www.ijte.net

Enhancing Problem-solving Skills in AI Game-based Learning Environment through Structural Equation Modelling and Artificial Neural Network

Anass Bayaga
University of Western Cape Town, South Africa

To cite this article:

Bayaga, A. (2025). Enhancing problem-solving skills in AI game-based learning environment through structural equation modelling and artificial neural network. *International Journal of Technology in Education (IJTE), 8*(2), 574-589. https://doi.org/10.46328/ijte.1114

IJTE International Journal of Technology in Education
2025, Vol. 8, No. 2, 574-589 https://doi.org/10.46328/ijte.1114

Enhancing Problem-solving Skills in AI Game-based Learning Environment through Structural Equation Modelling and Artificial Neural Network

Anass Bayaga

Article Info

Article History

Received:
20 November 2024
Accepted:
21 April 2025

Keywords

Gamification
AI-driven technologies
Education
Higher-order thinking

Abstract

This investigation explored the role of artificial intelligence (AI)-powered gamification on mathematics cognition through a mixed-methods design, blending an intervention with a gamified learning application (app) and a survey to evaluate student engagement and performance. The study explores the nexus of gamification, AI, and mathematics cognition through 71 students participating in an intervention using gamified app. It was designed to enhance both computational thinking and mathematical skills. Both multi-group partial least squares (MGA-PLS) as well as artificial neural networks (ANN) through multilayer perceptron (MLP) were employed for data analysis. The findings showed a significant positive influence on not just class engagement, attitudes toward mathematics, but overall student performance. The developed model discerned indirect gender-related variations, which affirmed a transformative potential of gamification, particularly in preparing teachers for the AI-driven digital society. Consequently, the implication validates the transformative potential of gamification, teacher preparation for an AI-driven digital society. The implications concurrently emphasise integrating gamified elements into educational strategies. Such incorporations tend to benefit not just educators, curriculum developers, but policymakers as well. Thus, resonating the demands of the 21st-century teaching landscape.

Introduction

Integrating artificial intelligence (AI) and gamification in teaching and learning has increasingly received recognition for not just engagement but performance. AI-powered gamification, whether used in science, technology, engineering or mathematics (STEM) or general terms (see Table 1) is considered to be the incorporation of AI technologies in educational gamification strategies for the creation of an engaged as well as effective teaching learning experiences.

However, gamification itself tends to leverage on game design elements including points, badges, leaderboards, together with challenges in motivating and engaging teaching and learning processes. Combined, both can provide personalised and adapted experiences for users' needs, consequently making the teaching learning experience more effective and enjoyable. By adjusting the difficulty level of tasks and challenges based on the users' progress and performance, the combined effect also analyses' data and behaviour, which intend tailors the gamified learning experience to each users' preferences, strengths, and areas for improvement. On the other hand, mathematics problem-solving skills includes understanding, analysing, and solving mathematical problems through logical reasoning, critical thinking, and the application of mathematical concepts and techniques through key components as;

- **Understanding the Problem:** Comprehending what the problem is asking.
- **Planning a Strategy**: Deciding on the methods and steps needed to solve the problem**.**
- **Executing the Plan:** Applying mathematical techniques to carry out the plan.
- **Reviewing/Reflecting:** Checking the solution for accuracy and understanding the reasoning behind it.

Table 1. Examples of AI-powered Gamification in STEM and General Education

STEM Platform	STEM Description	General Platform	General Description
Carnegie Learning	Uses AI to provide personalised math tutoring and gamified practice sessions.	Duolingo	A language learning platform that uses AI to personalise lessons and exercises, incorporating gamified elements like streaks, points, and leaderboards.
ALEKS	An adaptive learning platform for math and science that uses AI to tailor instruction and provide gamified assessments.	Khan Academy	An educational platform that uses AI to create personalised learning paths, with gamified elements such as badges and points.
DreamBox Learning	An adaptive math program that uses AI to personalise learning experiences and includes gamified elements to keep students engaged.	Classcraft	A classroom management tool that gamifies the learning experience by turning the classroom into a role-playing game, using AI to track progress and customise challenges.
Zyrobotics	Offers gamified STEM learning tools that use AI to adapt to the needs of students with different learning abilities.	Smart Sparrow	An adaptive learning platform that uses AI to tailor educational content, with gamified features like interactive simulations and challenges.
CodeCombat	A platform that teaches coding through a gamified environment, using AI to adjust challenges based on the learner's progress.		

Number of reasons led to the current study. Though extensive receive has been conducted in unravelling different use cases of AI-powered gamification and mathematics problem-solving skills, however, key areas rife for research include the role of game-based learning (GBL) in increasing student engagement, and the relationship between computational thinking (CT) and mathematics education through AI-powered gamification and mathematics problem-solving skills. Additionally, the specific influences of these technologies on mathematics problem-solving skills, particularly regarding gender differences, remain underexplored.

This study addresses this gap by examining the gender-specific effects of AI-powered gamification on mathematical cognition, providing valuable insights for educators and policymakers. While there is yet little research to comprehend enhancement of computational problem-solving skills in AI powered environment, several attempts are underway. For instance, there have been three key themes. One theme that has received wide attention is AI in education, its role in students' performance and AI applications in teaching (Aljohani, 2019; Chen et al., 2020; Huang, 2023; Lee, 2020; Kim et al. 2019; Krstić et al., 2022). The second theme, GBL and engagement proposed by Gonzalez et al. (2017), Vygotski (1978), Wu and Yang (2022), Ye et al. (2023) and Zovko and Gudlin (2019) tends to be contextualised under use of game-based learning, design frameworks, engagement strategies. The third theme was computational thinking in mathematics education in the context of relationship between computational thinking and mathematics education proposed by Durksen et al. (2017), Roman-Gonzalez et al. (2017), Vygotski (1978), Wu and Yang (2022), Ye et al. (2023), Zovko and Gudlin (2019) as well as GeoGebra and mobile (m) - learning for learning and teaching (Mthethwa, Bayaga, Bossé & Williams, 2020; Mutambara & Bayaga, 2020).

