 students provided an additional margin, strengthening the robustness of the results and ensuring that conclusions regarding the intervention's effectiveness are reliable and generalizable.

Table 1.

Characteristics of the group involved

		Control Group		**Intervention Group**		**Total**	
Variable	**Category**	**n**	**%**	**n**	**%**	**n**	**%**
Sierra Region	Amazon Region	423	21%	121	6%	544	27%
	Sierra Region	1131	56%	336	17%	1467	73%
Shift	Afternoon	351	18%	98	5%	449	22%
	Morning	1203	60%	359	18%	1562	78%
Grade	4th Basic Education	452	23%	132	7%	584	29%
	7th Basic Education	426	21%	117	6%	543	27%
	10th Basic Education	437	22%	132	7%	569	28%
	3rd High School	239	12%	76	4%	315	16%
Moment	January 2023	423	21%	121	6%	544	27%
	June 2023	1007	50%	298	15%	1305	65%
	October 2023	124	6%	38	2%	162	8%
Total		1554	77%	457	23%	2011	100%

The sample is distributed by region, shift, grade, intervention time, and intervention condition. The sample includes both the control and intervention groups to evaluate the effectiveness of an educational intervention for students in the lowest quartile of academic

performance. The sample includes students from the Amazon region (four institutions from the Sucúa canton in the province of Morona Santiago) and the Sierra region (two from the Tambo de Cañar canton, two from the Nabón canton in the province of Azuay, and four from the Hermano Miguel urban parish in the city of Cuenca). Students were distributed across two school shifts: morning and afternoon. The sample included students from four educational levels: 4th grade (last year of lower basic education), 7th grade (last year of middle basic education), 10th grade of Basic Education (last year of upper basic education), and 3rd year of Baccalaureate (last year of secondary school). The interventions were carried out at three different time points during the academic year: January, June, and October 2023. The first two occurred in the 2022–2023 school year, during which the intervention covered all Amazon institutions and basic education institutions in the Sierra region; the last was carried out in the 2023–2024 school year, targeting high school students in the Sierra region. Most of the interventions were concentrated in June, coinciding with the period prior to the final examinations (Table 1).

Data source and measurement

The grades used to calculate academic performance in this study were obtained directly from the records of each participating educational institution. The study did not employ a researcher-developed instrument. Instead of it, the data source consisted of official academic records maintained in accordance with the regulations of the Ecuadorian Ministry of Education and provided directly by the 12 participating schools. Accordingly, psychometric validation was not applicable; data verification focused on the completeness, consistency, and correspondence of the records with the relevant grades and subject areas. From these grades, performance was assessed in four core subjects as follows: natural sciences, mathematics, social sciences, and language and literature.

Research procedure

Fieldwork was carried out in three phases throughout 2023. In January, the intervention took place across four educational institutions in the canton of Sucúa—three in urban areas and one in the rural parish of Huambi—and targeted four educational levels. Preservice teachers from various degree programs within the Faculty of Philosophy, Literature, and Educational Sciences at the University of Cuenca participated in the implementation, applying instructional plans developed by the research team. Each session was conducted by two preservice teachers, working with small groups of up to 10 students from the lowest academic performance quartile in each class at both the primary and secondary levels. In June, the intervention was extended for one month across eight institutions in the Sierra region (Cuenca, Nabón, El Tambo, and Cañar), focusing on 4^{th}-, 7^{th}-, and 10^{th}-grade students. Finally, in October, the intervention targeted 3rd-year high school students in the same locations.

The intervention was implemented over 96 hours, with 20 hours dedicated to each of the four core subject areas. To support the program, 16 leveling guides were developed with the assistance of the University of Cuenca's Department of Community Engagement. Figure 1

shows the process of design and implementation of the leveling guides, including the core subjects, target grades, integrated subject areas, development and validation teams, and classroom coverage. These guides included four dedicated to Language and Literature for the 4th, 7th, and 10th grades of Basic General Education (EGB) and the 3rd grade of Unified General Baccalaureate (BGU); four for Social Sciences at the same levels; four for Mathematics; and four for Natural Sciences. The Natural Sciences guide for Baccalaureate consolidated subjects such as physics, chemistry, and biology, while the Social Sciences guide integrated philosophy, citizenship education, and history. The guides were jointly developed by 80 preservice teachers and 12 university professors from the University of Cuenca. To validate the guides, 16 expert teachers from the primary and secondary levels, specializing in the four subject areas, were engaged to provide feedback for revision. This material provided students with structured tools to address their academic weaknesses, facilitating a systematic, targeted approach to remediation and educational improvement. The 80 preservice teachers from the University of Cuenca implemented the guides across 75 student groups drawn from the 12 participating institutions.

Figure 1.

Creation process for leveling guides



The intervention sessions were conducted in three distinct periods during the 2022–2023 and 2023–2024 academic years: January (prior to the first-semester examination), June (prior to the second-semester examination), and October (at the start of the following academic year). These three implementation windows enabled the program to adapt to institutional schedules and to maximize both coverage and continuity of support for participating students.

Grade Point Average (GPA) is a robust, multidimensional indicator of students' future academic success. According to Easton et al. (2017), GPA is not only a strong predictor of high school graduation, but also of college enrollment and one-year college retention. Importantly, it outperforms standardized test scores in predicting these long-term outcomes. This predictive power stems from GPA's capacity to reflect not only academic achievement but also key behavioral traits such as effort, attendance, and engagement—factors that standardized tests often overlook. These findings demonstrate that GPA captures both content mastery and broader learning behaviors, making it a valid and meaningful measure of student performance. As such, GPA should be viewed not just as an evaluative tool, but as an early signal of students' academic trajectories and a basis for timely, targeted interventions.

Ethical considerations

To ensure participants' inclusion in the study, informed consent was obtained in advance from parents and school principals, authorizing the withdrawal of children from their regular classes for integration into the intensive remediation process.

Data analysis

To ensure the validity of the estimation and control for potential confounds, a Difference-in-Differences (DiD) model was implemented following the framework proposed by Schatschneider and Compton (2013).

Within the RStudio environment, mixed-effects models were estimated using the lmerTest package (Kuznetsova et al., 2017), which enables significance testing of fixed effects in the lme4 framework. This approach enables accurate comparisons of changes in academic performance between the intervention and control groups over time, adjusting for the initially lower academic performance observed in students from the intervention group. By isolating the interaction between time and intervention, the DiD approach helps distinguish the specific effect of the educational intervention from other external influences on academic outcomes.

