



# Learning Outcome Analysis of Instructional Grouping on Mathematical Skills for Education 4.0 in AI-Enhanced Classrooms

Darwin D. Apolinario 

International Journal of Interdisciplinary Viewpoints  
Vol. 2, No. 4, pp. 54-60, 2026

Correspondence: [d.apolinario@usls.edu.ph](mailto:d.apolinario@usls.edu.ph)

<sup>1</sup> University of St. La Salle, Bacolod City, Philippines

## Abstract

The integration of artificial intelligence (AI) in Philippine basic education under DepEd Order No. 003, s. 2026 has heightened the need for evidence-based guidance on how students should be organized in AI-enhanced classrooms. This study examined whether significant differences exist in mathematical skills—critical thinking, problem-solving, creativity, and application—between students in homogeneous and heterogeneous Grade 10 mathematics sections using AI-driven instruction. A causal-comparative design with statistical control was employed involving 94 students from a private secondary school in Negros Occidental, Philippines. Data were gathered using the researcher-developed Mathematical Competencies Assessment for Education 4.0, which underwent expert validation. Prior mathematics achievement was treated as a covariate, and the data were analyzed using multivariate and univariate analyses of covariance. Results revealed a significant overall difference in mathematical skills between the two grouping arrangements after controlling for prior achievement. Students in homogeneous sections demonstrated significantly higher levels of critical thinking, problem-solving, and application than those in heterogeneous sections. No significant difference was found in creativity, and both groups showed only modest performance in this domain. The findings suggest that homogeneous grouping may be more effective for developing analytical and applied mathematical competencies in AI-enhanced classrooms, whereas creativity requires explicit instructional strategies beyond grouping decisions alone. These results provide preliminary evidence that grouping strategy is an important pedagogical consideration in the implementation of AI in mathematics education.

## Keywords

homogeneous grouping; heterogeneous grouping; AI-enhanced classroom; mathematical skills; Education 4.0

## Citation

Apolinario, D. (2026). Learning Outcome Analysis of Instructional Grouping on Mathematical Skills for Education 4.0 in AI-Enhanced Classrooms. *International Journal of Interdisciplinary Viewpoints*, 2(4), 54–60.

<https://doi.org/10.64612/ijiv.v2i4.120>

Received: 4 March 2026

Accepted: 23 April 2026

Published: 30 April 2026

## Copyright and License

© 2026 The Author. Published by Edukar Publishing. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Peer Review Statement

This article underwent a double-anonymous peer review process in accordance with the editorial policies of the *International Journal of Interdisciplinary Viewpoints*. Both author and reviewers remained anonymous throughout the review process.

## Open Access Statement

The *International Journal of Interdisciplinary Viewpoints* is a fully open-access, peer-reviewed journal. All articles are freely available online immediately upon publication without subscription or access fees.

## Funding Statement

The author received no financial support for the research, authorship, and/or publication of this article.

## Conflict of Interest

The author declared no conflicts of interest.

## Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

## Author Contributions

The author contributed to conceptualization, methodology, investigation, writing—original draft preparation, writing—review and editing, and supervision. The author approved the final manuscript.

## Ethics Statement

This study was conducted in accordance with ethical standards.

## INTRODUCTION

Artificial intelligence is steadily reshaping the landscape of Mathematics Education, transforming classrooms into adaptive environments where instruction can respond to learners with unprecedented precision. Across the world,

education systems are reimagining how mathematical understanding is cultivated in an era often described as Education 4.0, where learners are expected not only to compute accurately but also to think critically, solve complex problems, apply concepts to authentic contexts, and create original solutions. Artificial Intelligence in Education has emerged as a central force in this transformation, offering intelligent tutoring systems, adaptive feedback, and learning analytics that personalize instruction and support higher-order thinking (Luckin & Holmes, 2016; Holmes et al., 2019; Hussin, 2018; World Economic Forum, 2020). Recent studies suggest that AI can strengthen mathematical word-problem solving, reduce conceptual gaps, and support inclusive learning for students with diverse needs (Liu et al., 2025; Naseer & Khawaja, 2025; Ondap & Simpal, 2025; Polydoros et al., 2026). Bibliometric evidence further confirms that intelligent tutoring systems are becoming increasingly influential in mathematics education worldwide (Kuncoro et al., 2026). Yet beneath this technological promise lies a profoundly human question: when machines become learning companions, how should students be grouped so that each learner has the greatest opportunity to grow?