The first theme as suggested by Chen et al. (2020), Zovko and Gudlin (2019) identifies gaps in AI application and theory in education, areas needing research. Yet, the second warrants the exploration in gaps in understanding GBL, need for more comprehensive studies as suggested by Adipat et al. (2021) and Tsarava et al. (2017), while Roman-Gonzalez et al. (2017), Wu and Yang (2022), Ye et al. (2023), Zovko and Gudlin (2019) identify and suggest integrating computational thinking in mathematics education. Other studies such as Aljohani (2019) suggest careful analysis of implications for the role of AI in education and its role on student learning as with practical implications. While Lee and Hannafin (2016) developed a design framework for enhancing engagement in student-centered learning, the suggestion is need for framework for enhancing engagement in student-centered learning. On the other hand, Durksen et al. (2017) investigated motivation and engagement in mathematics through a qualitative framework for teacher-student interactions through qualitative framework for teacher-student interactions to explore motivation and engagement and concluded that framework for understanding motivation and engagement in mathematics education is important. These three themes have unanimously proposed the examination of the hypothesis of integrating AI to enhance student performance as well as adoption of AI to improve performance through GBL, hence the current study.

Background and Development of the Model

Based on the background, three key themes are discussed: Research suggests the integration of AI to enhance student performance (Aljohani, 2019; Krstić et al. 2022). Collectively, the studies underscore a dual narrative that is, integration of AI is both transformative and bolsters content skills, which intend leads to improved educational performance (Chen et al., 2020; Kim et al., 2019; Lee, 2020). Additional studies based on GBL amplify' student engagement (Gonzalez et al., 2017; Wu & Yang, 2022). The assertion is that implementation GBL not only captivates but also sustains student interest which intend elevates the learning experience. The last theme is computational thinking (CT) by Durksen et al. (2017) and Roman-Gonzalez et al. (2017) advocates that the cultivating CT skills significantly improves mathematical understanding and prowess. Regardless of the articulation of the themes, critical research gaps persist. For instance, we still do not have firm inquiry into both practical and theoretical underpinnings of AI through GBL (Chen et al., 2020; Zovko and Gudlin, 2019). Equally, investigating the dynamic range of teacher-student interactions in the midst of AI-GBL remains under researched ((Adipat et al., 2021; Durksen et al., 2017; Tsarava et al., 2017).

The are several ramifications worthy of considering. The implication so long is that integrating of AI, the employment of GBL, and the development of CT are not just additive but synergistic in enhancing the educational landscape. Educators are also urged to harness such innovation not in isolation but in concert to create a more dynamic, engaging, and effective learning environment. Such an approach promises to captivate and elevate academic achievements. Though the aforementioned sources highlight AI, GBL, and CT into the educational curriculum, however, there are notable gaps for future research. In terms of AI, Aljohani (2019) and Krstić et al. (2022) suggest a positive influence with student performance and teaching methodologies. Nevertheless, the recognised gap in the application of AI in education is limited (Chen et al., 2020; Zovko & Gudlin, 2019). While theoretically, it is implied that AI has transformative role, yet practically, careful integration of AI to enhance student skills is limited. In terms of GB L and engagement, Lee and Hannafin (2016) highlighted the role of GBL in increasing student engagement. Though theoretically that appears to be accurate in terms of GBL enhancing student-centered learning strategies, yet in practice, guidelines for educational experience design are not firmly established. That is, there are still gaps remaining in understanding the full role of GBL on learning outcomes, needing further studies (Adipat et al., 2021; Tsarava et al., 2017). On the part of CT, it has long been established that the relationship between CT and mathematics education is positively related (Durksen et al., 2017). Additionally, the theoretical implications suggest teacher-student interactions importance in enhancing motivation and engagement in mathematics. Yet, the gap is that the integration of CT into K-12 mathematics education, remains underexplored (Roman-Gonzalez et al., 2017; Wu & Yang, 2022; Ye et al., 2023). Comparative analysis this far suggests the need to bridge these identified gaps leading to the objective of examining the degree to which to enhance mathematics problem-solving skills in AI-powered GBL environment.

In conclusion, while the benefits of AI, GBL, and CT in education are increasingly recognised, substantial research gaps remain. Addressing these gaps could significantly contribute to the creation of more engaging, effective, and equitable educational experiences. Future research must strive to close these gaps with innovative methodologies and practical applications that are informed by the evolving landscape of technology in education. Particularly notable is the use of multi-group analysis within PLS-SEM, known as MGA-PLS, as the analysis considered gender categories (male and female) to explore potential variations in the structural relationships among constructs. PLS-SEM, including MGA-PLS, was chosen based on the work of Hair, Hult, Ringle, and Sarstedt (2022), who argued its suitability for assessing challenging-to-measure and unobservable latent variables. Particularly well-suited for analysing both direct and indirect effects, PLS-SEM accommodates mediated (and moderated) relationships, as illustrated in Figure 1. This conceptual model explores various aspects of learning and understanding in mathematics, addressing the effectiveness of problem-solving activities, the role of prior knowledge, the influence of instructional strategies, and the use of analogical reasoning.