The models included fixed effects for time, intervention condition, and the specific moments of intervention delivery, as well as their interactions. Random effects were specified at multiple levels, including grade level, educational institution, and student nesting, to capture hierarchical variability in the data. Model diagnostics and assumption checks were conducted in RStudio, including tests for normality of residuals (Kolmogorov–Smirnov and Shapiro–Wilk), independence (Durbin–Watson), and the inspection of standardized residual plots to confirm the absence of problematic outliers and heteroscedasticity.

Although the residuals did not follow a perfectly normal distribution, the literature supports the robustness of mixed models to such deviations, especially when random effects are appropriately specified, and the sample size is large. As Pinheiro and Bates (2009) argue, mixed-effects models remain reliable under modest violations of the normality assumption, since random effects absorb much of the unexplained variance. This is reinforced by Gelman and Hill (2006) and Raudenbush and Bryk (2001), who highlight the robustness of multilevel modeling under real-world data conditions.

Before implementing the mixed-effects model with a difference-in-differences (DiD) structure, we assessed key assumptions both graphically and statistically, including normality of residuals, homoscedasticity, independence, and the presence of extreme outliers.

RESULTS

To address two research questions stated in the Introduction (RQ1, concerning the intervention's differential effect on students in the lowest quartile relative to those in the higher quartiles, and RQ2, concerning the contextual, geographic, and demographic factors shaping its effectiveness within the low-performing group) our results are organized into three stages. First, we describe the evolution of the dependent variable concerning academic performance. Next, we present the statistical assumptions underpinning the use of multilevel mixed models. Finally, we elucidate the hypothesis test employing a difference-in-differences approach within a mixed-effects framework.

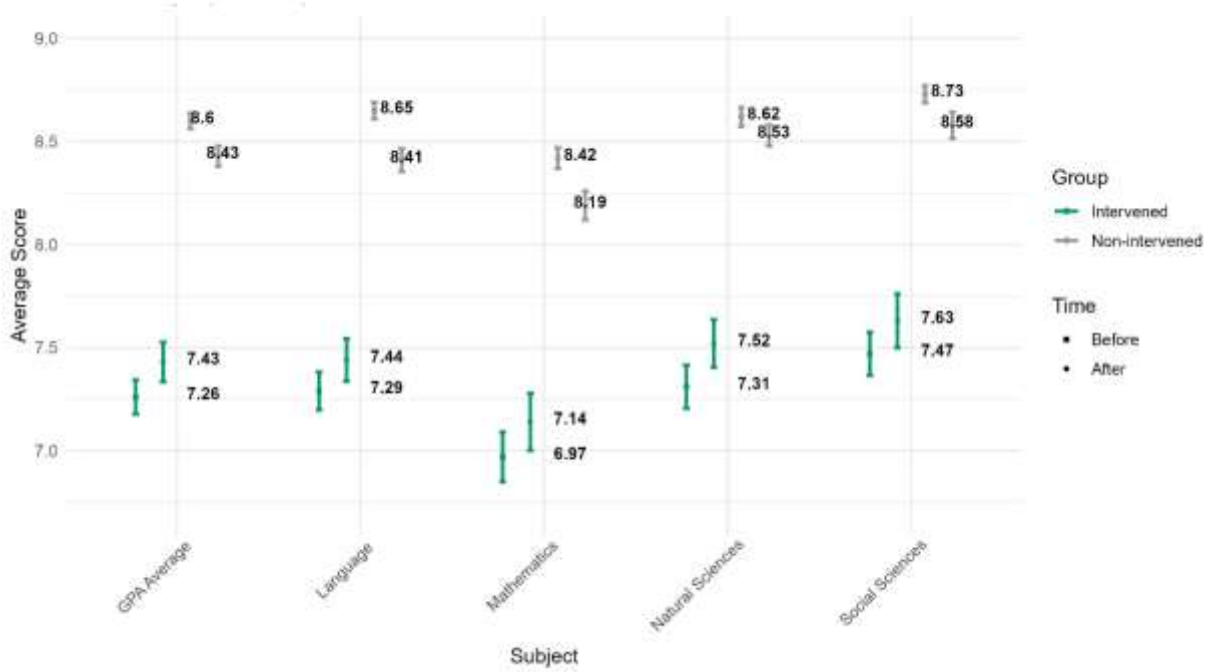
To address the two research questions stated in the Introduction—RQ1, concerning the intervention's differential effect on students in the lowest quartile relative to those in the higher quartiles, and RQ2, concerning the contextual, geographic, and demographic factors shaping its effectiveness within the low-performing group—our results are organized into three stages.

Descriptive statistics

Figure 2 displays the average scores before and after the intervention, disaggregated by group (intervention and control) and by subject: Overall GPA, Language and Literature, Mathematics, Natural Sciences, and Social Sciences. Vertical bars represent 95% confidence intervals.

Figure 2.

Academic performance comparison (pre and post), by group and subject



Overall, the control group—comprising students outside the lowest quartile—began with higher academic performance but showed a slight decline across all subjects from pre- to post-intervention. For instance, the average GPA dropped from 8.60 to 8.43, and Mathematics scores decreased from 8.42 to 8.19. In contrast, the intervention group—composed of students in the lowest performance quartile—started with lower baseline scores but demonstrated improvements in all subjects. The average GPA increased from 7.26 to 7.43, with more pronounced gains in Natural Sciences (from 7.31 to 7.52) and Mathematics (from 6.97 to 7.14). While modest, these gains suggest that the intervention helped mitigate academic decline, particularly relative to the downward trends observed in the control group.

Although the graph clearly visualizes the mean grades before and after the intervention for both groups, it is not sufficient on its own to draw conclusions about the intervention's effectiveness. DiD analysis provided a more rigorous, controlled approach by adjusting for baseline differences and external factors, enabling a more precise assessment of the educational intervention's impact.

Assumption test

The Kolmogorov–Smirnov test ($D = 0.026$, $p = 0.009$) and the Shapiro–Wilk test ($W = 0.976$, $p < 0.001$) indicated deviations from perfect normality, which are expected in large samples and are not considered problematic for linear mixed models. Standardized residual plots showed values centered around zero, with the vast majority falling within ±2, suggesting no problematic outliers. The residuals-versus-fitted-values plot exhibited a relatively constant spread, indicating acceptable homoscedasticity. Additionally, the Durbin–Watson test ($DW = 1.973$, $p = 0.101$) confirmed the absence of significant autocorrelation in the residuals. Based on these diagnostics, linear mixed-effects models with nested random effects were deemed appropriate for robustly estimating the intervention effect while accounting for the hierarchical structure of the data.

