



JOURNAL OF DIGITAL SYSTEM DEVELOPMENT

<https://e-journal.uum.edu.my/index.php/jdsd>

How to cite this article:

Annamalai, S. (2026). Design, Development and Evaluation of a Two-Level AI-Generated Quiz Game for Teaching Multimedia Elements using Gemini Canvas. *Journal of Digital System Development*, 4(2), 78–90. <https://doi.org/10.32890/jdsd2026.4.2.6>

DESIGN, DEVELOPMENT AND EVALUATION OF A TWO-LEVEL AI-GENERATED QUIZ GAME FOR TEACHING MULTIMEDIA ELEMENTS USING GEMINI CANVAS

¹Subashini Annamalai

¹School of Multimedia Technology and Communication, Universiti Utara Malaysia, Malaysia

¹Corresponding author: subashini@uum.edu.my

Received: 29/7/2026

Revised: 1/9/2026

Accepted: 26/10/2026

Published: 28/10/2026

ABSTRACT

Foundation-level multimedia courses require students to master the five core media elements, text, graphics, audio, video, and animation, yet abstract terminology and formats are often difficult to retain through lectures alone. This study reports the design, development, and evaluation of a two-level, AI-generated quiz game built with Gemini Canvas to support teaching multimedia elements in the Multimedia System Foundation course. The game was developed using the ADDIE model and organised into five element modules, each offering an easy recall level and a hard application level, with instant explanatory feedback after every item. We used a one-group pre-test/post-test design with 80 undergraduate students, complemented by a perception survey grounded in the Technology Acceptance Model. Paired-samples analysis showed a significant improvement from pre-test to post-test, rising from a mean of 10.29 to 15.51 out of 20, $t(79) = 40.35$, $p < .001$, with a normalised gain of 0.54, indicating a medium learning effect. Gains were larger for application than for recall. Learner perceptions were highly positive across perceived usefulness, perceived ease of use, behavioural intention, and engagement, with all construct means ranging from 4.29 to 4.37 on a five-point scale. The findings indicate that an AI-generated quiz game is a well-accepted supplement associated with improved post-test performance in teaching multimedia elements, and they demonstrate a practical, low-cost workflow for educators to produce classroom games using generative AI tools.

Keywords: AI-generated game, game-based learning, Gemini Canvas, multimedia elements, technology acceptance.

INTRODUCTION

Multimedia has become a foundational literacy across computing, communication, and creative disciplines, and introductory multimedia courses commonly begin by teaching the five core elements from which every project is assembled: text, graphics, audio, video, and animation (Vaughan, 2011). In the Multimedia System Foundation course offered at a public university in Malaysia, Topic 3 introduces these elements, along with the terminology, characteristics, and file formats associated with each. Although the content is conceptually accessible, it is dense with definitions and distinctions, for example, between bitmap and vector graphics, between amplitude and frequency in audio, and among the many video container formats. Delivered only through lectures, such material is prone to shallow processing and rapid forgetting, particularly among first-year learners who may not yet connect the terms to concrete production tasks.

Game-based learning offers a well-established response to this problem (Clark et al., 2016). By embedding retrieval practice, immediate feedback and a sense of progression within an enjoyable activity, educational games can strengthen recall and transfer while sustaining attention (Plass et al., 2015; Soriano-Sánchez et al., 2026; Wouters et al., 2013). Quiz games are a natural fit for factual and conceptual content because they combine repeated testing, a documented driver of long-term retention, with the motivational affordances of scoring and levels (Hamari et al., 2016). However, creating a custom classroom game has traditionally required programming skills and considerable time, putting it out of reach for many instructors.

