

Comparing Mathematics Performance Outcomes Between Learners Exposed and Not Exposed to a Blended Digital Learning Intervention



Research Article

Authors: Nyasha Patrick Tarakino¹, Alex Mugala², Africa Research University

Corresponding Author: Nyasha Patrick Tarakino, Africa Research University email: ntarakino@gmail.com,

Citation: Tarakino, N.P, and Mugala, A., (2026). Comparing Mathematics Performance Outcomes Between Learners Exposed and Not Exposed to a Blended Digital Learning Intervention . AJIESS 2 (1):1-8.

Received: 04/05/ 2026. **Published:** 09/08/ 2026

Abstract

Introduction: This study investigates differences in mathematics performance between learners exposed to a blended DeltaMath–Quizizz intervention and those who were not.

Methods: Using a quasi-experimental design with 70 learners equally divided between exposed and non-exposed groups, four continuous outcome measures were analysed: Check-in 1, Check-in 2, mean check-in score, and end-of-course (EOC) score. Descriptive statistics, Welch independent-samples *t*-tests, one-way ANOVA, and Cohen's *d* effect sizes were computed to compare the groups. Correlations between interim assessments and EOC scores contextualised the findings.

Results: The results indicated no statistically significant differences in Check-in 1, Check-in 2, or mean check-in scores. The EOC score showed a directionally favourable difference for the exposed group (mean = 81.17, SD = 6.71) compared with the non-exposed group (mean = 76.43, SD = 15.03), approaching but not reaching conventional significance ($t = 1.705$, $p = 0.095$; $F = 2.906$, $p = 0.093$), with a small-to-moderate effect size ($d = 0.408$). Significant positive correlations were observed between interim check-ins and EOC scores.

Conclusion: Future research should examine the long-term impact of blended digital interventions on mathematics achievement and learner performance. Longitudinal, qualitative,

and comparative studies across different contexts could provide deeper insights into effectiveness, implementation, and scalability.

Keywords: mathematics achievement, blended learning, DeltaMath, Quizizz, formative assessment, effect size, rural education.

Introduction

Educational research increasingly recognises the complexity of evaluating interventions in authentic classroom settings, where sample sizes are often modest and random assignment is impractical (Shadish, Cook, and Campbell 2002). In mathematics education, the impact of instructional innovations is frequently assessed through learner performance outcomes. However, single summary indicators and binary classifications may obscure important variations in learner achievement and learning progress. Continuous outcome measures provide a more nuanced understanding of learning progress and intervention effects than single summary indicators, allowing researchers to identify patterns of achievement that may otherwise remain hidden (Messick, 1995; Cumming, 2014; Field, 2018).

This study examined the differences in mathematics performance between

learners exposed to and not exposed to a blended DeltaMath–Quizizz intervention. DeltaMath is a digital platform that offers structured mathematics practice and problem-solving tasks, whereas Quizizz provides gamified review and formative assessment opportunities. The blended use of these platforms aims to combine deliberate practice with engaging, low-stakes assessment, aligning with formative assessment and feedback theories that emphasise timely, actionable information to support learning (Black and Wiliam, 1998; Hattie and Timperley, 2007).

The rationale for the intervention is grounded in the premise that repeated practice, immediate feedback, and motivational engagement foster mathematical understanding and skill acquisition (Bloom, 1976; Guskey, 2007). By integrating DeltaMath and Quizizz into classroom routines, teachers can potentially enhance formative

assessment processes, enabling targeted instructional adjustments and supporting mastery of learning (Black and Wiliam, 1998; Guskey, 2007).

The present study compared mathematics performance outcomes between learners exposed and not exposed to the blended intervention using inferential statistics. The key research question is as follows: How do mathematics performance outcomes differ between learners exposed to the blended DeltaMath–Quizizz intervention and those not exposed, and what do inferential statistics reveal about the magnitude and certainty of these differences?

Theoretical and Empirical Literature Review

Formative Assessment and Feedback in Mathematics Learning

Formative assessment is widely acknowledged as a pivotal factor in improving student learning outcomes by providing ongoing information about learner understanding that informs instructional decisions (Black and Wiliam, 1998). In mathematics education, formative assessment facilitates the identification of

misconceptions and guides corrective feedback, which is critical given the cumulative and hierarchical nature of mathematical knowledge (Hattie and Timperley, 2007).

Hattie and Timperley (2007) emphasised that feedback is most effective when it clarifies the gap between current and desired performance and offers guidance on how to close that gap. This is particularly relevant in mathematics, where procedural fluency and conceptual understanding must be developed concurrently. Digital platforms such as DeltaMath and Quizizz can support these processes by delivering immediate feedback and enabling repeated practice, thereby reinforcing learning and promoting self-regulation (Nicol and Macfarlane-Dick, 2006; Zimmerman, 2002).