Hypothesis test

Table 2 presents the comparative performance of five hierarchical linear models estimated via restricted maximum likelihood (REML). These models tested varying combinations of fixed effects and interaction terms to evaluate the academic impact of a targeted educational intervention.

Model 1, the most complex specification, includes a three-way interaction between Time, Intervention, and Moment. Model 2 simplifies this by retaining only two-way interactions. Model 3 isolates the core interaction of interest—Time × Intervention—while Models 4 and 5 progressively reduce complexity, with Model 5 testing only the main effects of Time and Intervention.

In terms of conditional R^2 (variance explained by both fixed and random effects), Models 1 and 2 are equivalent ($R^2 = 0.527$), suggesting similar explanatory power. However, Model 2 is preferred due to its superior performance across multiple indices. It demonstrates the best trade-off between fit and parsimony, reflected in the highest AIC weight (0.770) and

Performance Score (91.63%), indicating that it is over nine times more likely than Model 1 to be the best model in the set. Moreover, it achieves a marginal R^2 of 0.266, signifying that its fixed effects alone explain approximately 26.6% of the variance in GPA.

In contrast, Model 3 (focused only on the Time × Intervention interaction) shows diminished explanatory power, particularly in marginal R^2 (0.230), and receives minimal weight across information criteria. Models 4 and 5—lacking interactions—perform poorly across all comparative metrics, confirming that modeling interactions is critical to capturing the complex dynamics of the intervention.

The Intraclass Correlation Coefficient (ICC) values indicate substantial clustering at the group level (ranging from 0.345 to 0.376), reinforcing the appropriateness of multilevel modeling. RMSE and residual standard deviations (Sigma) are consistently low across all models, indicating acceptable residual variability.

Table 2.

Model Information

Measure	Model 1	Model 2	Model 3	Model 4	Model 5
R^2 Conditional	0.527	0.527	0.519	0.51	0.514
R^2 Marginal	0.267	0.266	0.23	0.252	0.226
ICC	0.355	0.355	0.376	0.345	0.373
RMSE	0.726	0.727	0.735	0.738	0.738
Sigma	0.734	0.734	0.742	0.745	0.745
AIC Weights	0.23	0.77	2.02e-18	8.63e-26	2.07e-26
AICc Weights	0.227	0.773	2.1e-18	8.92e-26	2.16e-26
BIC Weights	0.000548	0.999	4.23e-10	7.74e-19	1.01e-16
Performance Score (%)	61.69	91.63	28.28	8.54	14.71

Note. AIC (Akaike Information Criterion) and *BIC* (Bayesian Information Criterion): both measure model fit, with lower values indicating better fit, though BIC penalizes model complexity more heavily. *R^2 Marginal* represents the proportion of variance explained by fixed effects only, while *R^2 Conditional* includes both fixed and random effects. *ICC* (Intraclass Correlation Coefficient) reflects the proportion of variance attributable to the grouping structure. *RMSE* and *Sigma* represent residual variation; lower values indicate better predictive accuracy.

Table 3 reports the omnibus F-tests and significance values (p) for each fixed effect in the models. These tests assess whether each variable (or interaction) significantly contributes to explaining variance in the outcome variable, GPA.

In Model 1, the three-way interaction term (Time × Intervention × Moment) is not statistically significant ($F = 1.56$, $p = .212$), suggesting that the added complexity does not provide meaningful explanatory power. This is consistent with its lower ranking on information criteria (see Table 2). However, all two-way interactions—Time × Intervention, Time × Moment, and Intervention × Moment—are significant across Models 1 and 2 ($p < .001$ to $p = .018$), affirming their importance.

The variable Moment, treated as a fixed effect comparing different academic cycles (2022–2023 vs. 2023–2024), is not significant in Models 1 and 2, further questioning the need for its inclusion as a main effect when interactions are already modeled. Nevertheless, Time and Intervention remain highly significant across all models (F values from 17.84 to over 410.28, all $p < .001$), underscoring the robustness of their influence.

In summary, Table 3 validates the statistical relevance of the two-way interaction structure implemented in Model 2, while providing no support for higher-order terms or a simplified main-effects-only approach.

Table 3.

Fixed Effects and Omnibus Tests

Effect	M 1 (F, p)	M 2 (F, p)	M 3 (F, p)	M 4 (F, p)	M 5 (F, p)
Time	98.35, <.001	106.45, <.001	45.00, <.001	17.84, <.001	17.84, <.001
Intervention	285.34, <.001	410.28, <.001	1215.60, <.001	1827.50, <.001	1827.60, <.001
Moment	0.96, .385	1.82, .164	—	2.92, .058	—
Time × Intervention	17.99, <.001	38.96, <.001	38.96, <.001	—	—
Time × Moment	41.55, <.001	52.24, <.001	—	—	—
Intervention × Moment	3.21, .040	4.32, .018	—	—	—
Time × Intervention × Moment	1.56, .212	—	—	—	—

Note. Each cell reports the F statistic, followed by its p-value. These omnibus tests evaluate the contribution of each fixed effect (and its interactions) to explaining variability in GPA. Higher F values indicate stronger explanatory power, while lower p values denote greater statistical significance. A $p < .001$ suggests that the effect is highly significant.

Table 4 presents the estimated coefficients and standard errors for all fixed effects included in the five models. These coefficients quantify the direction and magnitude of the relationships between predictors and student academic performance, measured by GPA.

Across all models, the Intercept term is stable (ranging from 8.38 to 8.66), representing the expected GPA at the baseline (pre-intervention, control group, first academic cycle). The Time coefficient is consistently negative, indicating a general decline in academic performance over time in the absence of any intervention. This decline is most pronounced in Model 1 ($\beta = -0.50$) and attenuates in reduced models, confirming the moderating role of intervention timing.

The Intervention coefficient is significantly negative in all models (ranging from -1.21 to -1.38), reflecting the initially lower performance levels of students selected for the intervention. However, the Time × Intervention interaction is consistently positive ($\beta = 0.34$ to 0.45), suggesting that, although initially behind, students in the intervention group improved at a faster rate over time than their control-group peers.

In Model 2, the Time × Moment (June–October) interaction yields a coefficient of 0.62 (SE = 0.09), indicating that interventions administered in the later cycle were associated with a steeper positive trajectory in academic performance over time. The interaction Intervention × Moment (January–June) is significantly negative (-0.23, SE = 0.06), suggesting that early implementation was less effective for the treated group—possibly due to planning challenges during the program's initial phase and the difficulty of applying the program in the Amazon region over a period of only 15 days.

Finally, the three-way interaction terms in Model 1, though theoretically rich, are not significant, which aligns with the model selection results in Table 2 and justifies their exclusion in Model 2.