The emergence of generative artificial intelligence tools has begun to change this equation. Tools such as Gemini Canvas can generate functional, interactive web applications from natural-language prompts, allowing an educator without formal software-development training to produce a working game rapidly and at negligible cost (Chen & Law, 2025; Kasneci et al., 2023). Despite growing interest, few documented accounts track a single AI-generated learning game through a complete design, development, and evaluation cycle with authentic students and quantitative outcomes (Chen & Law, 2025). Recent work has begun to evaluate AI-generated quiz content and educator-built tools: Durgungoz and Durgungoz (2025) assessed the quality and engagement value of AI-generated gamified quizzes, while Bland (2026) reported an educator building a course-specific learning app with Gemini and evaluating student perception. These studies, however, either rely on template-based quiz platforms that assemble items into a fixed format, or stop at feasibility and perception without paired pre-test and post-test learning gains. The present study differs from generic quiz applications and no-code authoring tools in three ways: the game is generated end-to-end as a self-contained interactive application rather than populated into an existing template; it is deliberately structured by both content element and cognitive level (recall versus application) with item-level explanatory feedback; and it is taken through a full ADDIE cycle with measured learning outcomes as well as acceptance.

This study addresses that gap. It reports the design, development and evaluation of a two-level, AI-generated quiz game for teaching multimedia elements, produced with Gemini Canvas and deployed with 80 undergraduate students. The study is guided by three objectives: first, to design and develop an AI-generated quiz game aligned to the multimedia-elements content of this course; second, to evaluate its effect on student learning using a pre-test and post-test design; and third, to examine student perception of the game in terms of perceived usefulness, perceived ease of use, behavioural intention and engagement. By reporting both the development workflow and the measured outcomes, the study

contributes a replicable model for educators seeking to build accepted, effective games with generative AI.

LITERATURE REVIEW

Game-Based Learning and Retrieval Practice

Game-based learning uses game elements and mechanics to achieve defined learning outcomes (Soriano-Sánchez et al., 2026). Foundational work argues that well-designed games embody sound learning principles, including clear goals, incremental challenge, and frequent, actionable feedback (Gee, 2003; Prensky, 2001). Meta-analytic evidence indicates that serious games can produce moderate improvements in learning and retention relative to conventional instruction, with the strongest effects when games are used to supplement rather than replace teaching and when they incorporate feedback and repetition (Clark et al., 2016; Hamari et al., 2016; Sailer & Homner, 2020; Wouters et al., 2013). Quiz-style games also draw strength from the testing effect, where retrieving information from memory improves later retention more than restudying the same material (Soriano-Sánchez et al., 2026). Coupling retrieval with immediate corrective feedback, as a quiz game does, further consolidates learning by correcting misconceptions as they surface (Hamari et al., 2016).

Teaching Multimedia Elements

The five elements of multimedia are typically taught as definitions, characteristics, and formats (Shelly & Vermaat, 2012; Vaughan, 2011). Text is the most easily manipulated element, structured through typefaces, fonts, and attributes such as kerning and leading (Vaughan, 2011). Graphics are divided into bitmap (raster) images, composed of pixels and realistic but degraded when enlarged, and vector images, defined mathematically and therefore scalable without loss (Vaughan, 2011). Audio is characterised by amplitude, which determines volume, and frequency, which determines pitch, and by formats ranging from compressed MP3 to instruction-based MIDI (Shelly & Vermaat, 2012). Video is defined by properties such as frame rate, image size, and colour depth; broadcast standards including NTSC, PAL, and SECAM; and container formats such as MPEG-2 for DVD and MPEG-4 for high-density media (Vaughan, 2011). Animation is presented as the display of consecutive images to create movement, spanning traditional techniques such as cel animation and rotoscoping and computer techniques such as vector and three-dimensional animation (Shelly & Vermaat, 2012). Because this material is terminology-rich, it is a strong candidate for retrieval-based reinforcement.

Generative AI for Educational Content Creation

Generative artificial intelligence has rapidly become relevant to teaching practice, both as a subject and as a production tool (Chen & Law, 2025). Large language models and their multimodal extensions can now generate code and complete interactive artefacts from natural-language descriptions, lowering the barrier for non-programmers to create bespoke learning resources (Kasneci et al., 2023). Tools positioned as AI canvases, including Gemini Canvas, let an author specify an interface and its behaviour in prose and iterate conversationally until they produce a working application (Kasneci et al., 2023). For educators, this affords a practical pathway to custom, content-aligned games without a development team. What remains under-documented is the end-to-end use of such tools in a rigorous instructional cycle, including alignment to learning outcomes, systematic development, and quantitative evaluation with students.