Mastery Learning and Digital Interventions

Bloom's (1976) mastery learning model posits that, given sufficient time and appropriate instructional support, most learners can achieve high levels of mastery. Guskey (2007) extended this framework by highlighting the importance of formative assessment, followed by

targeted corrective instruction to close achievement gaps. Digital tools can facilitate mastery learning by providing adaptive practice opportunities and timely feedback, enabling learners to progress at their own pace (Faber, Luyten, and Visscher, 2017).

Empirical studies have demonstrated the potential of digital formative assessment tools to enhance mathematics achievement and motivation among students. Faber et al. (2017) found that a digital formative assessment intervention led to significant improvements in student achievement and engagement. Similarly, Roschelle et al. (2016) reported that online mathematics homework increased students' achievement, underscoring the value of technology-supported practice.

Technology-Enhanced Learning and Gamification

Meta-analyses by Li and Ma (2010) and Cheung and Slavin (2013) provide evidence that computer technology can positively impact mathematics learning, although effects vary depending on the nature of the intervention and implementation context. Gamification, as implemented in Quizizz, has been shown to enhance cognitive, motivational, and

behavioural outcomes, provided that game elements are pedagogically coherent and reinforce learning objectives (Sailer and Homner, 2020).

The integration of gamified formative assessment with structured practice platforms such as DeltaMath may thus offer complementary benefits: increased engagement through gamification and improved procedural fluency through deliberate practice (Mayer, 2009; Higgins, Xiao and Katsipataki, 2012).

Statistical Considerations in Educational Research

Interpreting the intervention effects requires attention to both statistical and practical significance. Cohen (1988) advocated for the use of effect sizes to assess the magnitude of effects beyond p-values. Cumming (2014) further argued that overreliance on null hypothesis significance testing can obscure meaningful educational effects, particularly in studies with limited power or complex designs. Effect sizes, confidence intervals, and correlation patterns provide a richer interpretive framework for understanding intervention outcomes (Field, 2018).

Methodology

Research Design and Participants

This study employed a quasi-experimental design to compare the mathematics performance of learners exposed to and not exposed to a blended DeltaMath–Quizizz intervention. The sample comprised 70 learners from a rural schooling context, with 35 learners in the exposed group and 35 in the non-exposed group. Group assignment was based on participation in the intervention rather than random allocation, reflecting the practical constraints of the educational setting.

Intervention Description

The blended intervention combined structured problem-solving exercises with gamified formative assessments. DeltaMath provided learners with opportunities for deliberate practice and immediate feedback on mathematical

procedures, whereas Quizizz offered engaging review activities designed to reinforce content and assess understanding in a low-stakes environment. The integration aimed to support formative assessment cycles and mastery of learning principles.

Measures

Four continuous outcome variables were analysed.

- Check-in 1 score: An interim assessment early in the course.
- Check-in 2 score: Subsequent interim assessment.
- Mean check-in score: Average of Check-in 1 and 2 scores.
- End-of-course (EOC) score: The final summative assessment score.

These measures allowed for a comparison of short-cycle interim performance and cumulative course-level achievement.

Data Analysis

Descriptive statistics (means and standard deviations) were used to characterise group performance. Welch independent-samples t-tests were used to compare group means, chosen for

robustness against unequal variances (Field, 2018). One-way ANOVA was used for confirmatory group comparisons. Cohen's *d* quantified the standardised effect sizes, which were interpreted using conventional benchmarks (Cohen, 1988). Pearson's

correlation coefficients were used to examine the relationships between interim assessments, platform-specific performance indicators, and EOC scores to contextualise the findings.

All analyses were conducted with a significance threshold of $p < 0.05$, and effect sizes and confidence intervals were used for practical interpretation.

Sample characteristic	N
Total learners	70
Exposed learners	35
Non-exposed learners	35
Outcome variables	Check-in 1, Check-in 2, mean check-in, EOC

Results

This section reports the performance evidence in a structured way. It first compares the mean scores for the exposed and non-exposed groups across the interim check-ins, the mean check-in score, and the end-of-course assessment. Inferential tests and effect

sizes are then presented to indicate whether the observed differences are statistically reliable and educationally meaningful. Finally, it reports correlations showing how interim performance and platform indicators relate to end-of-course achievements.

Descriptive and Inferential Statistics

Table 1 summarises the descriptive and inferential statistics for the four outcome measures used in this study.

Outcome	Exposed Mean (SD)	Non-exposed Mean (SD)	T	p (t-test)	F	p (ANOVA)	Cohen's d

Check-in 1	35.14 (14.86)	36.60 (18.41)	-0.364	0.717	0.133	0.717	-0.087
Check-in 2	40.29 (15.40)	38.77 (17.96)	0.379	0.706	0.143	0.706	0.091
Mean check-in	37.71 (12.86)	37.69 (13.41)	0.009	0.993	0.000	0.993	0.002
EOC score	81.17 (6.71)	76.43 (15.03)	1.705	0.095	2.906	0.093	0.408

The interim Check-in 1 and Check-in 2 scores showed negligible differences between the groups, with trivial effect sizes and non-significant p-values. The mean check-in score was virtually identical across the groups. The EOC score exhibited a directionally favourable difference for the exposed group, with a

mean score approximately 4.74 points higher than that of the non-exposed group. Although this difference did not reach conventional statistical significance, the effect size ($d = 0.408$) suggests a small to moderate practical effect.