For schools in the Philippines, this question has become more urgent with the release of DepEd Order No. 003, series of 2026 by the Department of Education, which formally recognizes AI as part of the nation's educational future. As teachers navigate this transition, they do so while balancing instructional realities, professional collaboration, and inclusive teaching practices (Cole, 2025; Lazaga, 2025; Saro et al., 2025). Local scholarship highlights both the promise and barriers of AI integration, emphasizing that successful implementation depends on teacher competence, pedagogical knowledge, and institutional support (Bayo Jr. & Doronio, 2025; Cariaga et al., 2025; Alwakid et al., 2025; Arefian, 2026). Within mathematics classrooms, educators continue to organize students into Homogeneous Grouping and Heterogeneous Grouping, each arrangement carrying distinct pedagogical assumptions. Research from Europe, Nigeria, and the Philippines has shown that grouping structures can shape achievement and participation, although results remain mixed and context dependent (Gabaldón-Estevan, 2020; Gambari et al., 2018; Černilec et al., 2023; de Guzman & Balmeo, 2023; Ghanbari & Abdolrezapour, 2020). In many classrooms, grouping decisions are still guided by tradition, intuition, or logistical convenience rather than by evidence generated under AI-enhanced conditions.

Despite the growing body of literature on AI in education and instructional grouping, an important gap remains. Existing studies have documented how AI improves planning, feedback, and student performance (Liu et al., 2025; Jaya et al., 2026), while other investigations have examined grouping effects on achievement in conventional or technology-assisted settings (Gambari et al., 2018; Černilec et al., 2023). However, few studies have directly tested whether grouping type influences the development of higher-order mathematical skills—critical thinking, problem-solving, application, and creativity—within AI-enhanced classrooms. This gap is particularly significant in the Philippine context, where national policy now encourages AI adoption but offers limited empirical guidance regarding classroom organization. Although creativity is widely recognized as a core dimension of mathematical talent (Sriraman, 2005) and international assessments such as Organisation for Economic Co-operation and Development's Programme for International Student Assessment emphasize the integrated nature of mathematical literacy (OECD, 2019), little is known about whether AI and grouping strategies work together to nurture these competencies. The present study addresses this gap by examining how homogeneous and heterogeneous grouping are associated with students' mathematical outcomes in AI-enhanced Grade 10 classrooms.

This study is significant because it brings together two transformative forces in contemporary education: the rise of artificial intelligence and the enduring influence of classroom grouping. At a practical level, the findings may guide school leaders, teachers, and policymakers in making evidence-based decisions about how students are organized as AI becomes embedded in daily instruction. At a theoretical level, the study extends the social constructivist insights of Lev Vygotsky by exploring how AI may interact with learners' Zones of Proximal Development. More importantly, the study recognizes that grouping decisions shape more than test scores; they shape the kinds of thinkers students are becoming. In an era where technology can personalize learning at scale, understanding which instructional arrangements best cultivate reasoning, application, and creativity offers a meaningful contribution to the continuing effort to ensure that innovation remains deeply connected to human learning and possibility.

## **METHODOLOGY**

### **Design**

This study employed a causal-comparative research design using a nonequivalent group structure and analysis of covariance (ANCOVA) to examine the effects of classroom grouping type on students' mathematical skill development. The independent variable was classroom grouping type, categorized as homogeneous and heterogeneous sections, which were pre-existing groups established by school policy. Because random assignment of students was not possible, the study compared naturally occurring groups while statistically controlling for differences in prior mathematical ability. This design was appropriate for investigating potential differences in critical thinking, problem-solving, application, and creativity in AI-enhanced mathematics classrooms. By utilizing intact class sections and applying covariance analysis, the study maintained ecological validity while reducing the influence of selection bias and enhancing the rigor of group comparisons.

### **Locale**

The study was conducted in a private secondary school in Negros Occidental, Philippines, that had implemented AI-driven mathematics instruction for at least one academic quarter. The school utilized adaptive learning platforms, intelligent tutoring systems, and AI-powered analytics dashboards to support instruction and monitor student progress. The institution was selected because it maintained both homogeneous and heterogeneous mathematics sections, allowing for meaningful comparison under similar instructional conditions. The focus on a private school was justified by the broader implementation of educational technologies in private institutions, which typically possess greater access to technological infrastructure and digital resources. Although the setting provided a suitable environment for examining AI-enhanced

instruction, the use of a single private school limits the generalizability of the findings to other contexts, particularly public schools with different student populations, resource levels, and instructional conditions.