Technology Acceptance Model

The Technology Acceptance Model (TAM) is among the most widely used frameworks for explaining user acceptance of information technology (Davis, 1989; Venkatesh & Davis, 2000). It proposes that perceived usefulness and perceived ease of use jointly shape attitudes and behavioural intention to use a system. In educational technology research, researchers commonly use TAM constructs to gauge whether learners find a tool valuable and usable and whether they intend to continue using it (Scherer et al., 2019). Because acceptance is a prerequisite for sustained classroom use, this study adopts perceived usefulness, perceived ease of use, and behavioural intention from TAM and adds engagement and enjoyment dimensions appropriate to the game context to evaluate learners' perceptions of the developed game.

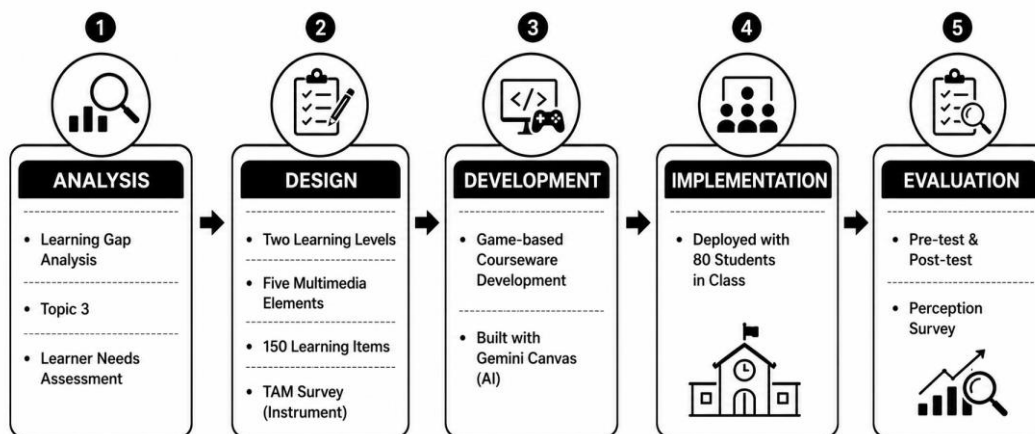
METHODOLOGY

Research Design

The study followed a design-and-development approach to build the game and a one-group pre-test/post-test quasi-experimental design to evaluate its effect on learning, complemented by a cross-sectional perception survey. The development followed the ADDIE model, comprising the phases of Analysis, Design, Development, Implementation, and Evaluation (Branch, 2009). ADDIE was selected because its evaluation phase maps directly onto the study objectives and because its systematic, phase-based structure suits the production of an instructional artefact. Figure 1 summarises how the model was applied in this study.

Figure 1

The ADDIE model as applied in this study



Analysis

The analysis phase identified the learning need and the learner profile. Topic 3 of the Multimedia System Foundation course requires students to recall and apply the concepts, terminology and formats of the five multimedia elements. Prior cohorts had found the volume of terminology difficult to retain through lectures alone, and the topic lends itself to retrieval practice. The target learners were first-year

undergraduates with mixed prior exposure to digital games. These findings motivated a quiz game structured by element and by cognitive demand, offering both recall and application practice with explanatory feedback.

Design

The design phase specified the game's structure, content, and assessment logic. Content was drawn directly from the Topic 3 lecture materials to ensure curricular alignment. The game was organised into five modules, one for each multimedia element, and each module offered two levels: an easy level targeting recall and recognition, and a hard level targeting application through short scenarios. Each level contained 15 items, for 30 items per element and 150 total. Every item was a four-option multiple-choice question with a single correct answer and a one-sentence explanation shown immediately after a response, so that the game functioned as a learning tool rather than a test. A concept refresher summarising the key facts of each element appeared before the level choice, and the system recorded a per-set score with a mastery indicator so progress could be tracked separately for each element and level.

To keep the pre-test and post-test measurement valid, the game items were written to be distinct from the test items while covering the same concepts. This design avoids inflating the post-test through rote memorisation of specific questions and instead measures whether students learned the underlying concepts.