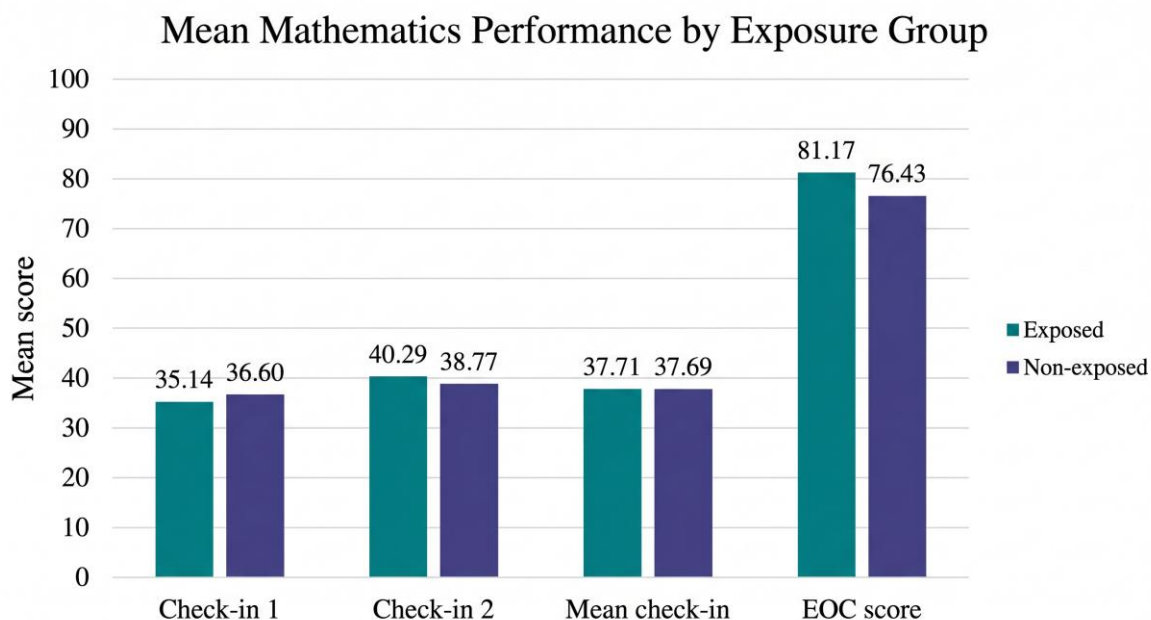


Figure 1. Mean mathematics performance by exposure group.

Figure 2 shows the relationship between mean check-in scores and EOC mathematics scores. The scatterplot

indicates a positive association between formative assessment performance and final achievement.



Figure 2: Relationship between the mean check-in score and EOC.

Correlation Analysis

Table 2 presents Pearson's correlation coefficients between the interim assessments, platform-specific performance indicators, and EOC scores.

Variable 1	Variable 2	R	P	N
Check-in 1	EOC	0.321	0.007	70
Check-in 2	EOC	0.466	<0.001	70
Mean check-in	EOC	0.502	<0.001	70
DeltaMath	EOC	0.200	0.097	70
Quizizz	EOC	0.198	0.100	70

Significant positive correlations were observed between all interim check-in scores and EOC performance, with the strongest association being the mean check-in ($r = 0.502$, $p < 0.001$). DeltaMath and Quizizz performance indicators showed positive but non-significant correlations with the EOC scores.