In support of the Figure 1, various studies such as but not exclusively Aljohani (2019) and Krstić et al. (2022) suggested that AI enhances student performance, but Gonzalez et al. (2017) and Wu and Yang (2022) further highlight the probable additional effect of GBL engagement. Furthermore, others such as Durksen et al. (2017) and Roman-Gonzalez et al. (2017) have also advocated for the role of CT in improving mathematical understanding. Such ongoing debates inform the basis of the hypotheses that gender moderates the relationship between AI-powered gamification and mathematical cognition. At the moment, limited studies have unravelled gender differences in AI-powered gamification and mathematical cognition, nor do we comprehend how specifically, to what extend female learners differ from male considering the model. Consequently, it is hypothesised that female learners will exhibit a stronger relationship between AI-powered gamification and mathematical performance due to enhanced engagement and motivational aspects.

As depicted in the model, the conceptual model examines different pathways. That it is posited that Mathematical and Computational Algorithms (MCA) and Mathematical Modelling and Simulation (MMS) are antecedent constructs, which influence Abstract and Concrete Representations of Mathematical Principles (MP). MP, in turn, mediate the relationship between MCA and MMS and the Analogical Comparison Principle (ACP). Furthermore, notice the direct effects both from MCA and MMS to ACP as an indication of the constructs immediate influence on analogical comparison abilities beyond their indirect influence through MP. The consideration of both pathways (direct and indirect) model tends to facilitate the examination necessary and sufficient conditional relationships between computational algorithms, modelling and simulation techniques, as well as role in the development of analogical reasoning within mathematical learning contexts.

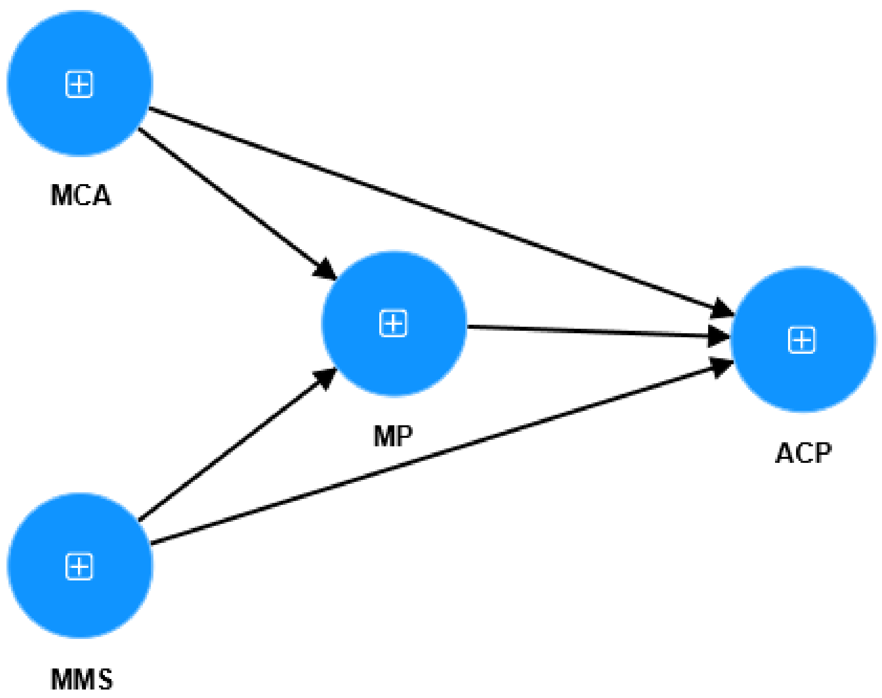

Figure. 1. Model of Gamification and Enhanced -AI Mathematics Cognition.

The results from the current model could have several implications. For instance, if a path is stronger for one gender, it suggests more effective engagement that gender.

Research Question

Guided by the model (see Fig 1), the question that arises is how do MCA, MMS and MP influence ACP on mathematical cognition across gender groups considering AI enabled gamification and computational thinking?

Hypothesis

Thus, the study explores the intricate connections between MMS, MCA, MP, and ACP, focusing on gender differences. It is then hypothesized that while measurement properties may slightly vary, they remain largely consistent across genders. The relationships between MCA, MMS, MP, and ACP are expected to differ between male and female groups. We anticipate gender-specific variations in the mediating role of MP and the total indirect effects of MCA and MMS on ACP. This hypothesis aims to uncover nuanced insights into gender-specific cognitive processes in mathematical learning.

H1: The relationship between gamification elements and MCA will be moderated by gender, with female learners exhibiting a stronger relationship as compared to male learners.

H2: Gender will moderate the relationship between MCA and mathematical performance (MP), with male learners demonstrating a higher path coefficient, suggesting a more significant impact of motivation on performance due to gamification.