Development

The team developed the game with Gemini Canvas, an AI tool that generates interactive web applications from natural-language prompts. The instructor specified the required structure, content, levels, feedback behaviour, and visual design in prose and iteratively refined the generated application. Before deployment, the same two subject-matter experts who validated the knowledge test reviewed all 150 AI-generated questions, along with their answer keys and explanatory feedback, for accuracy and curricular alignment. The team corrected or replaced inaccurate, ambiguous, or misaligned items, so the study used only expert-verified items. The result is a single, self-contained web application that runs in any browser without installation, making it suitable for use on students' own devices or through a projector. The interface uses a clean, light visual theme, with each of the five elements assigned a distinct colour and icon that is carried consistently through the application. Figures 2 to 6 show the game's principal interfaces.

Figure 2

The home interface, where the learner enters an optional identifier and starts the game

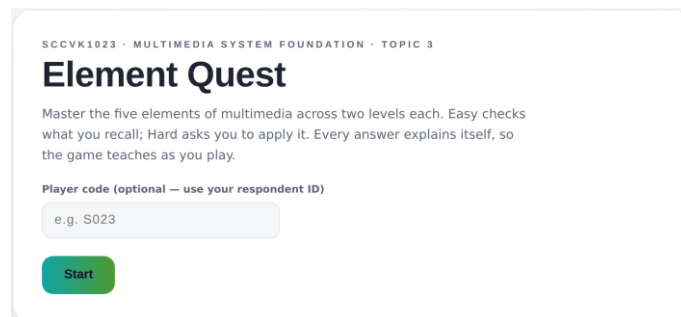


Figure 3

The element-selection interface, showing the five multimedia elements and per-set progress

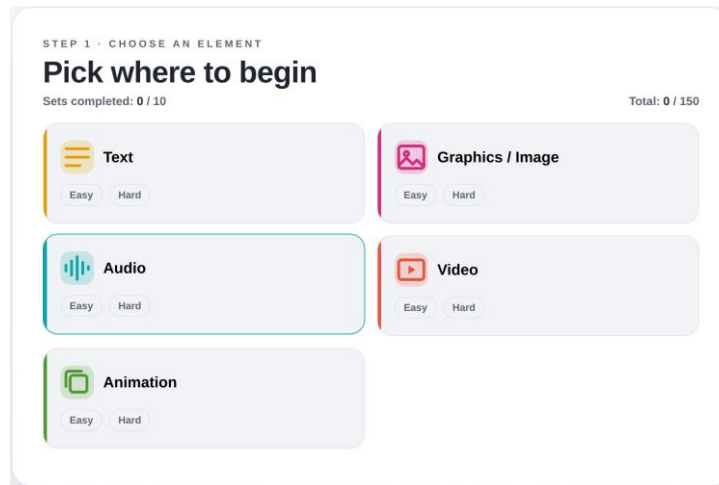


Figure 4

The level-selection interface, with a concept refresher and a choice of easy or hard level

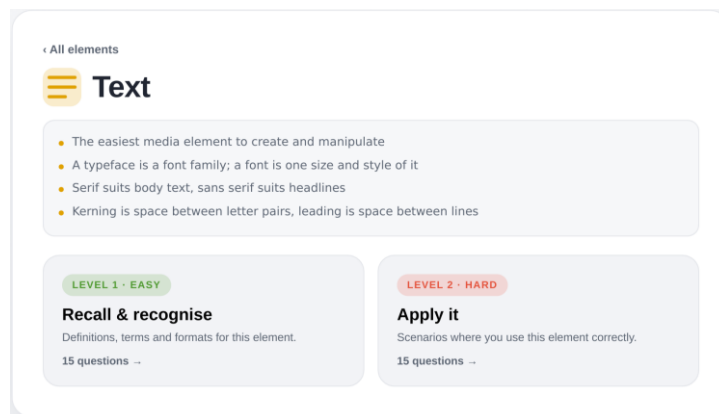


Figure 5

The gameplay interface, showing a question with instant explanatory feedback after a response