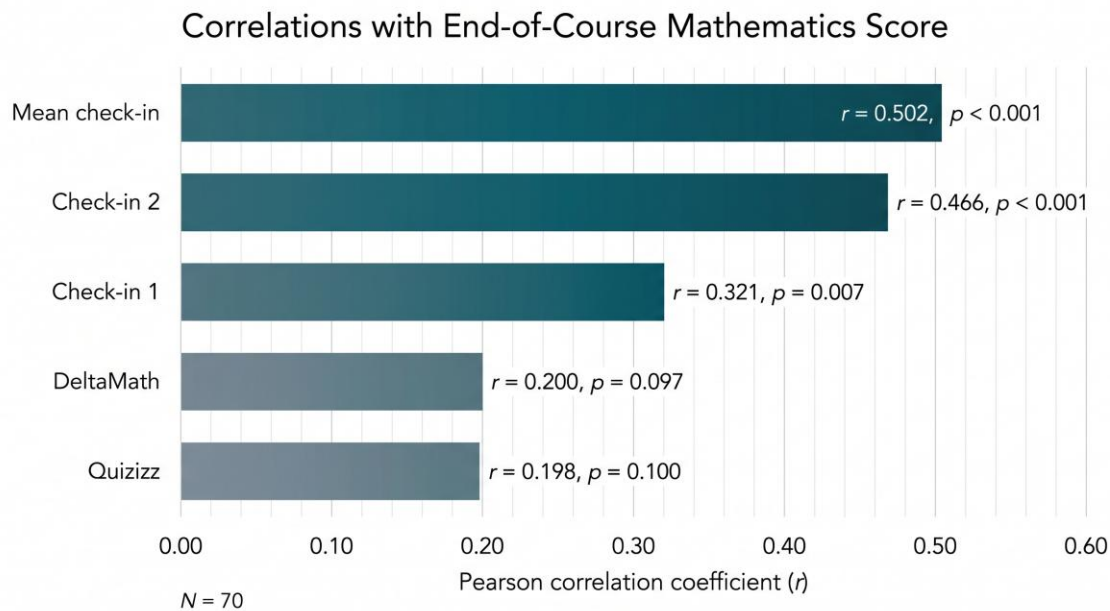


Figure 3. Correlations between performance measures and EOC scores.

Visualisation of EOC Scores

Error bar plots illustrate the mean EOC scores and variability by group, highlighting the directionally higher performance of exposed learners.

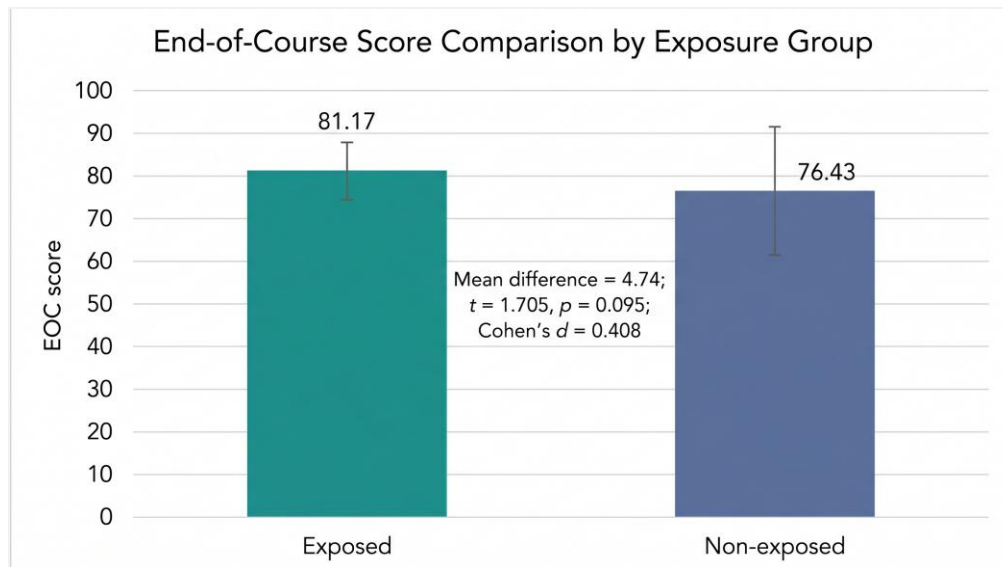


Figure 4: End-of-course score comparison by exposure group.

Figure 5 presents the distribution of end-of-course (EOC) mathematics scores for learners exposed and not exposed to the blended DeltaMath–Quizizz intervention.

The boxplot provides a visual comparison of score distributions, central tendency, variability, and potential outliers across the two groups.

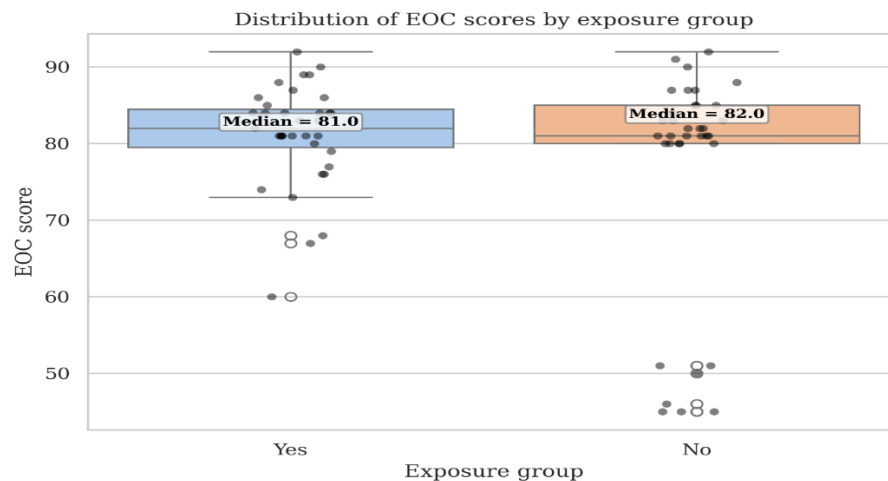


Figure 5: Distribution of EOC scores by exposure group.

Discussion

The discussion explains what statistical results mean for blended mathematics learning, formative assessment, and instructional decision-making. It distinguishes between statistical significance, practical significance, and classroom usefulness, and clarifies why the end-of-course pattern should be viewed as promising but not conclusive.