Methodology

The research evaluated the following variables: student engagement, mathematical problem-solving skills, and computational thinking abilities. These were evaluated utilising a sequence of observational checklists and self-reported surveys. The observational checklists focused on specific behaviors indicative of engagement, such as involvement in gamified tasks and completion of problem-solving exercises. The surveys were conceived to capture students' noticed engagement and their self-efficacy in mathematics. This research employed a mixed-methods approach, integrating mutually an intervention and a survey design. The intervention necessitated a gamified learning app operated over a period of two weeks to improve mathematical cognition. Simultaneously, a survey was administered to gather data on student experiences and engagement levels before and after the intervention

The methodology employed in this study aligns with the research hypotheses, utilizing a survey design involving 71 learners, complemented by an observational study conducted by the researcher. The collected data underwent meticulous analysis, encompassing demographic information such as age categories (ranging from 15-17 to 30 or more), gender identification (Male, Female), current education levels (High school, Certificate, Diploma, Degree, Postgraduate), and geolocation (Urban, Rural) of the participants. Notably, a MGA-PLS was applied, distinguishing between gender categories (Male and Female) to capture potential variations in the relationships between constructs for different gender groups as reflected in Figure 1. The study focused on key themes: MMS, MCA, MP, and the ACP. For each theme, specific indicators were outlined, accompanied by associated indicator questions. For instance, within the MMS theme, indicators emphasised understanding dependent on problem-solving activities. Analogously, MCA indicators focussed on the fact that understanding is improved when explaining the process used in answering questions. Jointly, both constructs and indicators formed a comprehensive model. Ethical considerations obtained from institutional review board approval from XXX University (penname) (Reference: H21-EDU-PGE-026) was secured and clearance from the research ethics committee of XXX University's Faculty of Education with rigorous steps underscored the commitment to upholding ethical standards and ensuring participant protection in the research endeavour by ensuring, privacy, anonymity, voluntary withdrawal, and confidentiality of participants information. Introduction of computational thinking processes through Table 1 gameplay was employed to enhance problem-solving skills for understanding of mathematical concepts, and overall student performance. The study employed MGA using the PLS approach, which allowed for the comparison of path coefficients between distinct groups, provided mechanism for detecting variations in the effects of the independent variables across different dependent variables (Hair et al., 2022).

Leveraging PLS-SEM made it possible to assess hypothesized model across different groups systematically and simultaneously (Henseler, Ringle, & Sinkovics, 2009), thus examined gender-specific variations in the structural relationships among constructs. Moreover, artificial neural networks (ANN) through multilayer perceptron (MLP), with its multiple layers of interconnected nodes or neurons were employed in modelling complex non-linear relationships and predict outcomes (Sarstedt, Ringle, & Hair, 2017), which allowed for an analysis that not only highlighted the key factors driving the effectiveness of educational setting but also uncovered non-linear relationships between these factors. The dual approach validated hypotheses with a higher degree of accuracy. The MLP is structured with an input layer, wherein neurons represent input data features, followed by one, sometimes more hidden layers in which each neuron processes a weighted sum of inputs using an activation function. The final layer, termed as the output layer, produces the network's output. Each connection in this network has a weight, and each neuron includes a bias term, both of which are adjusted during training to minimise prediction errors.

In this study, the common activation functions used in MLP include sigmoid hyperbolic tangent (tanh), and rectified linear unit (ReLU). The process of implementing an MLP in the current study involved several steps:

- preprocessing the data (which included cleaning, normalising, and splitting),
- defining the model architecture (that is, number of layers and neurons, and activation functions), compiling the model (meaning, choosing the optimiser, loss function, and evaluation metric),
- training (using backpropagation and optimisation algorithms), and
- finally, evaluating the model's performance on a test set using metrics appropriate to the research question (such as accuracy or mean squared error) (Hair et al., 2017).

Results

Assessment of Measurement Model

The measurement model revealed all constructs-ACP, MCA, MMS, and MP-were reliably (internal consistency) measured through Cronbach's alpha (α) and composite reliability (CR) with scores exceeding the recommended threshold of 0.7 (Hair et al., 2022). Convergent validity is established using average variance extracted (AVE) with threshold above 0.5. Discriminant validity is confirmed through both the HTMT criterion, with values below the threshold of 0.85, as well as the Fornell-Larcker criterion, signifying that construct share more variance with their own indicators than with other constructs (Fornell & Larcker, 1981). For the entire sample, the R^2 for ACP is 0.53 (R^2 adjusted = 0.51), indicating that the model elucidates 53% of the variance in ACP. Regarding MP, the R^2 is 0.17 (R^2 adjusted = 0.14), explaining 17% of the variance. Gender-specific analyses reveal that for females, the R^2 for ACP is 0.66 (R^2 adjusted = 0.62), while for males, it is 0.61 (R^2 adjusted = 0.57). In contrast, the R^2 for MP is 0.15 (R^2 adjusted = 0.09) for females and 0.25 (R^2 adjusted = 0.20) for males, indicating varying explanatory power for ACP and MP across gender groups.

Assessment of Structural Model

The examination of the structural model indicates that all Variance Inflation Factor (VIF) values are well within acceptable limits, signifying the absence of multicollinearity issues within the model. Specifically, the VIF values associated with each path - MCA to ACP, MCA to MP, MMS to ACP, MMS to MP, and MP to ACP - are all below the threshold of 5, affirming the model's validity across the complete, female, and male samples. Table 2 presents the path coefficients for the complete sample and gender-specific subgroups, highlighting the differential influence of mathematical constructs on analogical reasoning and problem-solving.