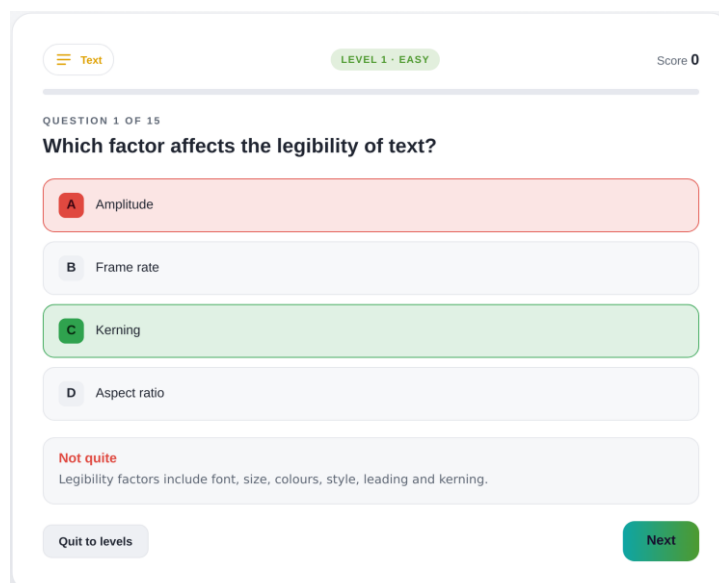
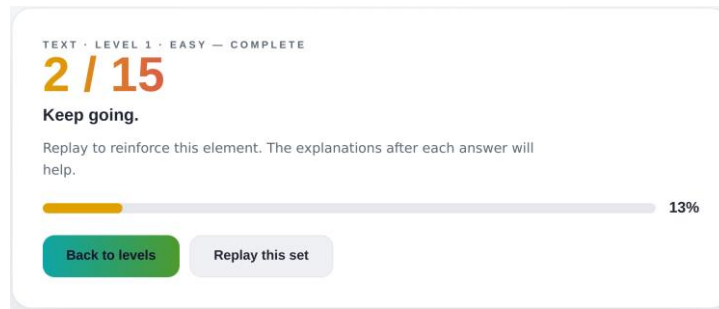


Figure 6

The results interface, reporting the score and a performance message for the completed set



Implementation

We implemented the developed game with a full cohort of 80 undergraduate students enrolled in the course during a regular Topic 3 session. Students first completed the pre-test, then engaged with the game across its levels, and finally completed the post-test using the same instrument, followed by the perception survey. The game served as a reinforcement activity after teaching each multimedia element, so students played the relevant element module after its topic was delivered rather than in a single continuous sitting. Students could attempt both the easy and hard levels of each element and were not required to complete all 150 items; the game recorded a per-set score, allowing each learner to revisit any element and level for further practice. The activity formed part of normal teaching and learning, and participation was voluntary and confidential.

Participants

The participants were the 80 students enrolled in the course. The sample comprised 28 male and 52 female students aged 19 to 22 years, with a mean age of 20.5 years. Most reported playing digital games often or very often, and 72 of 80 students had previously used an educational game. Table 1 summarises the demographic profile of the participants. We used purposive sampling to select the intact class and obtained informed consent from all participants before data collection.

Table 1

Demographic Profile of Participants (N = 80)

Characteristic	Category	N	Per cent
Gender	Male	28	35.0
	Female	52	65.0
Frequency of playing games	Very often	20	25.0
	Often	41	51.2
	Sometimes	18	22.5
	Rarely	1	1.3
Prior use of educational game	Yes	72	90.0
	No	8	10.0

Note. Mean age = 20.5 years (*SD* = 1.13).

Instruments

Knowledge Test

We measured learning with a 20-item multiple-choice knowledge test on the multimedia elements, administered unchanged as both the pre-test and post-test. The test had two parts that mirrored the game: 10 recall items (Level 1) and 10 application items (Level 2), each worth 1 mark, for a maximum of 20. This structure allowed reporting of overall learning alongside separate recall and application sub-scores. Two subject-matter experts validated the knowledge test for content against the Topic 3 learning outcomes, and it was pilot-tested before use.