Interpretation of Findings

The study examined whether exposure to the blended DeltaMath and Quizizz

intervention was associated with differences in continuous mathematics performance between exposed and non-exposed learners. This analysis was particularly significant because continuous score-based measures often reveal subtle patterns of learning growth that are not always evident when achievement is summarised using a single performance indicator. In this regard, four key outcome variables were examined, namely Check-in 1, Check-in 2, mean check-in scores, and End-of-Course (EOC) examination scores. The analysis adopted a multi-method

statistical approach that included descriptive statistics, independent samples t-tests, one-way ANOVA, and effect size estimation using Cohen's d. This triangulated approach ensured that the interpretation of results was not limited to statistical significance alone but also incorporated the magnitude, direction, and practical relevance of observed differences, in line with established recommendations in educational measurement and evaluation (Field, 2018; Cohen, 1988).

The descriptive results for the interim assessments indicated a high degree of similarity between the exposed and non-exposed groups across all short-cycle measures, suggesting that the intervention did not produce an immediate or clearly observable performance advantage in early or mid-course assessments. For Check-in 1, the non-exposed group recorded a slightly higher mean score than the exposed group, although this difference was minimal and statistically non-significant, indicating that initial exposure to the blended intervention did not translate into an early measurable performance improvement. For Check-in 2, the direction of difference shifted slightly in

favour of the exposed group; however, this advantage remained marginal and equally non-significant, reinforcing the absence of a consistent treatment effect across short-cycle assessments. When examining the mean check-in scores, the results revealed near-perfect equivalence between the two groups, further demonstrating that exposure to the intervention did not meaningfully differentiate learner performance across formative assessment points. Collectively, these findings suggest that the intervention did not produce a stable or detectable short-term academic advantage within the instructional period under review.

This pattern of near-equivalence across interim measures is methodologically important because it highlights the stability of learner performance across both groups under short-cycle assessment conditions. It suggests that daily or weekly fluctuations in learner performance were not meaningfully influenced by exposure to the blended DeltaMath and Quizizz intervention, thereby indicating that the intervention did not disrupt or significantly enhance short-term learning trajectories in a measurable way. However, this finding

should not be interpreted simplistically as evidence of pedagogical ineffectiveness. Rather, it reflects the well-established limitation of interim assessments in capturing transient and context-sensitive performance variations, which are often influenced by attendance, instructional pacing, classroom dynamics, and learner motivation. This finding is in tandem with Black and Wiliam (1998), who argue that formative assessment data often reflect momentary performance states rather than stable learning gains, thereby limiting their sensitivity in detecting sustained instructional effects over time.

The analysis of End-of-Course (EOC) performance, however, revealed a more pronounced and educationally meaningful pattern of difference between the two groups. The exposed learners achieved a higher mean EOC score (81.17) compared to the non-exposed learners (76.43), representing a mean difference of approximately 4.74 points in favour of the intervention group. Although this difference did not reach conventional levels of statistical significance under the independent samples t-test, the consistent direction of the effect across statistical procedures suggests a potentially meaningful trend in favour of

the exposed group. Additionally, the one-way ANOVA confirmed a similar non-significant pattern, while the calculated effect size (Cohen's $d = 0.408$) indicated a small-to-moderate practical effect. This combination of results suggests that while the observed difference cannot be described as statistically conclusive, it nevertheless reflects a non-trivial and potentially meaningful improvement in cumulative mathematics performance associated with exposure to the intervention.

This finding is in tandem with Hattie (2009), whose meta-analytic work on educational influences emphasizes that many instructional interventions produce effects that are more accurately interpreted through effect sizes rather than strict reliance on p-values. Similarly, Coe (2002) argues that educational significance often extends beyond statistical thresholds, particularly when effect sizes indicate consistent directional advantages that may have practical implications in classroom settings. In the context of this study, the observed EOC difference suggests that exposure to the blended DeltaMath and Quizizz intervention was associated with improved cumulative mathematics

performance, even if the sample size and variability limited the ability to establish statistical significance at conventional levels.

The distinction between statistical significance and educational meaning becomes particularly important in interpreting these findings, as reliance solely on p-values may obscure practically relevant patterns in learner performance. Although the EOC difference did not meet the threshold for statistical significance, the consistent directional advantage and moderate effect size suggest that the intervention may have had meaningful implications for student achievement in a real-world instructional context. In educational settings, even modest improvements in end-of-course performance can be highly consequential, particularly in high-stakes assessment environments where small score increases may influence course completion, academic progression, or placement decisions. This interpretation aligns with Cumming (2014), who argues that effect sizes and confidence intervals often provide more informative insights for educational decision-making than binary significance testing alone.

When the EOC findings are interpreted alongside the interim assessment results, a more complex and developmentally plausible pattern emerges. While Check-in 1, Check-in 2, and mean check-in scores showed no meaningful differences between exposed and non-exposed learners, the EOC outcome demonstrated a more favourable performance trend for the exposed group. This divergence suggests that the effects of the blended DeltaMath and Quizizz intervention may not have manifested immediately but instead accumulated gradually over time, becoming more visible in cumulative end-of-course performance. Such a pattern is consistent with the idea that adaptive learning systems and gamified platforms often support incremental knowledge acquisition, where learning gains develop progressively rather than instantaneously. This interpretation aligns with Messick's (1995) argument that cumulative assessments are often more valid indicators of sustained learning development than isolated short-term measures.