### **Sampling**

A two-stage sampling strategy was employed in this study. In the first stage, purposive sampling was used to select a school that met the study criteria, including documented AI-enhanced mathematics instruction and the existence of both homogeneous and heterogeneous class groupings. In the second stage, intact group sampling was applied by including all students enrolled in one homogeneous section and one heterogeneous section. This approach preserved the natural classroom structure and strengthened ecological validity. An a priori power analysis using G\*Power for multivariate analysis of covariance (MANCOVA), assuming a medium effect size of 0.10, a significance level of 0.05, and statistical power of 0.80, indicated a minimum sample size of 92 participants. To account for possible attrition, a target sample of 100 students was established. The final sample consisted of 94 Grade 10 students, including 46 students in the homogeneous section and 48 students in the heterogeneous section, which exceeded the minimum required sample size.

### **Participants**

The participants were Grade 10 students enrolled in a private secondary school in Negros Occidental. The school was selected based on four criteria: (1) classification as a private secondary institution; (2) documented use of AI-driven mathematics instruction for at least one academic quarter; (3) availability of both homogeneous and heterogeneous mathematics sections; and (4) willingness to participate in the study. Students in the selected sections served as the respondents because they had experienced comparable curricular content and instructional strategies, differing primarily in classroom grouping type. Their participation provided the basis for comparing the influence of grouping arrangements on the development of higher-order mathematical competencies in AI-supported learning environments.

### **Instrument and Procedure**

Data were collected using the researcher-developed Mathematical Competencies Assessment for Education 4.0 (MCAE-4), which was specifically designed to measure four dimensions of mathematical skill development: critical thinking, problem-solving, application, and creativity. Critical thinking items required students to analyze claims and evaluate mathematical arguments, while problem-solving items involved multi-step strategic reasoning. Application items assessed students' ability to translate real-world situations into mathematical models, and creativity items measured flexibility in generating multiple valid solutions. A four-level analytic scoring rubric was developed to ensure consistent evaluation of responses.

To establish content validity, the instrument was reviewed by six experts in mathematics education, educational assessment, AI-enhanced learning, and research methodology. Each expert rated the relevance of every item using a four-point scale. All items obtained a Content Validity Ratio (CVR) of 1.00 and an Item-level Content Validity Index (I-CVI) of 1.00, surpassing the critical standards recommended by Lawshe (1975) and Polit and Beck (2006). These results indicated that the instrument demonstrated excellent content validity and was appropriate for measuring the targeted competencies.

### **Ethical Considerations**

The study adhered to recognized ethical principles governing research involving human participants, including respect for persons, beneficence, and justice. Written informed consent was obtained from parents or legal guardians, and student assent was secured prior to participation. The consent and assent forms clearly explained the study's purpose, procedures, confidentiality safeguards, and the voluntary nature of participation. Respondents were informed that they could withdraw from the study at any point without academic or personal consequences. To protect confidentiality, student names were replaced with unique identification codes, and the master list linking names to codes was stored separately in a password-protected file. All results were reported in aggregate form to prevent the identification of individual respondents. Physical records were secured in a locked cabinet, while electronic files were password-protected and retained only in accordance with institutional data retention policies before secure destruction.

### **Statistical Analysis**

Descriptive and inferential statistical techniques were used to analyze the data. Mean scores and standard deviations were computed to determine the level of students' mathematical skill development in the areas of critical thinking, problem-solving, application, and creativity. To interpret the results, the following descriptive scale was used: 8.10–9.00 (Very High), 7.20–8.09 (High), 6.30–7.19 (Moderately High), 5.40–6.29 (Moderate), 4.50–5.39 (Low Moderate), 3.60–4.49 (Low), and below 3.60 (Very Low). Analysis of covariance (ANCOVA) was employed to determine whether significant differences existed between homogeneous and heterogeneous groups while controlling for prior mathematical ability. The level of significance for all statistical tests was set at 0.05.

## **RESULTS AND DISCUSSION**

A clear descriptive pattern emerges when comparing the performance of homogeneous and heterogeneous sections across the four mathematical skill domains. Across critical thinking, problem-solving, and application, students in homogeneous sections consistently outperformed their heterogeneous counterparts, with the gap being most pronounced in the application domain. For creativity, however, both groups demonstrated comparable, modest performance levels. In the critical thinking domain, students in homogeneous sections achieved a mean score of 6.70 ( $SD = 0.553$ ), placing them in the "moderately high" range, indicating above-average performance approaching proficiency. In contrast, the heterogeneous group scored 5.21 ( $SD = 0.771$ ), which falls within the "moderate" range, reflecting developing skills with basic understanding. A similar disparity was observed in problem-solving, where the homogeneous group's mean of 6.89 ( $SD = 0.823$ ) also fell within the "moderately high" range, while the heterogeneous group's mean of 5.52 ( $SD = 0.618$ ) remained in the "moderate" range. The most striking difference appeared in the application domain. The homogeneous group scored