Perception Questionnaire

Learner perception was measured with a 17-item questionnaire using a five-point Likert scale from strongly disagree to strongly agree. The instrument comprised four constructs: perceived usefulness (five items), perceived ease of use (five items) and behavioural intention (three items), adopted from the Technology Acceptance Model, together with an engagement and enjoyment dimension (four items) appropriate to the game context. The perceived usefulness, perceived ease of use, and behavioural intention items were adapted from Davis (1989) and Venkatesh and Davis (2000), and the engagement items were developed for this study; two experts reviewed the instrument, and its internal consistency was examined (see Table 3).

Data Analysis

We analysed the data using descriptive and inferential statistics. We compared pre-test and post-test scores using a paired-samples *t*-test, reported the effect size as Cohen's *d* (computed for the paired design as the mean pre-to-post difference divided by the standard deviation of the difference scores), and reported the class-level learning effect as Hake's normalised gain (Hake, 1998). We examined the normality of the difference scores with the Shapiro-Wilk test and used the Wilcoxon signed-rank test as a non-parametric confirmation. Perception data were summarised as construct means and standard deviations, and Cronbach's alpha was used to assess the internal consistency of each construct (Cronbach, 1951). All statistics were rounded to two decimal places.

RESULTS

Effect on Learning

The knowledge test showed a marked improvement from pre-test to post-test. The mean score rose from 10.29 (*SD* = 2.07) to 15.51 (*SD* = 2.46) out of 20, a mean gain of 5.22 marks. A paired-samples *t*-test confirmed this improvement was statistically significant, $t(79) = 40.35$, $p < .001$, with a very large effect size by Cohen's conventions ($d > 0.8$), reflecting the tightly clustered gain scores typical of a single-group design (Cohen's $d = 4.51$). Because the Shapiro-Wilk test indicated that the difference scores departed from normality ($W = 0.93$, $p < .001$), we also conducted a Wilcoxon signed-rank test, which confirmed the same significant improvement ($p < .001$). Hake's normalised gain was 0.54, representing a medium learning effect and indicating that the game recovered slightly more than half of the improvement available to the class.

Separating recall and application items showed that both kinds of knowledge improved, and that application showed the larger absolute gain. Recall (Level 1) rose from 6.06 ($SD = 1.40$) to 8.26 ($SD = 1.38$), a gain of 2.20 marks, $t(79) = 30.54$, $p < .001$, $d = 3.41$. Application (Level 2) rose from 4.22 ($SD = 1.03$) to 7.25 ($SD = 1.36$), a gain of 3.02 marks, $t(79) = 34.03$, $p < .001$, $d = 3.80$. Table 2 reports these results, and Figure 7 illustrates the pre-test and post-test means.

Table 2

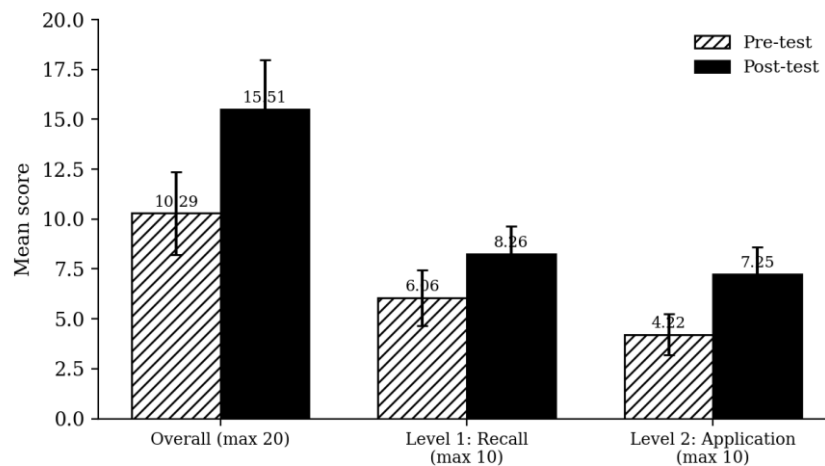
Pre-Test and Post-Test Descriptive Statistics and Paired-Samples t-Test

Measure	Pre M (SD)	Post M (SD)	Gain	<i>t</i>	<i>d</i>
Overall (max 20)	10.29 (2.07)	15.51 (2.46)	5.22	40.35	4.51
Level 1: Recall (max 10)	6.06 (1.40)	8.26 (1.38)	2.20	30.54	3.41
Level 2: Application (max 10)	4.22 (1.03)	7.25 (1.36)	3.02	34.03	3.80

Note. $N = 80$; $df = 79$ for all tests; $p < .001$ for all comparisons. d = Cohen's d . Normalised gain (Hake) for the overall test = 0.54.