Further interpretive depth is provided by the correlation between interim assessments and EOC performance,

which demonstrated a coherent and internally consistent pattern of academic achievement across the assessment continuum. Specifically, Check-in 1 showed a positive correlation with EOC performance ($r = 0.321$, $p = 0.007$), while Check-in 2 exhibited a stronger relationship ($r = 0.466$, $p < 0.001$), and the mean check-in score demonstrated the strongest association ($r = 0.502$, $p < 0.001$). These findings indicate that interim performance measures were not isolated or random but were meaningfully aligned with final achievement outcomes, suggesting that learners who performed well during the instructional process were also more likely to achieve higher EOC scores. This strengthens the interpretive validity of the dataset and provides evidence that the observed EOC differences are embedded within a consistent performance structure rather than representing statistical anomalies.

The findings indicate that exposure to the blended intervention was not associated with statistically significant differences in short-cycle mathematics assessments and did not achieve conventional statistical significance in end-of-course comparisons. However, the presence of a consistent directional advantage in

EOC scores, combined with a small-to-moderate effect size, suggests that the intervention may have had a modest but meaningful influence on cumulative mathematics achievement. The most defensible interpretation is therefore that the intervention was associated with a weak-to-moderate but directionally positive pattern of performance, most clearly expressed in end-of-course outcomes rather than in interim assessments, underscoring the importance of using multiple complementary measures when evaluating the effectiveness of educational interventions in blended learning environments.

Correlations and Internal Consistency

The significant positive correlations between interim check-in scores and EOC performance support the internal validity of assessment measures. These relationships indicate that interim assessments captured meaningful variations in learner achievement and were predictive of final course outcomes. The weaker and non-significant correlations for DeltaMath and Quizizz platform-specific indicators suggest that isolated platform usage metrics may be insufficient proxies for overall

mathematics attainment. This finding aligns with the view that digital tools must be integrated into coherent pedagogical routines to realise their potential (Black and Wiliam, 1998; Faber, Luyten and Visscher, 2017).

Comparison with Prior Research

This finding is supported by Field (2018), who notes that strong inter-correlations among educational measures enhance confidence in the validity of performance interpretations. It is also in tandem with Messick's (1995) unified validity framework, which emphasizes that consistency across multiple measurement points strengthens the argument that observed outcomes reflect genuine differences in learning rather than measurement error or chance variation. In this case, the consistent relationship between interim assessments and EOC performance suggests that the dataset captures a stable academic performance trajectory, thereby reinforcing the credibility of the observed trends.

Despite the absence of statistically significant differences in most comparisons, the overall pattern of findings suggests that the blended

DeltaMath and Quizizz intervention may have contributed to modest improvements in cumulative mathematics performance. The lack of significant effects in interim assessments indicates that the intervention did not produce immediate or short-term performance gains, while the more favourable EOC outcomes suggest that its impact may have accumulated gradually over time. This pattern is consistent with contemporary understandings of digital and adaptive learning environments, which often support long-term learning development through repeated practice, feedback loops, and learner engagement rather than immediate test score improvements.

Implications of the Study Findings

The findings of this study have important implications for mathematics instruction, assessment practice, and the integration of digital learning tools in secondary education. Although statistically significant differences were not observed across most performance measures, the directionally favourable end-of-course results for learners exposed to the blended DeltaMath–Quizizz intervention

suggest that such tools may support cumulative mathematics learning when implemented consistently over time. These findings indicate that digital learning platforms may contribute to gradual improvements in learner achievement through repeated practice, immediate feedback, and increased opportunities for engagement. Consequently, educators should view blended digital interventions as complementary instructional supports that enhance existing teaching practices rather than as stand-alone solutions for improving mathematics performance.

From an instructional practice perspective, the findings suggest that teachers implementing blended learning tools such as DeltaMath and Quizizz should not expect immediate or uniform improvements in mathematics performance. The absence of statistically significant differences in short-cycle assessments indicates that the effects of the intervention may emerge gradually rather than immediately. However, the directionally favourable end-of-course results and small-to-moderate effect size suggest that sustained engagement with the platforms may support cumulative learning over time. Teachers should

therefore prioritise consistent and structured integration of these tools within daily instructional routines.

In terms of assessment practice, the findings highlight the importance of using multiple measures when evaluating instructional interventions. While no statistically significant differences were observed across most short-cycle assessments, the directionally favourable end-of-course results suggest that cumulative performance measures may provide valuable insights into learner achievement. Consequently, schools should consider combining formative and summative assessment evidence when evaluating the effectiveness of instructional innovations.