Table 2. Direct Relation

	Paths	MCA -> ACP	MCA -> MP	MMS -> ACP	MMS -> MP	MP -> ACP
Complete	Beta	0.73	0.37	-0.07	-0.14	-0.03
	SD	0.08	0.17	0.1	0.15	0.11
	T values	9.54	2.12	0.66	0.91	0.28
	P values	0	0.03	0.51	0.36	0.78
Female	Beta	0.85	0.39	-0.1	0.02	-0.17
	SD	0.09	0.32	0.13	0.27	0.15
	T values	9.13	1.22	0.76	0.09	1.14
	P values	0	0.22	0.45	0.93	0.25
Male	Beta	0.67	0.39	0	-0.25	0.21
	SD	0.12	0.21	0.14	0.19	0.14
	T values	5.62	1.81	0.03	1.28	1.53
	P values	0	0.07	0.98	0.2	0.13

For the complete sample, the path from MCA to ACP exhibits a significantly positive effect ($\beta = .73$, SD = .08, $t = 9.54$, $p < .001$). This effect is even more pronounced in the female subgroup ($\beta = .85$, SD = .09, $t = 9.13$, $p < .001$) but slightly reduced in the male subgroup ($\beta = .67$, SD = .12, $t = 5.62$, $p < .001$). The influence of MCA on MP is significant for the complete sample ($\beta = .37$, SD = .17, $t = 2.12$, $p = .03$) but not for the female ($\beta = .39$, SD = .32, $t = 1.22$, $p = .22$) and male subgroups ($\beta = .39$, SD = .21, $t = 1.81$, $p = .07$), suggesting gender-related variability in the influence of MCA on MP. Conversely, the relationship between MMS and ACP is non-significant across all samples: complete ($\beta = -.07$, SD = .10, $t = 0.66$, $p = .51$), female ($\beta = -.10$, SD = .13, $t = 0.76$, p = .45), and male (β = 0, SD = .14, t = 0.03, p = .98). Similarly, the effect of MMS on MP is not significant in any group: complete (β = -.14, SD = .15, t = 0.91, p = .36), female (β = .02, SD = .27, t = 0.09, p = .93), and male (β = -.25, SD = .19, t = 1.28, p = .20). Lastly, the path from MP to ACP exhibits a non-significant effect in the complete (β = -.03, SD = .11, t = 0.28, p = .78) and female samples (β = -.17, SD = .15, t = 1.14, p = .25), but a marginally non-significant positive effect in the male sample (β = .21, SD = .14, t = 1.53, p = .13). In summary, these findings underscore a robust influence of MCA on ACP, especially in females, while other relationships display mixed or non-significant effects, emphasising the potential influence of gender on these pathways.

The analysis of mediated effects, encompassing total indirect effects and specific indirect effects, was conducted for the entire sample and gender subgroups. For the complete sample, the indirect effect of MCA on ACP through MP was found to be non-significant (β = -0.01, SD = 0.05, t = 0.24, p = 0.81), indicating that the mediation of MP does not significantly influence the relationship between MCA and ACP. Similarly, the indirect effect of MMS on ACP through MP was also non-significant (β = 0, SD = 0.02, t = 0.18, p = 0.86), suggesting that MP does not mediate the relationship between MMS and ACP in the complete sample. In the female subgroup, both the indirect effect of MCA on ACP through MP (β = -0.07, SD = 0.09, t = 0.76, p = 0.45) and the indirect effect of MMS on ACP through MP (β = 0, SD = 0.05, t = 0.08, p = 0.94) remained non-significant. In the male subgroup, the indirect effect of MCA on ACP through MP became marginally significant (β = 0.08, SD = 0.08, t = 1.05, p = 0.29), suggesting a potential mediation effect in this group. However, the indirect effect of MMS on ACP through MP remained non-significant (β = -0.05, SD = 0.06, t = 0.81, p = 0.42). In summary, the mediation analysis indicates that MP does not significantly mediate the relationship between MCA and ACP or between MMS and ACP for the complete sample and the female subgroup. In the male subgroup, there is a marginal indication of mediation between MCA and ACP through MP.

Bootstrapping Multigroup Analysis

Assessment of Measurement Invariance

In the context of multigroup analysis using MGA-PLS, an evaluation of measurement invariance was conducted, as summarised in Table 2. The findings reveal that the measurement properties of the constructs ACP, MCA, MMS, and MP demonstrate a substantial level of invariance across gender groups. Minor variations were observed, but these do not exert a significant influence on the overall measurement invariance.

Table 3 presents an evaluation of measurement invariance for the constructs ACP, MCA, MMS, and MP across gender groups. The original correlations (original correlation) indicate the initial correlation coefficients for each construct. The correlation per-mutation mean represents the average correlation coefficients obtained through permutations, providing insight into the stability of the measures. The subsequent columns (5.00%, Permutation p value, 2.50%, 97.50%, Permutation p value) offer information about the significance of the differences observed in the permutation mean compared to the original correlation. The presented values demonstrate that the measurement properties of the constructs maintain a substantial level of invariance across genders, with minor variations that do not significantly impact measurement invariance.

Table 3. Evaluating Measurement Invariance

		ACP	MCA	MMS	MP
MICOM STEP 2	Original correlation	1	0.99	0.93	0.99
	Correlation permutation mean	1	1	0.91	0.98
	5.00%	0.99	0.99	0.58	0.93
	Permutation p value	0.13	0.22	0.3	0.42
MICOM STEP 3a (mean)	Original difference	-0.17	0.35	-0.22	0.28
	Permutation mean difference	0.01	0.01	-0.01	0.01
	2.50%	-0.47	-0.46	-0.46	-0.47
	97.50%	0.47	0.43	0.47	0.49
	Permutation p value	0.5	0.14	0.37	0.24
MICOM STEP 3b (variance)	Original difference	-0.06	0.4	0.05	0.16
	Permutation mean difference	0.01	0	-0.01	0.02
	2.50%	-0.6	-0.81	-0.62	-1.17
	97.50%	0.66	0.86	0.62	1.33
	Permutation p value	0.85	0.43	0.86	0.8

In MGA-PLS, gender differences in specific paths within the structural model were examined through Bootstrap MGA, parametric tests, and Welch Satterthwaite tests. The results indicated minimal gender discrepancies in the paths from MCA to ACP, MCA to MP, MMS to ACP, MMS to MP, and MP to ACP. The differences in the paths ranged from 0 to 0.38, with associated p values exceeding the significance threshold, implying non-significant gender variations in the examined paths. The findings suggest the robustness of the measurement model and the absence of substantial gender-related differences in the specific paths of the structural model (see Table 4).