Table 5 displays the estimated variances and corresponding standard deviations for each of the random effects included in the five hierarchical models: Nest (student groups), UE (educational institutions), Grade (academic level), and Residual variance. These components capture the multilevel structure of the data and help explain the variability in student GPA beyond the fixed effects.

Across all models, the Nest-level variance (i.e., the grouping of students within the same intervention cohort) is consistently high, with values around 0.15 and standard deviations close to 0.39. This indicates that a substantial portion of the variability in academic performance is attributable to differences between these small instructional groups. Given that 80 preservice teachers across 75 groups delivered the intervention, this result reflects the heterogeneity in delivery quality, group dynamics, and possibly local contextual conditions (e.g., classroom environment, student motivation, or teacher preparedness).

The UE-level variance (institution-level differences) ranges from 0.05 to 0.07, with standard deviations ranging from 0.21 to 0.27. This suggests moderate variability in student GPA across the 12 participating institutions. While this variance is less pronounced than at the group level, it highlights institutional effects—such as leadership quality, infrastructure, school culture, or availability of support services—that may influence academic achievement.

The grade-level variance remains relatively stable across all models, ranging from 0.10 to 0.11 (standard deviations ~0.31–0.33). This reflects differences in GPA attributable to students' academic levels—specifically between the 4th, 7th, and 10th grades of Basic Education and the 3rd year of Baccalaureate (secondary education). These differences likely stem from developmental factors, curricular complexity, and baseline academic expectations, with younger students tending to respond more positively to interventions than older students, as supported by the fixed effects analysis.

Lastly, the Residual variance, which captures the portion of GPA variability unexplained by both fixed and random effects, is consistently around 0.54-0.56, with associated standard deviations of approximately 0.73-0.75. These values indicate that although the models capture considerable structured variation across levels, individual-specific factors—such as personal

motivation, family environment, health status, or unmeasured cognitive abilities—still play a meaningful role in shaping student performance.

Table 4.

Estimated Fixed Effects Parameters

Name	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept)	8.55 (0.21)	8.53 (0.21)	8.66 (0.19)	8.38 (0.21)	8.62 (0.19)
Time	-0.50 (0.05)	-0.48 (0.05)	-0.18 (0.03)	-0.10 (0.02)	-0.10 (0.02)
Intervention	-1.29 (0.08)	-1.23 (0.06)	-1.38 (0.04)	-1.21 (0.03)	-1.21 (0.03)
Moment (January-June)	0.14 (0.18)	0.16 (0.18)	—	0.30 (0.18)	—
Moment (June-October)	0.29 (0.24)	0.31 (0.24)	—	0.61 (0.23)	—
Time × Intervention	0.45 (0.11)	0.34 (0.06)	0.35 (0.06)	—	—
Time × Moment (January-June)	0.41 (0.06)	0.38 (0.05)	—	—	—
Time × Moment (June-October)	0.67 (0.11)	0.62 (0.09)	—	—	—
Intervention × Moment (January-June)	-0.16 (0.09)	-0.23 (0.06)	—	—	—
Intervention × Moment (June-October)	0.17 (0.16)	0.06 (0.11)	—	—	—
Time × Intervention × Moment (January-June)	-0.14 (0.13)	—	—	—	—
Time × Intervention × Moment (June-October)	-0.22 (0.22)	—	—	—	—

Note. Each cell reports the estimated regression coefficient followed by its standard error in parentheses. Coefficients represent the expected change in GPA associated with a one-unit increase in the corresponding variable, holding all others constant. Positive values indicate a favorable effect on academic performance, while negative values suggest a detrimental effect. Interaction terms quantify how the effect of one variable (e.g., Time) varies across levels of another variable (e.g., Intervention or Moment).

In sum, the random components analysis underscores the need for multilevel modeling in this educational context. The strongest sources of between-group variability occur at the group (Nest) level, which justifies the focus on classroom-based interventions. Institutional and grade-level differences are also meaningful but secondary. The residual variability confirms that both measurable and unobservable individual characteristics shape academic performance, further reinforcing the importance of personalized support strategies within broader remedial frameworks.

Table 5.