6.61 ( $SD = 0.614$ ), again in the "moderately high" range, whereas the heterogeneous group's mean of 4.29 ( $SD = 0.824$ ) fell within the "low" range, indicating beginning skills with limited understanding. This performance gap represents a substantial descriptive difference in students' capacity to translate real-world situations into mathematical models and interpret results within context. In contrast, creativity showed minimal descriptive differences between groups. The homogeneous group scored 5.26 ( $SD = 0.713$ ), while the heterogeneous group scored 4.75 ( $SD = 0.729$ ), both falling within the "low moderate" range. This indicates that students in both grouping arrangements demonstrated only emerging skills with inconsistent performance in creative mathematical thinking. When examining overall performance across all four domains, the homogeneous group achieved a combined mean of 6.37 ( $SD = 0.716$ ), placing them in the "moderately high" range and indicating above-average performance approaching proficiency across the integrated skill set. The heterogeneous group, with a combined mean of 4.94 ( $SD = 0.798$ ), fell within the "low moderate" range, reflecting emerging skills with inconsistent performance. These descriptive findings align with and extend prior research. The observed advantages for homogeneous groups in analytical and applied domains are consistent with Cernilec et al. (2023), who found that homogeneous groups demonstrated stronger performance in structured mathematical tasks among European secondary students, and with Gambari et al. (2018), who reported higher achievement scores for homogeneous groupings in Nigerian virtual laboratories. Locally, de Guzman and Balmeo (2023) similarly observed that homogeneous Philippine classrooms initially outperformed heterogeneous counterparts in academic achievement. However, these findings contrast with Ghanbari and Abdolrezapour (2020) and Gabaldón-Estevan (2020), who found that heterogeneous grouping benefited collaborative and applied learning outcomes in traditional classrooms. This contrast suggests that AI-enhanced environments may amplify the descriptive advantages of homogeneous grouping, perhaps by enabling more precisely calibrated instruction for students with similar developmental levels. For creativity, the comparable scores across groups in the "low moderate" range align with Sriraman's (2005) argument that mathematical creativity requires explicit instructional support beyond ability grouping alone, a theme that will be further examined in the inferential analyses that follow.

Table 1  
Descriptive Statistics of Mathematical Skills by Grouping Type

Skill Domain	Section Type	n	Mean	SD	Verbal Description
Critical Thinking	Homogeneous	46	6.70	0.553	Moderately High
	Heterogeneous	48	5.21	0.771	Low Moderate
Problem Solving	Homogeneous	46	6.89	0.823	Moderately High
	Heterogeneous	48	5.52	0.618	Moderate
Application	Homogeneous	46	6.61	0.614	Moderately High
	Heterogeneous	48	4.29	0.824	Low
Creativity	Homogeneous	46	5.26	0.713	Low Moderate
	Heterogeneous	48	4.75	0.729	Low Moderate
Overall Mean	Homogeneous	46	6.37	0.716	Moderately High
	Heterogeneous	48	4.94	0.798	Low Moderate

Legend: 8.10–9.00 = Very High; 7.20–8.09 = High; 6.30–7.19 = Moderately High; 5.40–6.29 = Moderate; 4.50–5.39 = Low Moderate; 3.60–4.49 = Low; Below 3.60 = Very Low.

A multivariate analysis was conducted to determine whether homogeneous and heterogeneous sections differ across the four mathematical skill domains, critical thinking, problem-solving, creativity, and application, when considered simultaneously, after accounting for prior mathematics achievement. The results reveal a robust and exceptionally strong overall effect. The multivariate relationship between section type and the combined skill profile, with prior achievement controlled, was statistically significant across all three tests ( $p < .001$ ). The consistency of this finding across Pillai's Trace, Wilks' Lambda, and Hotelling's Trace confirms its robustness, indicating that the result is not an artifact of the specific statistical test employed. The effect was exceptionally large, the Wilks' Lambda value of 0.0897 indicates that only approximately 9% of the variance in the combined dependent variables remains unexplained by grouping type after controlling for prior achievement. In practical terms, students in homogeneous sections demonstrated a consistently higher skill profile across all four domains compared to their peers in heterogeneous sections, independent of their entering ability levels.