Table 4. Bootstrap MGA, Parametric Test, Welch Satterthwaite

		MCA -> ACP	MCA -> MP	MMS -> ACP	MMS -> MP	MP -> ACP
Bootstrap MGA	Difference (Female - male)	0.18	0	-0.1	0.27	-0.38
	1-tailed (Female vs male) p value	0.1	0.46	0.69	0.22	0.97
	2-tailed (Female vs male) p value	0.2	0.93	0.61	0.45	0.07
Parametric test	Difference (Female - male)	0.18	0	-0.1	0.27	-0.38
	t value (|Female vs male|)	1.19	0	0.51	0.82	1.89
	p value (Female vs male)	0.24	1	0.61	0.41	0.06
Welch Satterthwaite test	Difference (Female - male)	0.18	0	-0.1	0.27	-0.4
	t value (|Female vs male|)	1.18	0	0.51	0.83	1.9
	p value (Female vs male)	0.25	1	0.61	0.41	0.07

Artificial Neural Network (ANN), through Multilayer Perceptron (MLP)

In this study, a MLP ANN was deployed for predictive modelling, aiming to forecast categorical outcomes across seven dependent variables (ACP1, ACP2, ACP3, ACP4, ACP5, ACP7, ACP6) based on eight input factors (MCA1 through MCA8) (Table 5). The utilised dataset comprised 72 entries. The dataset was partitioned into a training set (comprising 70.3% of cases, n = 45) and a testing set (comprising 29.7% of cases, n = 19), following a 7:3 ratio. The MLP ANN architecture consisted of an input layer with eight units corresponding to the input factors, a single hidden layer comprising 15 units, activated by the hyperbolic tangent function, and an output layer comprising seven units activated by the softmax function. The model's training employed the cross-entropy error function, and optimization criteria were based on scaled conjugate gradient descent. During training, the MLP ANN demonstrated a training cross-entropy error of 74.671, yielding an average percent incorrect prediction of 7.0%. In the testing phase, the model exhibited a cross-entropy error of 57.729, resulting in an average percent incorrect prediction of 12.0%.

Table 5. Constructs Independent Variable Importance

Constructs	Importance	Normalised Importance
MCA1	.128	83.9%
MCA2	.152	100.0%
MCA3	.086	56.4%
MCA4	.129	84.9%
MCA5	.118	77.5%
MCA6	.109	71.3%
MCA7	.137	90.3%
MCA8	.141	92.7%

Discussion

The study set out to examine AI-powered GBL and mathematics cognition through 71 students through gender-specific subgroups. For the path from MCA to ACP, the Beta values are 0.73 for the complete sample, 0.85 for females, and 0.67 for males, indicating a strong positive ($p < .001$), suggesting robustness across all groups. Similarly, the path from MCA to MP shows significance in the complete sample (Beta = 0.37, $p = 0.03$) but exhibits variability in gender subgroups, with significance in females (Beta = 0.39, $p = 0.22$) and a marginally significant effect in males (Beta = 0.39, $p = 0.07$). The paths from MMS to ACP and MP, as well as from MP to ACP, generally lack significance across all samples, suggesting mixed or negligible effects as reflected in Table 6.

Table 6. Key Elements and Contributions on AI enabled GBL in Mathematics Cognition

Category	Details
Key Hypotheses	- H1: The relationship between gamification elements and motivational cognitive aspects MCA will be moderated by gender, with female learners exhibiting a stronger relationship compared to male learners. - H2: Gender will moderate the relationship between motivational cognitive aspects

Category	Details
	MCA and MP, with male learners demonstrating a higher path coefficient.
Key Gaps in Literature	- Lack of comprehensive studies on the integration of AI in educational settings (Aljohani, 2019; Chen et al., 2020; Huang, 2023; Sparks, 2023; Krstić et al., 2022; Zovko & Gudlin, 2019). - Need for more in-depth studies on the full impact of GBL on learning outcomes (Adipat et al., 2021; Gonzalez et al., 2017; Tsarava et al., 2017; Vygotski, 1978; Wu & Yang, 2022; Ye et al., 2023; Zovko & Gudlin, 2019). - Underexplored integration of CT into K-12 mathematics education (Durksen et al., 2017; Roman-Gonzalez et al., 2017; Wu & Yang, 2022; Ye et al., 2023;).
Key Findings Using Pathways	- Significant positive impact of gamification on class engagement, attitudes toward mathematics, and overall student performance. → Improved student performance and engagement. - Subtle gender-related variations affirming the model's consistency across diverse groups. → Gender-specific effects in gamification.
Contributions	- Further research on the long-term impacts of gamification on mathematical cognition. → Investigate long-term effects. - Exploration of other demographic variables such as age and socioeconomic status in the context of gamified learning environments. → Broaden demographic scope. - Gender-specific strategies in educational design where gamification could be tailored to bridge cognitive gaps or leverage motivational strengths. → Develop tailored strategies - Current findings do not support significant relationships between MMS and MP or ACP across all samples. → Reevaluate pathways - Non-significant effects of MMS on MP and ACP suggest that certain pathways in the current model might be redundant or need to be reconsidered. → Reconsider model components
Theoretical and Practical Implications	- The integration of AI and gamification in education is transformative, enhancing content skills and educational performance. → AI and gamification impact. Development of computational thinking skills significantly contributes to mathematical understanding and prowess. → Enhance computational thinking - Educators and curriculum developers should consider integrating gamified elements into mathematics education to boost engagement and performance, particularly for female students. → Implement gamified elements - Importance of adaptive learning systems tailored to the developmental stage of learners to maximize cognitive engagement and learning outcomes. → Tailor adaptive learning systems.