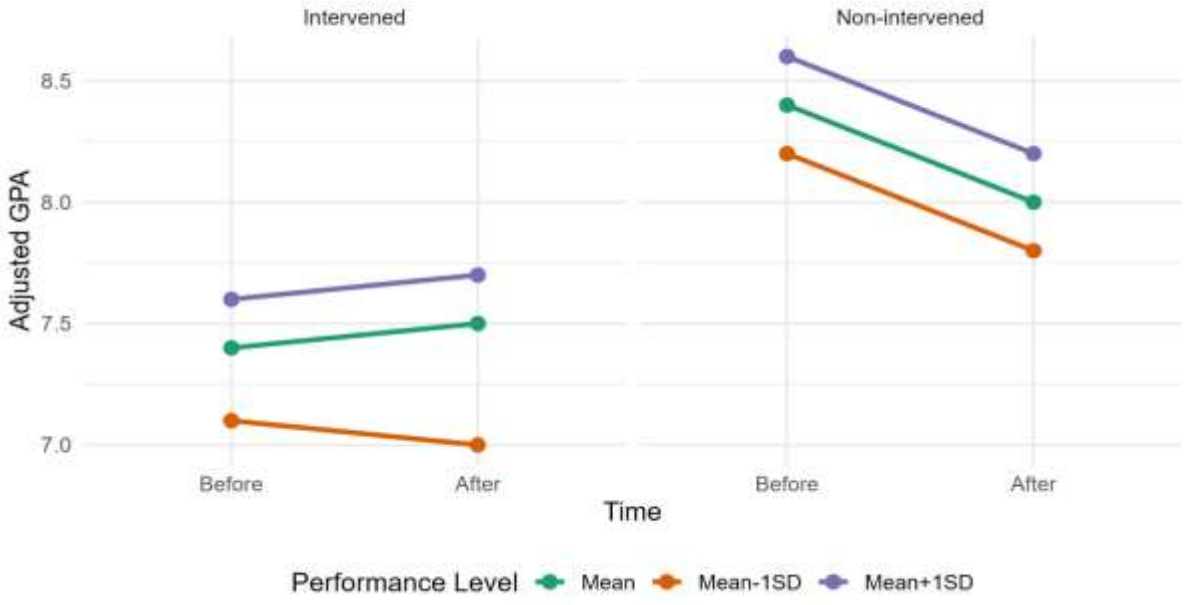
Random Components

Cluster	Model 1	Model 2	Model 3	Model 4	Model 5
Nest	0.151 (0.39)	0.151 (0.39)	0.148 (0.38)	0.151 (0.39)	0.148 (0.38)
	0.047 (0.22)	0.047 (0.22)	0.071 (0.27)	0.045 (0.21)	0.071 (0.27)
UE	0.099 (0.31)	0.099 (0.31)	0.112 (0.33)	0.096 (0.31)	0.111 (0.33)
	0.539 (0.73)	0.538 (0.73)	0.550 (0.74)	0.555 (0.75)	0.555 (0.75)
Grade	0.151 (0.39)	0.151 (0.39)	0.148 (0.38)	0.151 (0.39)	0.148 (0.38)
	0.047 (0.22)	0.047 (0.22)	0.071 (0.27)	0.045 (0.21)	0.071 (0.27)
Residual	0.099 (0.31)	0.099 (0.31)	0.112 (0.33)	0.096 (0.31)	0.111 (0.33)
	0.539 (0.73)	0.538 (0.73)	0.550 (0.74)	0.555 (0.75)	0.555 (0.75)

Note. UE: Educational institutions or units (Unidades Educativas). The main value represents the variance between groups (e.g., classes or schools), indicating how much they differ from one another. The number in parentheses indicates the standard error, which reflects the precision of this estimate (smaller values indicate greater precision). The ICC percentage % quantifies how much of the total variability is attributable to these group differences. The residual variance captures unexplained variability, representing factors not accounted for in the model.

Figure 3.

Adjusted academic performance (GPA) over time by initial performance level



Note. Each line represents the trajectory of adjusted academic performance (GPA) over time for subgroups defined by their relative performance level within the group (mean ±1 standard deviation). The horizontal axis is centered to facilitate the interpretation of pre- and post-

intervention changes. Estimates are derived from a mixed-effects model including interaction terms and hierarchical controls.

Figure 3 shows the evolution of adjusted GPA over time, comparing the intervention and control groups. Within each group, three subgroups are displayed: students performing at the mean (green line), one standard deviation below the mean (orange line), and one standard deviation above the mean (purple line). The values represent estimates from Model 5, adjusted for hierarchical effects.

In the intervention group, all subgroups exhibit a slight upward trend, particularly students with initially lower performance (Mean –1SD), suggesting that the intervention helped prevent further academic decline. In contrast, the control group shows a consistent downward trajectory across all subgroups, with the steepest drop observed among those starting below the mean. This divergence between groups supports the notion that the intervention not only helped improve academic outcomes in absolute terms but also buffered against the sharper decline observed in similar populations without intervention.

DISCUSSION

The results of this study offer a substantive basis for evaluating the effectiveness of educational interventions in relation to the existing literature. Prior theory and research indicate that targeted educational interventions are often shaped by a combination of contextual, structural, and methodological factors (Dussel, 2014; Quiroga et al., 2025). This is reflected in the divergent academic performance (GPA) trajectories observed between the intervention and control groups.

It is noteworthy that the coefficient for the Time × Intervention interaction was positive and statistically significant across all evaluated models, indicating that, despite an initial negative impact, the intervention's effect progressively improved over time. This finding is consistent with the arguments of Suárez-Guerrero et al. (2020), who contend that the effectiveness of educational interventions depends on their sustainability and adaptability over time. Similarly, Loyalka et al. (2015) emphasize the progressive nature of targeted interventions, noting significant improvements in student performance from sustained, structured tutoring programs, despite initial adaptation challenges.

The initial decline in performance likely reflects students' difficulty in adapting to unfamiliar pedagogical approaches, particularly in academically vulnerable contexts such as the Amazon region. From the perspective of Cognitive Load Theory (Sweller, 1988; Sweller et al., 2011), the early phase of the intervention may have imposed an additional extraneous load on students who were simultaneously processing new instructional formats, unfamiliar tutors, and consolidated subject content—conditions under which working memory becomes more easily overwhelmed. As students became familiar with the structure of the leveling sessions, this extraneous load presumably decreased, allowing germane cognitive resources to be redirected toward learning. The subsequent positive trajectory also resonates with Hattie and Timperley's

(2007) model of feedback: the small-group, tutor-mediated format of the intervention enabled preservice teachers to provide timely, targeted feedback at the *task* and *process* levels, which the literature identifies as the most effective forms for narrowing performance gaps. This interpretation is empirically reinforced by Paunesku et al. (2015), who found that brief, well-targeted interventions significantly improved academic outcomes over time among low-performing students.

The mixed-model analysis conducted in RStudio using the lme4, lmerTest, and car packages compared five hierarchical models. Model 2, which retained the two-way interactions Time × Intervention and Time × Moment while excluding the non-significant three-way interaction, was selected as the preferred specification based on its superior performance across multiple information criteria (AICc weight = 0.773; BIC weight = 0.999; Performance Score = 91.63%). Beyond statistical fit, Model 2 also offers the greatest conceptual clarity for answering the core research questions, as it captures both the differential trajectory of the intervention group over time and the moderating role of the timing of intervention delivery, while controlling for institutional and grade-level clustering. Specifically, the significant Time × Intervention interaction ($\beta = 0.34$, SE = 0.06, $p < .001$ in Model 2) suggests that the cumulative effect of pedagogical strategies may not be evident in the short term, underscoring the importance of conducting longitudinal evaluations. Additionally, the significant Time × Moment interaction ($\beta = 0.62$, SE = 0.09 for June–October) reinforces the presence of a differential impact by the timing of the intervention. This finding aligns with McEwan's (2015) meta-analysis, which indicates that pedagogically oriented interventions yield stronger effects when applied consistently over time rather than as isolated short-term actions, and with de Bruijn and Meeter's (2023) evidence that structured tutoring programs produce more robust learning gains when their timing is aligned with curricular milestones.

The study also highlighted the importance of considering variability across student groups, institutions, and geographic regions. As shown in Table 5, variability across student groups (Nest) is particularly notable, with a high intraclass correlation coefficient (ICC) in Model 5. This suggests that within-group differences are a critical factor in academic performance, consistent with Graizer (2018), who highlighted that structural inequalities within educational institutions significantly impact learning outcomes. Similarly, Arnaiz-Sánchez et al. (2020) emphasize that collaborative methodologies and inclusive pedagogical strategies effectively reduce academic gaps, addressing institutional and contextual differences.

Although the residuals showed slight deviations from normality, this did not compromise the analysis because linear mixed-effects models are robust under such conditions, especially when sample sizes are large and random effects are correctly specified (Gelman & Hill, 2006; Pinheiro & Bates, 2009).All other assumptions were fully met. This approach captures the hierarchical structure of the data and addresses variability at multiple levels, reinforcing the reliability of the conclusions obtained (Maggio, 2012).