The covariate, prior grade, also showed a statistically significant relationship with the combined dependent variables (Wilks'  $\Lambda = 0.4261$ ,  $F(4, 88) = 29.6$ ,  $p < .001$ ). This confirms the appropriateness of including prior achievement as a covariate and underscores that students' previous mathematics performance strongly predicts their current mathematical skill levels. Importantly, even with this strong covariate effect partialled out, the grouping variable remained a powerful differentiator. Taken together, these multivariate findings indicate that the way students are grouped for instruction is associated with fundamentally different patterns of mathematical competency development that cannot be explained by students' entering ability levels alone. Within AI-enhanced classrooms, homogeneous grouping appears to create learning conditions that more effectively foster the integrated development of higher-order mathematical competencies, a holistic outcome that aligns with the Education 4.0 paradigm (Hussin, 2018; World Economic Forum, 2020), which conceptualizes mathematical proficiency as an interconnected set of competencies rather than isolated skills. These results align with and extend prior research. Cernilec et al. (2023) employed MANOVA to examine grouping effects on mathematics knowledge among European secondary students, demonstrating that grouping type significantly influences the multivariate profile of mathematical competencies. The present study extends this by showing that within AI-enhanced classrooms, the multivariate effect was

exceptionally large (Pillai's Trace = 0.910), with approximately 91% of the variance in combined skills explained by section type after controlling for prior achievement. Similarly, Gambari et al. (2018) utilized MANOVA to analyze grouping effects on chemistry achievement in Nigerian virtual laboratory settings, finding significant multivariate effects for group composition and confirming that technology-enhanced learning environments produce complex, multi-dimensional outcomes best analyzed through multivariate techniques. Locally, de Guzman and Balmeo (2023) employed ANCOVA in their Philippine-based study of grouping effects on academic achievement persistence, demonstrating the importance of statistical control for prior performance. Their finding that homogeneous grouping initially outperformed heterogeneous grouping, with significant effects after controlling for covariates, aligns with the present results. The current study extends this local work by demonstrating that grouping effects extend beyond simple achievement to encompass multiple higher-order skill domains critical for Education 4.0.

Theoretically, these findings resonate with Vygotsky's (1978) social constructivist framework, wherein homogeneous grouping may create more coherent learning environments where students share similar Zones of Proximal Development, allowing for coordinated instruction that simultaneously addresses multiple skill domains. The results also align with the Programme for International Student Assessment (PISA) framework (OECD, 2019), which emphasizes that mathematical literacy encompasses the capacity to reason mathematically, solve problems, apply concepts in real-world contexts, and think creatively about mathematical situations. The significant grouping effect after controlling for prior achievement demonstrates that instructional organization independently influences this comprehensive skill profile.

Several limitations should be considered when interpreting these multivariate results. The causal-comparative design limits the ability to draw causal conclusions, as other unmeasured variables such as teacher quality, instructional time, student motivation, or classroom climate may contribute to the observed multivariate differences between groups. While the inclusion of prior grades as a covariate strengthens causal inference by controlling for pre-existing ability differences, it does not eliminate the possibility of selection bias or other unmeasured confounding variables. The significant multivariate relationship was observed within AI-enhanced classrooms in Philippine private schools, and findings may not generalize to non-AI environments, different grade levels, public school settings, or other cultural contexts. Additionally, the cross-sectional design captures current skill levels rather than developmental trajectories. De Guzman and Balmeo (2023) found that while homogeneous grouping showed initial advantages, heterogeneous grouping promoted long-term retention, suggesting that the multivariate relationship observed in this study may reflect immediate performance differences that could change over time.

Table 2  
 Multivariate Differences in Mathematical Skills by Grouping Type Controlling for Prior Achievement

Source	Multivariate Test	Value	F	df1	df2	p-value
Section Type	Pillai's Trace	0.910	223.20	4	88	<.001
	Wilks' Lambda	0.0897	223.20	4	88	<.001
	Hotelling's Trace	10.14	223.20	4	88	<.001
Prior Grades	Pillai's Trace	0.574	29.60	4	88	<.001
	Wilks' Lambda	0.4261	29.60	4	88	<.001
	Hotelling's Trace	1.35	29.60	4	88	<.001

Following the multivariate analysis, univariate tests were conducted to pinpoint where specific differences lie across the four mathematical skill domains. These results reveal a nuanced picture: instructional grouping exerts a powerful and statistically significant influence on three of the four competencies, critical thinking, problem-solving, and application, but does not reliably differentiate performance in creativity. The most substantial grouping effect emerged for application, where section type explained a remarkable 34.3% of the variance ( $\eta^2p = 0.343$ ,  $p < .001$ ). This very large effect size indicates that students in homogeneous sections scored, on average, 2.32 points higher on the MCAE-4 rubric, equivalent to moving from "low" to "moderately high", after controlling for prior achievement. This finding aligns with Gambari et al. (2018), who reported that homogeneous groupings in Nigerian virtual laboratory settings resulted in significantly higher achievement on applied tasks. The magnitude of the present effect suggests that homogeneous grouping is particularly potent for developing the real-world mathematical literacy emphasized in Education 4.0 and the PISA framework.