Overall, these results highlight the differential influence of mathematical constructs on analogical reasoning and problem-solving, with gender-related nuances in specific relationships. The results presented in Table 1 reveal significant direct relations within the structural model for the complete sample and gender subgroups. Specifically, the paths from MCA to ACP demonstrate a substantial positive effect, which is notably pronounced in females. However, the impact of MCA on MP is significant only in the complete sample, indicating gender-related variability. The paths from MMS to ACP and MP, as well as from MP to ACP, mostly lack significance across all samples, suggesting mixed or negligible effects in these relationships (see Table 1 for details). Table 2 provides an assessment of measurement invariance using MICOM steps. The results indicate a substantial level of invariance across gender groups for the constructs ACP, MCA, MMS, and MP. Minor variations observed in correlation coefficients and permutation mean differences do not significantly impact measurement invariance, reinforcing the reliability of the measurement model (see Table 2 for details). Furthermore, Table 3 displays the Bootstrap MGA results, revealing minimal gender differences in specific paths within the structural model. The non-significant differences in paths from MCA to ACP, MCA to MP, MMS to ACP, MMS to MP, and MP to ACP, with p values exceeding the significance threshold, suggest the model's consistency and the absence of substantial gender-related variations in the examined paths (see Table 3 for details). In this study, MLP ANN was employed to predict categorical outcomes represented by seven dependent variables based on eight input factors. The dataset consisted of 72 rows, with a training-testing split of 7:3, resulting in 45 cases for training and 19 for testing. The architecture of the MLP ANN included an input layer with eight units corresponding to the input factors, a single hidden layer with 15 units using the hyperbolic tangent activation function, and an output layer with seven units employing the softmax activation function. The error function used for training was cross-entropy, and the optimization criteria involved scaled conjugate gradient descent. The model demonstrated satisfactory performance during training, with a training cross-entropy error of 74.671 and an average percent incorrect prediction of 7.0%. The testing phase exhibited a cross-entropy error of 57.729 and an average percent incorrect prediction of 12.0%. Notably, the overall percent correct for training and testing were 93.0% and 88.0%, respectively. The classification results for each dependent variable revealed varying levels of accuracy, with ACP3 achieving 100% accuracy during training. Additionally, ROC curves were employed to evaluate the model's discrimination ability, demonstrating high AUC values for each dependent variable.

Summary

In summary, this study (Table 7) explores the integration of AI and gamification in enhancing mathematics problem-solving skills. It specifically examines gender-specific effects, aiming to fill the gaps in existing research that has not sufficiently addressed the influence of these technologies on mathematical cognition. The research employs a combination of SEM and ANN to analyse data and draw conclusions. Key findings indicate that AI-enabled gamification significantly improves student engagement and performance, with notable gender-specific variations.

Table 7. Enhancing Problem-solving Skills in an AI game-based Learning Environment

Section	Details
Background	The study addresses the gap in understanding how AI-enabled gamification can enhance mathematics problem-solving skills, particularly focusing on gender-specific effects. Existing research has not sufficiently explored the influence of these technologies on

Section	Details
	mathematical cognition, making this study crucial for educators and policymakers. Key sources include Aljohani (2019), Chen et al. (2020), and Roman-Gonzalez et al. (2017), which highlight the transformative role of AI and the need for integrating computational thinking into mathematics education.
Objectives	The primary aim is to examine the gender-specific effects of AI-enabled gamification on mathematical cognition. Key findings show significant positive impacts on student performance and engagement, particularly in female students.
Key Findings	The study found that AI-enabled gamification significantly improves student engagement and performance in mathematics. Gender-specific variations were observed, with females showing a stronger relationship between gamification elements and mathematical performance.
Pathways	- Gamification elements ➜ Increased engagement ➜ Improved performance- AI-enabled learning ➜ Enhanced cognitive skills ➜ Better problem-solving- Gender-specific analysis ➜ Females exhibit stronger relationship ➜ Tailored educational strategies
Results and Conclusions	The findings highlight the transformative potential of gamification in education, particularly for female students, by enhancing engagement and performance. The study suggests that incorporating gamified elements into educational strategies can benefit educators, curriculum developers, and policymakers, meeting the demands of the 21st-century teaching landscape. The research emphasises the need for further studies to explore other demographic variables and long-term impacts of gamification on mathematical cognition.
What is currently known about this topic?	Existing research indicates that AI and gamification can enhance student engagement and performance, but there is limited understanding of their gender-specific effects and long-term impacts on mathematical cognition.
What does this paper add?	This paper provides new insights into the gender-specific effects of AI-enabled gamification on mathematical cognition, highlighting significant improvements in student engagement and performance, particularly among female students.
Implications for practice/or policy	The study suggests that integrating gamified elements into educational strategies can significantly benefit educators, curriculum developers, and policymakers. Tailored approaches based on gender-specific analysis can enhance student engagement and performance, meeting the evolving demands of the 21st-century educational landscape. Further research is needed to explore other demographic variables and long-term impacts of these technologies.