The differential effects observed across regions, educational levels, and intervention timing in this study can be read as empirical traces of structural inequalities that operate beyond the classroom. The weaker intervention effect observed in the Amazon region, where the program was implemented under condensed timeframes and across geographically dispersed institutions, is consistent with evidence that territorial dispersion and uneven distribution of pedagogical resources lead to accumulated learning disadvantages (Murillo, 2007). Similarly, the stronger gains observed at lower educational levels (4th and 7th grades) compared to upper basic education and Baccalaureate align with longitudinal evidence that the cumulative impact of structural disadvantage on academic trajectories tends to widen with age (Dickerson & Popli, 2016; Nikulina et al., 2011), making early targeted interventions more effective than later remediation. The substantial group-level variability captured by the random effects (Nest variance ≈ 0.151 across all models) further suggests that the conditions under which the intervention is delivered—group composition, tutor preparedness, and classroom environment—mediate its effectiveness in ways that purely individual-level explanations cannot fully account for. These patterns motivate a broader reading of the results in light of the literature on the structural effects of poverty and cultural reproduction on academic performance.

An extensive literature on the structural effects of poverty on academic performance supports this broader reading. Reimers (2000) argues that poverty directly affects students' health and nutrition, leading to reduced learning capacity and limited academic progress. Furthermore, students from poor households often have less time for school activities due to their early entry into the workforce, which constrains the development of foundational skills such as reading, writing, and arithmetic—skills that are essential for success at higher levels of the education system (Santillán Hernández & Vargas Sánchez, 2022). Accumulated evidence indicates that children in poverty tend to achieve lower learning outcomes than their more advantaged peers, even when other contextual variables are controlled for (Murillo, 2007).

Furthermore, poverty has cumulative impacts from childhood, negatively affecting cognitive development (Dickerson & Popli, 2016) and increasing the likelihood that adults who experienced poverty in childhood will remain poor in adulthood (Nikulina et al., 2011). These effects are not only economic but also symbolic and cultural. According to Bourdieu (1986), the cultural capital transmitted in homes shapes teachers' perceptions and the attention students receive, which in turn affects their academic performance. This mechanism of cultural reproduction has been formalized by Jæger and Breen (2016), who explain how students'socioculturalenvironmentconditionsteachers'expectationsandattitudes. At the emotional level, Eamon (2002) notes that families under greater economic pressure tend to provide less stimulating environments and to be less involved in schooling, which exacerbates inequalities. Consequently, the intervention evaluated in this study should be understood not only as a pedagogical tool but as part of a broader effort to mitigate the multidimensional effects of poverty on learning. Sánchez-Almeida et al. (2021) explicitly reinforce this perspective,

highlighting the importance of socio-academic interventions that integrate both educational and psychosocial support to improve outcomes among vulnerable populations.

These findings have significant implications for the design and implementation of educational interventions. On the one hand, they emphasize the need for ongoing support and adaptive adjustments to ensure that pedagogical strategies benefit the most vulnerable students. On the other hand, they underscore the importance of evaluating medium- and long-term impacts, avoiding premature interpretations based on initial results.

Limitations of the Study

An important limitation of this study is the presence of floor and ceiling effects in grades, where students in the lowest quartile of academic performance face a "floor effect" because grades cannot fall below 6 points.[1]—a constraint that paradoxically prevents the observation of downward performance variations in this group, even when the general trend is declining. In contrast, students in the upper quartiles, particularly those in the highest, experience a "ceiling effect," in which their grades approach the maximum possible, limiting the ability to detect significant improvements. These grading boundaries may bias the evaluation of the intervention's impact, as they constrain the observable range of performance variation. Furthermore, consolidating Baccalaureate-level subjects into broad areas such as the natural and social sciences may not fully capture discipline-specific competencies, thereby affecting the precision of the analysis. The generalizability of the results is also limited by contextual and geographical differences among participating institutions and by variability in intervention implementation. Finally, uncontrolled external factors—such as family events, school transitions, and socioeconomic circumstances—could have influenced students' academic performance.

It is also important to reflect on how the leveling guides could be improved in future project replications. One of the main lessons learned was that the guides, although carefully designed, were underutilized in practice. This was largely due to the need to adapt them dynamically to each group's specific characteristics. As implementation progressed, preservice teachers increasingly relied on direct instruction methods accompanied by immediate, formative feedback, which appeared to be more suitable for addressing students' immediate academic difficulties. This approach aligns with Sweller's Cognitive Load Theory (1988), which emphasizes reducing extraneous cognitive demands in instructional materials, and with Hattie and Timperley's (2007) model of feedback, which identifies timely, targeted feedback as a powerful mechanism for improving learning. Future leveling interventions should incorporate more flexible and adaptable instructional materials, which will allow teachers to adjust their use according to students’ prior knowledge, educational level, and contextual needs.

[1]There is a provision from the Ministry of Education that children cannot fail a year of initial and secondary education.

These findings help close the gap identified at the outset of this study: unlike prior evaluations of leveling and remediation programs, most of which were conducted in urban, better-resourced, or higher-education settings, this study provides quasi-experimental evidence from a geographically and culturally heterogeneous region combining Andean and Amazonian territories, where such interventions had not previously been evaluated using a multilevel difference-in-differences approach. This design exanimated changes associated with the intervention considering the hierarchical structure of the data and examining variations across timing, educational level, and territorial context. In this sense, the novelty of the study lies less in demonstrating that leveling interventions can work, and more in showing how their effectiveness is conditioned by timing, educational level, and territorial context—precisely the moderating factors that motivated the study's second research question. For future research, these results point to at least three concrete directions: (a) longitudinal designs that track students beyond a single academic year to determine whether the buffering effect observed here translates into sustained gains; (b) qualitative studies examining why leveling guides were underutilized in practice, in order to inform the redesign of instructional materials for greater adaptability; and (c) disaggregated, subject-specific analyses that avoid consolidating Baccalaureate-level areas, thereby capturing discipline-specific effects that the present study's aggregated approach could not detect.