Figure 7

Mean pre-test and post-test scores overall and by level, with standard-deviation bars



Learner Perception

Learner perception of the game was highly positive across all four constructs. On the five-point scale, perceived usefulness averaged 4.29 ($SD = 0.26$), perceived ease of use 4.37 ($SD = 0.28$), behavioural intention 4.32 ($SD = 0.34$), and engagement and enjoyment 4.35 ($SD = 0.32$). All construct means exceeded 4.29, well above the high-agreement threshold of 3.67 (the conventional upper-third cut-off on a five-point Likert scale, corresponding to the boundary of the highest of three equal 1.33-point interpretive bands), indicating that students found the game useful and easy to use, intended to use similar games in future, and enjoyed the experience. Table 3 reports the construct statistics, and Figure 8 presents them graphically.

The overall instrument showed acceptable internal consistency, with a Cronbach's alpha of 0.69 across the 17 items. The alpha values for the individual constructs were lower, ranging from 0.39 to 0.46. This is attributable to a ceiling effect, since uniformly high and closely clustered ratings restricted the variance within each short construct, a condition known to deflate Cronbach's alpha. Because

Cronbach's alpha is a function of inter-item covariance, it falls when responses are compressed near the top of the scale even if the items are internally consistent, and this deflation is amplified for constructs measured with only three to five items, as here. The acceptable overall alpha of 0.69 across the 17 items, obtained after pooling variance across the full instrument, supports this interpretation rather than genuinely unreliable items. This pattern is discussed further as a study limitation.

Table 3

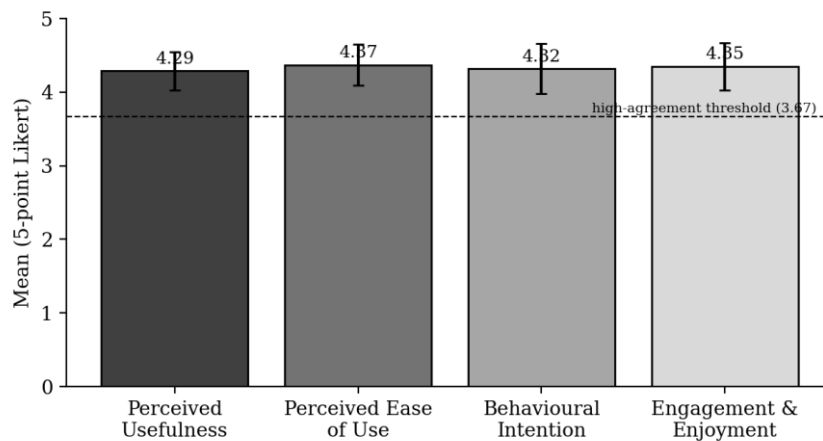
Perception Construct Means, Standard Deviations and Reliability

Construct	Items	<i>M</i>	SD	α
Perceived usefulness	5	4.29	0.26	0.39
Perceived ease of use	5	4.37	0.28	0.43
Behavioural intention	3	4.32	0.34	0.45
Engagement and enjoyment	4	4.35	0.32	0.46
Overall	17	4.33	0.20	0.69

Note. Ratings on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). α = Cronbach's alpha.

Figure 8

Mean perception scores by construct on a five-point scale, with standard-deviation bars



DISCUSSION

The results indicate that the AI-generated quiz game was associated with improved learning of multimedia elements. The significant pre-test to post-test improvement, with a very large effect size and a medium normalised gain, suggests that using the game was associated with substantial gains in the target content. Because the study used a single group without a control condition, we cannot attribute this improvement to the game alone. This aligns with broader evidence that games combining retrieval practice with immediate feedback support retention, and that such games are most effective when they supplement teaching rather than replace it (Sailer & Homner, 2020; Wouters et al., 2013).