Concluding and Future Work

This study highlights the significant positive influence of gamification and AI on educational outcomes, emphasising the need for tailored strategies and further research to explore demographic variables and long-term effects. The findings advocate for the integration of these innovative approaches to enhance engagement, performance, and computational thinking skills in educational settings. Future investigation should aim to expand the methodological methods utilised in studies investigating AI-powered gamification to ensure precision and robustness in design, specifically in deciding between intervention and survey-based approaches.

Statements and Declaration

Conflict of Interest: None

Funding: Declare no funding for current study.

Data Availability Statement: The datasets presented in this article are readily available on reasonable requests directed to the corresponding author.

Ethical approval: The study protocol followed ethical best practices, was reviewed and revised by a panel of experts, and approved by the faculty and university board

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author utilised quilbot in order to improve readability after which the author reviewed and edited the content as needed and took full responsibility for the content of the published article.

References

Adipat, S., Laksana, K., Busayanon, K., Asawasowan, A., & Adipat, B. (2021). Engaging students in the learning process with game-based learning: The fundamental concepts. *International Journal of Technology in Education*, *4*(3), 542-552. https://doi.org/10.46328/ijte.169

Benton, L., Kalas, I., Saunders, P., Hoyles, C., & Noss, R. (2018). Beyond jam sandwiches and cups of tea: An exploration of primary pupils algorithm-evaluation strategies. *Journal of Computer Assisted Learning*, *34*, 590-601. https://doi.org/10.1111/jcal.12266

Durksen, T., Way, J., Bobis, J., Anderson, J., Skilling, K., & Martin, A. (2017). Motivation and engagement in mathematics: A qualitative frameworkfor teacher-student interactions. *Mathematics Education Research Journal*, *29*(2). https://doi.org/10.1007/s13394-017-0199-1

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.2307/3151312

Greyling, J., & Tokosi, T. (2020, September 14). *Investigating the roles and competence of a Tangible game facilitator* [Paper presented at SAICSIT'20: Conference of the South African Institute of Computer Scientist and Information Technologies]. Cape Town, South Africa. https://doi.org/101145/3410886.3410894

Hagan, J. E., Amoaddai, S., Lawer, V. T., & Atteh, E. (2020). Students' Perception towards Mathematics and Its Effects on Academic Performance. *Asian Journal of Education and Social Studies*, *8*(1), 8-14. https://doi.org/10.9734/AJESS/2020/v8i130210

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). A primer on partial least squares structural equation modeling (PLS-SEM) (3rd ed.). Thousand Oaks, CA: Sage

Hamlen, K. R. (2013). Understanding children's choices and cognition in video game play A synthesis of three studies. *Zeitschrift fur Pscychologie*, *221*(3), 107-114. https://doi.org/10.1027/2151-2604/a000136

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing*, 20, 277-319.

Kim, S., Raza, M., & Seidman, E. (2019). Improving 21st-century teaching skills: The key to effective 21st-century learners. *Comparative and International Education*, *14*(1), 99-117. https://doi.org/10.4018/jig.2011070103

Mthethwa, M., Bayaga, A., Bossé, M.J., & Williams, D. (2020). GeoGebra for learning and teaching: A parallel investigation. *South African Journal of Education*, *40*(2), 1-12. https://doi.org/10.157/saje.v40nza1669

Mutambara, D. & Bayaga, A. (2020). Rural-based Science, Technology, Engineering and Mathematics teachers' and learners' acceptance of mobile learning. *South African Journal of Information Management, 22*(1), a1200. https://doi.org/10.4102/sajim.v22i1.1200

Ringle, C. M., & Sarstedt, M. (2016). Gain more insight from your PLS-SEM results: The importance-performance map analysis. Industrial Management & Data Systems.

Roman-Gonzalez, M., Perez-Gonzalez, J., & Jimenez-Fernandez, C. (2017). Which cognitive abilities underlie computational thinking? Criterion validity of the Computational Thinking Test. *Computers in Human Behavior*, *72*(2), 678-691. https://doi.org/10.1013/j.chb.2016.08.047

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2017). Partial least squares structural equation modeling. In Handbook of Market Research. Springer, Cham.

Tsarava, K., Moeller, K., Pinkwart, N., Butz, M., Trautwein, U., & Ninaus, M. (2017, October 1). *Training Computational Thinking: Game-Based Unplugged and Plugged-in Activities in Primary School* [Conference Presentation]. 11th European Conference in Game-Based Learning.

Vygotski, L. (1978). *Mind in Society. Development of Higher Psychological Processes*. Harvard University Press.

Wu, W., & Yang, K. (2022). The relationships between computational and mathematical thinking:A review study on tasks. *Cogent Education*, *9*(1), 1-19. https://doi.org/10.1080/23386X.2022.2098929

Ye, H., Liang, B., Ng, O., & Chai, C. (2023). Integration of computational thinking in K-12 mathematics education: A systematic review on CT-based mathematics instruction and student learning. *International Journal of STEM Education*, *10*(3), 1-26. https://doi.org/10.1186/s40594-023-00396-w

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, 14(1), 35-62.

Author Information

Anass Bayaga

https://orcid.org/0000-0003-0283-0262

University of Western Cape

South Africa

Contact e-mail: abayaga@uwc.ac.za