A large effect was also observed for problem-solving ( $\eta^2p = 0.153$ ,  $p < .001$ ), where 15.3% of the variance was uniquely attributable to section type. Similarly, a medium-to-large effect was found for critical thinking ( $\eta^2p = 0.110$ ,  $p = .001$ ), with grouping explaining 11.0% of the variance. In both domains, students in homogeneous sections demonstrated significantly stronger analytical and strategic skills, independent of their entering ability levels. These findings are consistent with Cernilec et al. (2023), who documented that homogeneous groups excelled in structured tasks requiring systematic application of learned procedures in European secondary mathematics classrooms. Locally, de Guzman and Balmeo (2023) employed ANCOVA in their Philippine-based study of grouping effects, finding that homogeneous grouping initially outperformed heterogeneous grouping with significant effects after controlling for covariates, a pattern that aligns with the present results. In contrast, the analysis revealed no significant grouping effect for creativity ( $p = .329$ ), with section type explaining only 1.0% of the variance ( $\eta^2p = 0.010$ ). This finding is noteworthy. It suggests that unlike analytical skills, creativity may not be effectively fostered through ability-aligned instruction alone. While prior studies on grouping effects have seldom included creativity measures, this result aligns with the broader literature on mathematical creativity (Sriraman, 2005; Torrance, 1974), which indicates that creative thinking develops through different mechanisms than analytical skills and requires targeted instructional interventions rather than emerging from grouping arrangements alone.

The finding implies that both grouping arrangements were equally (in)effective at cultivating this higher-order competency, highlighting creativity as a distinct instructional challenge within AI-enhanced classrooms. An examination of the covariate, prior achievement, reinforces the validity of these findings. Prior grades significantly predicted performance in critical thinking, problem-solving, and application, explaining between 6.0% and 27.1% of the variance. Crucially, for both problem-solving and application, the unique contribution of section type exceeded that of prior achievement (15.3% vs. 6.0% for problem-solving; 34.3% vs. 25.0% for application). This demonstrates that the instructional grouping strategy contributes uniquely to skill development, above and beyond what students bring to the classroom.

Several limitations should be considered when interpreting these univariate results. The causal-comparative design limits causal inference, as unmeasured variables such as teacher quality, instructional time, student motivation, or classroom climate may have contributed to the observed differences. While prior grades were included as a covariate to control for pre-existing ability, this does not eliminate the possibility of selection bias or other confounding factors. The non-significant finding for creativity may reflect instrument limitations, as the MCAE-4 instrument, while rigorously validated for analytical skills, may not fully capture the multidimensional nature of creative mathematical thinking (e.g., fluency, flexibility, originality). Alternatively, creativity may develop over longer timeframes than the single-quarter implementation captured in this cross-sectional study. Additionally, the cross-sectional design captures current skill levels rather than developmental trajectories; as de Guzman and Balmeo (2023) found, while homogeneous grouping showed initial advantages, heterogeneous grouping promoted long-term retention, suggesting that the advantages observed in this study may reflect immediate performance differences that could change over time. Finally, the residuals for each domain indicated substantial variance not explained by either grouping type or prior achievement, suggesting that other factors beyond those measured contribute to mathematical skill development.

Table 3  
Tests of Between-Subjects Effects for Individual Mathematical Skill Domains Controlling for Prior Achievement

Source	Dependent Variable	Sum of Squares	df	Mean Square	F	p-value	$\eta^2$	Partial $\eta^2$
Section Type	Critical Thinking	3.763	1	3.763	11.30	.001	0.083	0.110
	Problem Solving	8.222	1	8.222	16.43	< .001	0.145	0.153
	Creativity	0.494	1	0.494	0.964	.329	0.010	0.010
	Application	19.084	1	19.084	47.40	< .001	0.281	0.343
Prior Grade	Critical Thinking	11.288	1	11.288	33.80	< .001	0.249	0.271
	Problem Solving	2.904	1	2.904	5.80	.018	0.051	0.060
	Creativity	1.225	1	1.225	2.390	.126	0.025	0.026
	Application	12.242	1	12.242	30.40	< .001	0.180	0.250

### Conclusion and Recommendations

This study reveals that in AI-enhanced math classrooms, grouping strategy is a fundamental pedagogical condition, not an administrative convenience. Homogeneous grouping creates alignment among student readiness, AI calibration, and instructional coherence, significantly accelerating analytical, problem-solving, and application skills. Heterogeneous grouping was associated with lower outcomes in these domains, as AI systems struggle to balance scaffolding for struggling learners and enrichment for advanced ones. Notably, creativity showed no improvement under either arrangement, suggesting it is a distinct cognitive capacity requiring separate, intentional cultivation, AI does not automatically foster it.