From a curriculum implementation perspective, the findings imply that blended learning tools are most effective when used to support sustained practice and conceptual reinforcement rather than as short-term interventions aimed at rapid performance gains. The more favourable continuous achievement outcomes observed among exposed learners suggest that these platforms may be particularly useful for reinforcing procedural fluency, increasing practice

frequency, and supporting incremental mastery of mathematical concepts. This aligns with the idea that technology enhanced learning is most effective when it is embedded into daily instructional routines rather than treated as an occasional enrichment activity.

In terms of pedagogical implications, the results suggest that learner engagement plays a critical role in determining the effectiveness of the intervention. The directionally favourable end-of-course performance observed among exposed learners suggests that sustained participation in blended digital learning activities may support cumulative mathematics achievement over time. This implies that teachers should focus not only on providing access to digital tools but also on actively monitoring engagement, encouraging consistent use, and integrating platform activities with classroom instruction. Without sustained engagement, the potential benefits of adaptive feedback and gamified practice may not fully translate into measurable achievement gains.

The study also has implications for how educational success is evaluated in school systems. The findings highlight

the value of continuous performance measures in capturing patterns of learner achievement that may not be fully reflected through summary performance indicators alone. Learners may demonstrate meaningful academic growth over time, even when improvements are modest or gradual. This underscores the importance of using multiple measures of achievement when evaluating instructional interventions and supports the adoption of more comprehensive assessment frameworks that recognise both performance outcomes and learner progress.

Finally, the study contributes to broader theoretical and practical understandings of educational technology integration in mathematics education. It suggests that the effectiveness of interventions such as DeltaMath and Quizizz is best understood as part of a complex learning ecosystem rather than as isolated causal drivers of achievement. Their impact appears to be incremental and more evident in cumulative performance outcomes than in short-cycle assessment measures. This reinforces the importance of multi method evaluation approaches when assessing digital learning tools and supports the

view that educational technology should be positioned as a complement to, rather than a replacement for, high quality instructional practice.

Limitations and Directions for Further Inquiry

Several limitations must be acknowledged in this study, particularly in relation to research design, sample size, and the level of detail available in the dataset. The use of a quasi-experimental design rather than a fully randomized controlled design limited the ability to draw strong causal conclusions, as intact groups may have differed in unmeasured ways prior to the intervention. In addition, the analysis did not demonstrate statistically significant differences across most performance measures, limiting the strength of conclusions that can be drawn regarding intervention effectiveness, meaning that the modest end-of-course advantage observed among exposed learners cannot be interpreted as definitive evidence that the intervention directly caused improved attainment. Furthermore, the dataset did not capture detailed implementation data such as frequency of platform use, duration of engagement, or specific patterns of

interaction with each tool, which limited the ability to evaluate the mechanisms through which learning gains may have occurred.

These limitations were addressed through careful analytical and interpretive strategies aimed at strengthening the credibility of the findings despite the constraints. First, the study emphasized cautious interpretation of results, avoiding causal claims and instead focusing on associations and patterns of performance differences between exposed and non-exposed groups. To mitigate issues arising from small sample size and sparse subgroup data, analyses were supplemented with aggregated comparisons and simplified modeling approaches that prioritized stability over complexity, ensuring that reported results remained statistically defensible. Generally, these mitigation strategies helped ensure that the study

Recommendations

For Teachers

It is recommended that teachers integrate DeltaMath and Quizizz consistently into mathematics instruction to enhance opportunities for practice,

still provided meaningful, contextually relevant insights into blended learning in a rural mathematics classroom, even while acknowledging the inherent methodological and data limitations.

Finally, future research should explore the long-term effects of blended digital interventions on mathematics achievement. Longitudinal studies may provide insights into whether the observed incremental gains translate into stronger improvements over extended periods of implementation. Additionally, qualitative research could investigate teachers' and students' perceptions of these tools to provide a deeper understanding of implementation challenges and successes. Comparative studies across different educational contexts would further enhance understanding of the transferability of findings and effective implementation practices.

provide immediate feedback, and foster learner engagement. Sustained use of these platforms is likely to support gradual improvements in mathematics performance over time.

For Assessment and Evaluation

Future evaluations of mathematics interventions should incorporate multiple measures of achievement, including continuous performance measures and summative assessment outcomes. This comprehensive approach facilitates a deeper understanding of intervention effects and acknowledges that even modest score improvements can be educationally significant.

For School Leaders

School leaders are encouraged to facilitate the effective implementation of blended learning by offering professional development and promoting the consistent integration of digital learning tools in classroom instruction. It is also essential to ensure that all learners have equitable access to these resources.

For Researchers

Future research should examine the mechanisms by which blended digital interventions affect mathematics achievement. Special consideration should be given to factors such as implementation quality, learner engagement, and instructional practices that may contribute to differences in student performance.

Further Research

Future studies should investigate the long-term effects of blended digital interventions on mathematics achievement. Longitudinal, qualitative, and comparative research across diverse educational contexts may yield deeper insights into effectiveness, implementation challenges, and best practices for technology-enhanced mathematics instruction.