CONCLUSIONS

This study evaluated the effectiveness of an educational leveling intervention targeting students in the lowest-performing quartile across 12 institutions in Zone 6 of Ecuador, using a quasi-experimental design and a multilevel difference-in-differences approach. The findings allow us to return to the two objectives that guided the study. Regarding the first objective—estimating the impact of the intervention on the academic performance of students in the lowest performance quartile—the results provide evidence of moderate but meaningful effectiveness. Although the overall GPA of the intervention group declined slightly in absolute terms, the significant and positive Time × Intervention interaction ($\beta = 0.34$, $p < .001$ in Model 2) indicates that students in the intervention group followed a more favorable trajectory than their control-group peers, who showed a steeper decline across all subjects. In other words, the intervention did not produce dramatic absolute gains. However, it did buffer low-performing students against the broader downward trend observed in the control group, a relevant outcome in a post-pandemic context marked by widespread learning loss. Regarding the second objective—examining how contextual, geographic, and temporal factors moderate the effectiveness of the intervention—four patterns emerged. First, the *timing* of intervention delivery significantly shaped its effects: sessions implemented in June, immediately before final examinations, were associated with stronger gains than those delivered in January or October, suggesting that alignment with curricular milestones amplifies the impact of leveling programs. Second, *educational level* moderated the effect: students in lower grades (4th and 7th of Basic

Education) showed greater improvement than those in upper basic education and Baccalaureate, consistent with evidence that early targeted intervention is more effective than later remediation. Third, *regional context* mattered: students in the Sierra region showed relatively stronger outcomes than those in the Amazon region. However, this difference is partly attributable to implementation conditions in the Amazon—where the program began under condensed timeframes and faced specific logistical and territorial challenges—rather than to the intervention model itself. Fourth, *school shift* also moderated the effect, with students in the morning shift benefiting more than those in the afternoon shift.

These findings indicate that targeted leveling interventions can produce meaningful effects on the academic trajectories of low-performing students, but only when their design accounts for the timing of delivery, students' developmental stage, and the territorial conditions under which they are implemented. The results support a conception of leveling not as a uniform remedial protocol, but as a contextually adaptive strategy that must be aligned with curricular cycles, prioritized at lower educational levels, and accompanied by adequate planning and resourcing in geographically dispersed contexts. From a policy perspective, the evidence suggests that resource allocation for leveling programs in Ecuador and similar contexts should privilege early-grade interventions delivered in proximity to evaluation cycles, while explicitly addressing the structural conditions that constrain implementation in Amazonian and rural territories. Sustaining this kind of targeted allocation, however, depends on the stability of the policy environment in which it is embedded: comparative evidence indicates that education systems subject to frequent shifts in funding priorities and governance show more disrupted outcomes for vulnerable students than those operating under stable, consistently resourced policy frameworks (Mountford, 2025), reinforcing the case for multi-year institutional commitments to leveling programs rather than one-off interventions.

REFERENCES

Ajani, O., Govender, S., & Gamede, B. (2025). Investigating Emerging Technologies for Curriculum Delivery in a Rural South African University through the Lens of Diffusion of Innovation. *Journal of Curriculum Studies Research, 7*(2), 394-411. https://doi.org/10.46303/jcsr.2025.27

Arnaiz-Sánchez, P., de Haro, R., Alcaraz, S., & Mirete Ruiz, A. B. (2020). Schools that promote the improvement of academic performance and the success of all students. *Frontiers in Psychology, 10*, Article 2920. https://doi.org/10.3389/fpsyg.2019.02920

Ávila, M. (2024). Los recursos audiovisuales en el proceso de nivelación en las ciencias naturales para estudiantes de 10° grado de educación general básica. *Revista Científica Arbitrada Multidisciplinaria PENTACIENCIAS*, *6*(4), 257–269. https://doi.org/10.59169/pentaciencias.v6i4.1157

Bandura, A. (1997). Self-efficacy: The exercise of control. W. H. Freeman.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241–258). Greenwood.

Bronfenbrenner, U. (1979). The ecology of human development: Experiments by nature and design. Harvard University Press.

Contreras, J. A., Vera, J. A., & Villamizar, C. N. (2023). Pérdida de aprendizajes en estudiantes de educación básica en tiempos de pandemia: Una secuela académica durante el retorno a la presencialidad. *Dialogus*, *6*(11), 168–177.https://doi.org/10.37594/dialogus.v1i11.799

de Bruijn, A. G. M., & Meeter, M. (2023). Catching up after COVID-19: Do school programs for remediating pandemic-related learning loss work? *Frontiers in Education, 8*, Article 1298171. https://doi.org/10.3389/feduc.2023.1298171

Dickerson, A., & Popli, G. (2016). Persistent poverty and children's cognitive development: Evidence from the UK Millennium Cohort Study. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(2), 535–558.https://doi.org/10.1111/rssa.12128

Dussel, I. (2014). ¿Qué lugar tiene la escuela media en la producción de la desigualdad? Elementos para el debate. *Archivos Analíticos de Políticas Educativas*, *22*(43). https://doi.org/10.14507/epaa.v22n43.2014

Dyantyi, N., & Mkabile-Masebe, B. (2025). Closing the digital divide: Strategies for addressing inequalities in technology access in higher education institutions. *Research in Social Sciences and Technology, 10*(1), 412–425. https://doi.org/10.46303/ressat.2025.22

Eamon, M. (2002). Effects of poverty on mathematics and reading achievement of young adolescents. *Journal of Early Adolescence*, 22(1), 49–74. https://doi.org/10.1177/0272431602022001003

Easton, J. Q., Johnson, E., & Sartain, L. (2017). *The predictive power of ninth-grade GPA*. University of Chicago Consortium on School Research. https://consortium.uchicago.edu/publications/predictive-power-ninth-grade-gpa

Gallegos de Dios, O. (2022). Ausentismo, deserción escolar y rezago educativo en secundarias públicas en México durante la pandemia del Covid-19. *Sincronía*, *81*, 725-745. http://sincronia.cucsh.udg.mx/pdf/81/725_745_2022a.pdf

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press

Giler-Medina, P. (2023). Estrategias de nivelación y desarrollo del pensamiento numérico en estudiantes de educación secundaria. *Identidad Bolivariana*, *7*(1), 48-59. https://identidadbolivariana.itb.edu.ec/index.php/identidadbolivariana/article/view/205

Graizer, O. (2018). Sociología de la educación: tensiones en el campo educativo, en: Walker, V.; Alzamora, S.; Rosales, G. *Encuentro de cátedras de Sociología de la Educación*. Bahia Ediuns, 2018. Libro Digital. pp. 11-16. http://www.ediuns.uns.edu.ar/Files/Encuentro_Sociologia_Educacion.pdf

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Jabery, M.A.AL & Arabiat, D.H. (2024). Nurturing School Well-being: Insightful Perspectives on Key School Features for Students with Autism, *Journal of Social Studies Education Research, 15*(3), 88-117. https://jsser.org/index.php/jsser/article/view/5725/681

Jæger, M. M., & Breen, R. (2016). A dynamic model of cultural reproduction. *American Journal of Sociology*, 121(4), 1079–1115.https://doi.org/10.1086/684012

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Loyalka, P., Wei, J., Song, Y., Zhong, W., & Chu, J. (2015). The impacts of building elite high schools for students from disadvantaged areas. *Economic Development and Cultural Change, 63*(2), 393–422. https://doi.org/10.1086/678992

Maggio, M. (2012). *Enriquecer la enseñanza: Los ambientes con alta disposición tecnológica como oportunidad*. Paidós.