For practice, homogeneous grouping is recommended for developing analytical skills under DepEd Order No. 003, s. 2026, with two qualifications: creativity needs explicit support, and ongoing evaluation is necessary due to the study's causal-comparative design. Administrators should treat grouping as a strategic choice, allocate creativity-focused resources, and establish evaluation systems. Teachers should use AI analytics to calibrate collective ZPD in homogeneous sections or manage differentiated pathways in heterogeneous sections, while explicitly teaching creativity. Policymakers should incorporate flexible grouping guidelines into DepEd Order No. 003 and fund research on grouping effectiveness across Philippine contexts. Future researchers should investigate mechanisms underlying grouping effects, conduct longitudinal studies, and examine creativity with multiple measures. Ultimately, grouping decisions shape not only what students learn but what kind of thinkers they become, demonstrating that pedagogy and technology are deeply intertwined forces.

### References

- Alwakid, W. N., Dahri, N. A., Humayun, M., & Alwakid, G. N. (2025). Exploring the role of AI and teacher competencies on instructional planning and student performance in an outcome-based education system. *Systems*, 13(7), 517. <https://doi.org/10.3390/systems13070517>
- Arefian, M. H. (2026). AI-enhanced professional learning communities: a new era of personalized teacher education. *Cogent Education*, 13(1), 2626629. <https://doi.org/10.1080/2331186X.2026.2626629>
- Bayo Jr, R., & Doronio, R. (2025). Technological, Pedagogical, and Content Knowledge as a Predictor of Mathematics Pre-Service Teachers' Teaching Competency: A Multiple Regression Analysis. *International Journal of Interdisciplinary Viewpoints*, 1(3), 165–196. <https://doi.org/10.64612/ijiv.v1i3.17>
- Cariaga, R., Dagunan, M. A., Cariaga, V., Sabidalas, M. A., El Halaissi, M., & Bacatan, J. (2025). Rethinking Parental Involvement in Developing Countries: Toward Inclusive and Culturally Responsive Education. *International Journal of Interdisciplinary Viewpoints*, 1(5), 631–637. <https://doi.org/10.64612/ijiv.v1i5.39>
- Cariaga, R., El Halaissi, M., Refugio, C., Dagunan, M. A., Sabidalas, M. A., Cariaga, V., ... Gerodias, E. (2025). Local Voices, Global Technologies: AI Integration Barriers in K–12 Classrooms. *International Journal of Interdisciplinary Viewpoints*, 1(5), 672–680. <https://doi.org/10.64612/ijiv.v1i5.45>
- Černilec, B., Cotič, M., Felda, D., & Doz, D. (2023). Differences in students' mathematics knowledge in homogeneous and heterogeneous groups. *European Journal of Science and Mathematics Education*, 11(1), 15–32. <https://doi.org/10.30935/scimath/12431>