Conclusion

This study compared continuous mathematics performance outcomes between learners who received a blended DeltaMath–Quizizz intervention and those who did not. Learners in the intervention group exhibited more favorable performance patterns, particularly in End-of-Course (EOC) assessment outcomes. While the observed differences did not consistently achieve conventional levels of statistical significance, the direction and magnitude of the effect sizes indicate modest but educationally meaningful differences in achievement. The results suggest that the blended intervention was associated with incremental improvements in cumulative mathematics performance.

through sustained practice, timely feedback, and increased learner engagement. These findings offer cautious empirical support for the integration of digital learning platforms

Author Contributions: The author acknowledges sole ownership, conceptualization of the study, design, data collection, analysis, interpretation, results, and manuscript preparation.

Funding: The researchers received no funding.

Conflict of interest: The authors declare no conflicts of interests

References

Agresti, A. (2018) Statistical Methods for the Social Sciences. 5th edn. Boston: Pearson.

Black, P. and Wiliam, D. (1998) 'Assessment and classroom learning', *Assessment in Education: Principles, Policy & Practice*, 5(1), pp. 7–74.

Bloom, B.S. (1976). *Human Characteristics and School Learning*. New York: McGraw-Hill.

Cheung, A.C.K. & Slavin, R.E. (2013) 'The effectiveness of educational

into mathematics instruction and highlight the need for further research with larger samples and more rigorous designs to confirm the robustness and generalisability of this effect.

Acknowledgements: The Researcher extends sincere appreciation to the Africa Research University for granting approval for the study as part of a PhD dissertation.

technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis', *Educational Research Review*, 9, pp. 88–113.

Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cumming, G. (2014) 'The new statistics: Why and how', *Psychological Science*, 25(1), pp. 7–29.

Faber, J., Luyten, J.W. and Visscher, A.J. (2017) 'The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomised experiment', *Computers & Education*, 106, pp. 83–96.

Field, A. (2018) *Discovering Statistics Using IBM SPSS Statistics*. 5th edn. London: SAGE.

Guskey, T.R. (2007) 'Closing achievement gaps: Revisiting Benjamin S. Bloom's "Learning for Mastery"', *Journal of Advanced Academics*, 19(1), pp. 8–31.

Hattie, J. (2009) *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London: Routledge.

Hattie, J., & Timperley, H. (2007) 'The power of feedback', *Review of Educational Research*, 77(1), pp. 81–112.

Higgins, S., Xiao, Z. and Katsipatakis, M. (2012) *The Impact of Digital Technology on Learning: A Summary for the Education Endowment Foundation*. Durham: Durham University.

Johnson, R.B. and Christensen, L. (2017) *Educational Research: Quantitative, Qualitative, and Mixed Approaches*. 6th edn. Thousand Oaks, CA: SAGE.

Li, Q. and Ma, X. (2010) 'A meta-analysis of the effects of computer technology on school students' mathematics learning', *Review of Educational Research*, 80(3), pp. 215–243.

Lipsey, M.W. and Wilson, D.B. (2001) *Practical Meta Analysis*. Thousand Oaks, CA: SAGE.

Mayer, R.E. (2009) *Multimedia Learning*. 2nd edn. Cambridge: Cambridge University Press.

Means, B., Toyama, Y., Murphy, R. and Baki, M. (2013) 'The effectiveness of online and blended learning: A meta-analysis of the empirical literature', *Teachers College Record*, 115(3), pp. 1–47.

Messick, S. (1995) 'Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning', *American Psychologist*, 50(9), pp. 741–749.

Nicol, D.J. and Macfarlane-Dick, D. (2006) 'Formative assessment and self-regulated learning: A model and seven principles of good feedback practice', *Studies in Higher Education*, 31(2), pp. 199–218.

Panadero, E. (2017) 'A review of self-regulated learning: Six models and four directions for research', *Frontiers in Psychology*, 8, Article 422.

Roschelle, J., Feng, M., Murphy, R. F., & Mason, C. A. (2016) 'Online mathematics homework increases student achievement', *AERA Open*, 2(4), pp. 1–12.

Sailer, M. and Homner, L. (2020) 'The gamification of learning: A meta-analysis', *Educational Psychology Review*, 32, pp. 77–112.

Shadish, W.R., Cook, T.D., and Campbell, D.T. (2002) *Experimental and Quasi-Experimental Designs for Generalised Causal Inference*. Boston: Houghton Mifflin Press.

Slavin, R.E. (2008) 'Perspectives on evidence-based research in education: What works? Issues in synthesising educational program evaluations. *Educational Researcher*, 37(1), 5–14.

Van der Kleij, F.M., Feskens, R.C.W., and Eggen, T.J.H.M. (2015) 'Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis', *Review of Educational Research*, 85(4), pp. 475–511.

Vygotsky, L.S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Wiliam, D. (2011) *Embedded Formative Assessment*. Bloomington, IN: Solution Tree Press.

Zimmerman, B.J. (2002) 'Becoming a self-regulated learner: An overview', *Theory Into Practice*, 41(2), pp. 64–70.