Mapisa, B., & Makena, B. (2024). The Impact of ICT Adoption in Enhancing Teaching and Learning in Primary Schools of Amathole East District, Eastern Cape. *Research in Social Sciences and Technology, 9*(1), 213-231. https://doi.org/10.46303/ressat.2024.12

McEwan, P. J. (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research, 85*(3), 353–394. https://doi.org/10.3102/0034654314553127

Mendoza-Castro, A. M., & Arroba-Cárdenas, Á. E. (2021). Refuerzo académico y rezago escolar en estudiantes de educación básica superior. *Polo del Conocimiento*, *6*(4), 54–72.https://doi.org/10.23857/pc.v6i4.2538

Ministerio de Educación. (2023). *Estadística educativa volumen 4: Matrícula, docentes y establecimientos por nivel y modalidad educativa 2022–2023*. https://educacion.gob.ec/wp-content/uploads/downloads/2023/11/Estadistica-Educativa_Volumen-4.pdf

Ministerio de Educación. (2024). *Informe de gestión Zonal Zona 6, 2022–2023*. https://educacion.gob.ec/wp-content/uploads/downloads/2024/03/CZ6-Informe.pdf

Mokotjo, L. (2024). Fostering Inclusivity: A Critical Emancipatory Approach to Dyscalculia in Primary School Mathematics. *Research in Educational Policy and Management, 6*(2), 194-208. https://doi.org/10.46303/repam.2024.31

Mountford, M. (2025). The impact of partisanship on student performance in politically divided states. *Research in Educational Policy and Management, 7*(2), 56–72. https://doi.org/10.46303/repam.2025.10

Murillo, F. J. (Coord.). (2007). *Investigación iberoamericana sobre eficacia escolar*. Convenio Andrés Bello.

Nikulina, V., Widom, C. S., & Czaja, S. (2011). The Role of Childhood Neglect and Childhood Poverty in Predicting Mental Health, Academic Achievement and Crime in Adulthood. *American Journal of Community Psychology*, *48*(3–4), 309–321. https://doi.org/10.1007/s10464-010-9385-y

OECD. (2018). *Can equity in education foster social mobility?* OECD Publishing. https://www.oecd-ilibrary.org/content/dam/oecd/es/publications/reports/2018/10/can-equity-in-education-foster-social-mobility_169073f0/9335f0fa-es.pdf

OECD (2020). *Education responses to COVID-19: Embracing digital learning and online collaboration*. OECD Publishing.

Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement.*Psychological Science, 26*(6), 784–793. https://doi.org/10.1177/0956797615571017

Pinheiro, J. C., & Bates, D. (2009). *Mixed-Effects Models in S and S-PLUS*. Springer Science & Business Media.

Quiroga, A., Bracchi, C. C., & Parma, G. (2025). La teoría de los campos y las teorías socioeducativas: una presentación. *Libros de Cátedra*.

Raudenbush, S. W., & Bryk, A. S. (2001). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage

Reimers, F. (2000). ¿Pueden aprender los hijos de los pobres en las escuelas de América Latina? *Revista Mexicana de Investigación Educativa*, *5*(9), 11–69. http://www.redalyc.org/articulo.oa?id=14000902

Rosales-Reyes, L., Vásconez-Vásconez, A., & Zúñiga-Delgado, M. S. (2024). La nivelación en los aprendizajes de los estudiantes de octavo año de educación básica de la Unidad Educativa Teodoro Wolf post pandemia - Guía de recopilación de estrategias y técnicas de nivelación en Lecto-Escritura y Razonamiento Lógico Matemático. *MQRInvestigar*, *8*(1), 1863–1877. https://doi.org/10.56048/mqr20225.8.1.2024.1863-1877

Sánchez-Almeida, T., Naranjo, D., Gilar-Corbi, R., & Reina, J. (2021). Effects of socio-academic intervention on student performance in vulnerable groups. *Sustainability, 13*(14), 7673. https://doi.org/10.3390/su13147673

Santillán Hernández, A. S., & Vargas Sánchez, J. R. (2022). Descomposición de la desigualdad del rendimiento escolar por condición de pobreza en estudiantes mexicanos. *Análisis económico*, *37*(95), 125–141. https://doi.org/10.24275/uam/azc/dcsh/ae/2022v37n95/santillan

Sanz, J. R., & López, E. (2021). Consecuencias pedagógicas entre el alumnado de enseñanza básica derivadas de la COVID-19. Una reflexión en torno a los grandes olvidados de la

pandemia. *Teoría de la Educación. Revista Interuniversitaria*, *33*(2), 149–166. https://doi.org/10.14201/teri.25471

Schatschneider, C., & Compton, D. L. (2013). *Applied Quantitative Analysis in Education and the Social Sciences* (Y. Petscher, Ed.). Routledge

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Houghton Mifflin.

Suárez, G., & Cevallos, D. (2023). Programa Nivelación y Aceleración Pedagógica (NAP): Una posible solución para el rezago escolar en estudiantes de básica superior: Pedagogical Leveling and Acceleration Program (NAP): A possible solution for academic lag in upper elementary students. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, *4*(6). https://doi.org/10.56712/latam.v4i6.1504

Suárez-Guerrero, C., Rivera-Vargas, P., & Rebour, M. (2020). *Preguntas educativas para la tecnología digital como respuesta*. *Edutec. Revista Electrónica de Tecnología Educativa*, (73), 7–22. https://doi.org/10.21556/edutec.2020.73.1733

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer. https://doi.org/10.1007/978-1-4419-8126-4

United Nations. (2015). *Transformar nuestro mundo: la Agenda 2030 para el Desarrollo Sostenible*. https://www.un.org/sustainabledevelopment/es/

World Bank. (2021). *Remote learning during COVID-19: Lessons from today, principles for tomorrow*. The World Bank. https://www.worldbank.org/en/topic/edutech/brief/how-countries-are-using-edtech-to-support-remote-learning-during-the-covid-19-pandemic