- Cole, V. J. (2025). Perceptions on Inclusive Education, Professional Learning Community Practices and Teachers' Collaboration: Multiple Regression Analysis. *International Journal of Interdisciplinary Viewpoints*, 1(1), 26–37. <https://doi.org/10.64612/ijiv.v1i1.8>
- de Guzman, F. I., & Balmeo, M. G. (2023). Evaluating the Impact of Homogeneous Classrooms on the Persistence of Heterogeneous Academic Achievers: Implications for Developing Effective Support Systems. *Online Submission*, 3(1), 60–65. <https://eric.ed.gov/?id=ED631615>
- Denise F. Polit, & Cheryl Tatano Beck (2006). The content validity index: Are you sure you know what's being reported? *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Department of Education. (2026). Foundational Guidelines on Artificial Intelligence (AI) in Basic Education (DepEd Order No. 003, s. 2026). Republic of the Philippines.
- Gabaldón-Estevan, D. (2020). Heterogeneity versus homogeneity in schools: a study of the educational value of classroom interaction. *Education Sciences*, 10(11), 335. <https://doi.org/10.3390/educsci10110335>
- Galaura, R. J., & Simal, E. A. (2025). Challenges in the Implementation of K to 12 Program and Their Influence on the Instructional Competence of Teachers. *International Journal of Interdisciplinary Viewpoints*, 1(2), 121–132. <https://doi.org/10.64612/ijiv.v1i2.13>
- Gambari, A. I., Kawu, H., & Falode, O. C. (2018). Impact of virtual laboratory on the achievement of secondary school chemistry students in homogeneous and heterogeneous collaborative environments. *Contemporary Educational Technology*, 9(3), 246–263. <https://doi.org/10.30935/cet.444108>
- Ghanbari, N., & Abdolrezaipoor, P. (2020). Group Composition and Learner Ability in Cooperative Learning: A Mixed-Methods Study. *Tesl-Ej*, 24(2), n2. <https://eric.ed.gov/?id=EJ1268684>
- Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education promises and implications for teaching and learning. Center for Curriculum Redesign.
- Hussin, A. A. (2018). Education 4.0 made simple: Ideas for teaching. *International Journal of Education and Literacy Studies*, 6(3), 92–98. <https://doi.org/10.7575/aiac.ijels.v6n.3p.92>
- Jaya, S., Zaharudin, R., Bhartiya, S., & Alomari, M. (2026). Bridging Educational Gaps in Low-Resource Classrooms: AI-Enhanced Project-Based Learning for STEM Educators in Africa. *ECNU Review of Education*, 9(2), 20965311261446194. <https://doi.org/10.1177/20965311261446194>
- Kuncoro, K. S., Fitriyani, P., Erlangga, S. Y., Ario, M., & Kartika, D. L. (2026). Artificial Intelligence and Intelligent Tutoring Systems in Mathematics Education: A Bibliometric Analysis (2001–2025). *F1000Research*, 15, 107. <https://doi.org/10.12688/f1000research.176098.1>
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel psychology*, 28(4). <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Lazaga, R. J. (2025). Navigating the Engagement, Representation, and Expression in an Inclusive Learning Environment through the Eyes of Public School Teachers. *International Journal of Interdisciplinary Viewpoints*, 1(1), 9–25. <https://doi.org/10.64612/ijiv.v1i1.7>
- Liu, J., Keane, T., Sun, D., Yang, Y., & Yang, Y. (2025). Design and evaluation of ChatGPT-MWPS: an AI-enhanced learning system for improving primary students' mathematical word problem solving. *Smart Learning Environments*, 12(1), 68. <https://doi.org/10.1186/s40561-025-00419-9>
- Luckin, R., & Holmes, W. (2016). Intelligence unleashed: An argument for AI in education.
- Naseer, F., & Khawaja, S. (2025). Mitigating conceptual learning gaps in mixed-ability classrooms: A learning analytics-based evaluation of AI-driven adaptive feedback for struggling learners. *Applied Sciences*, 15(8), 4473. <https://doi.org/10.3390/app15084473>
- Ondap, A. J., & Simal, E. A. (2025). Mathematical Learning Intervention Program for Grade 9 Learners During the Resumption of In-Person Classes. *International Journal of Educational Viewpoints*, 1(1), 18–34.
- Organisation for Economic Co-operation and Development. (2019). PISA 2018 assessment and analytical framework. OECD Publishing
- Polydoros, G., Antoniou, A. S., & Polydoros, C. (2026). Inclusive AI-mediated mathematics education for students with learning difficulties: Reducing math anxiety in digital and smart-city learning ecosystems. *Encyclopedia*, 6(2), 39. <https://doi.org/10.3390/encyclopedia6020039>
- Saro, J., Pelesco, E., Abiao, J., Longaquit, J., Apique, J., & Palanog, J. G. (2025). Extent of Teachers' Manifestation Toward Instructional Supervision and Assistance from School Leaders: Its Relationship to Students' Academic Success. *International Journal of Interdisciplinary Viewpoints*, 1(5), 604–630. <https://doi.org/10.64612/ijiv.v1i5.38>
- Sriraman, B. (2005). Are giftedness and creativity synonyms in mathematics? *Journal of Secondary Gifted Education*, 17(1), 20–36. <https://doi.org/10.4219/jsge-2005-389>
- Torrance, E. P. (1974). Torrance tests of creative thinking. *Educational and psychological measurement*. <https://doi.org/10.1037/t05532-000>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard university press.
- World Economic Forum. (2020). *Schools of the Future: Defining New Models of Education for the Fourth Industrial Revolution*. World Economic Forum. [https://www3.weforum.org/docs/WEF\\_Schools\\_of\\_the\\_Future\\_Report\\_2019.pdf](https://www3.weforum.org/docs/WEF_Schools_of_the_Future_Report_2019.pdf)