A notable pattern is that absolute gains were larger for application than for recall. Because application items required students to select the correct element or format for a described scenario, this suggests the game supported not only memorising terminology but also transferring it to realistic decisions. Separating an easy recall level from a harder application level, and providing an explanation after every response, plausibly contributed to this outcome by giving students repeated, feedback-rich practice in applying concepts.

The perception results show that the game was well accepted. High scores on perceived usefulness and perceived ease of use, together with a strong behavioural intention to use similar games in future, indicate that acceptance is unlikely to be a barrier to adoption, which is important because acceptance is a precondition for sustained classroom use (Davis, 1989). The equally high engagement scores suggest the game was enjoyable, an affective condition that supports continued participation.

Beyond the learning and perception outcomes, the study demonstrates a practical development workflow. The complete game, spanning five elements and 150 items with instant feedback and progress tracking, was produced with Gemini Canvas from natural-language prompts, without a development team and at negligible cost. This illustrates how generative AI tools can let subject-matter experts create content-aligned learning games directly and offers a replicable model for other educators (Kasneci et al., 2023). From a digital system development perspective, the study contributes a documented, reproducible prompt-to-application workflow that translates a natural-language specification into a single, self-contained, browser-based interactive system with modular content, branching difficulty levels, item-level feedback logic, and per-set score tracking. This positions generative AI canvases as a rapid development environment for educational software and shows that educators can produce and validate a functional, deployable learning system without a conventional software-engineering pipeline, consistent with recent educator-driven development accounts (Bland, 2026). This effectiveness is consistent with meta-analytic evidence that game-based learning improves achievement (Clark et al., 2016; Soriano-Sánchez et al., 2026), with the engagement and flow reported for well-designed learning games (Hamari et al., 2016), and with emerging work that combines artificial intelligence and game-based learning (Chen & Law, 2025).

CONCLUSION

This study designed, developed and evaluated a two-level, AI-generated quiz game for teaching the five elements of multimedia, produced with Gemini Canvas and deployed with 80 undergraduate students. Using a pre-test/post-test design, the game was associated with significant improvement in knowledge, with larger gains in application than in recall, and learner perceptions were highly positive across usefulness, ease of use, behavioural intention, and engagement. Together, the findings indicate that an AI-generated quiz game is a well-accepted supplement associated with improved learning outcomes for teaching multimedia elements, though the single-group design limits causal claims.

Several limitations should be acknowledged. The study used a one-group pre-test and post-test design without a control group, so the gains cannot be attributed to the game alone with full certainty, and a testing effect from the repeated instrument cannot be excluded. The perception ratings showed a ceiling effect, which restricted variance and deflated the internal-consistency estimates for the individual constructs; future work should use more discriminating items and more heterogeneous samples. The study also measured learning immediately after the intervention, so retention over time was not assessed. Future research should include a control group, measure delayed retention, and extend the approach to other topics and courses, while further examining how AI-generated game design influences learning and acceptance.

ACKNOWLEDGMENT

The author thanks the students of the Multimedia System Foundation course at Universiti Utara Malaysia for their participation in this study. This study received no external funding.

REFERENCES

- Bland, T. (2026). Harnessing “vibe coding” to rapidly develop tailored educational apps: A generative AI-driven ECG interpretation tool in medical education. *AI*, 7(6), Article 223. <https://doi.org/10.3390/ai7060223>
- Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer. <https://doi.org/10.1007/978-0-387-09506-6>
- Chen, C. H., & Law, V. (2025). The role of help-seeking from ChatGPT in digital game-based learning. *Educational Technology Research and Development*, 73, 1703–1721. <https://doi.org/10.1007/s11423-025-10464-3>
- Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, 86(1), 79–122. <https://doi.org/10.3102/0034654315582065>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Durgungoz, A., & Durgungoz, F. C. (2025). Exploring effortless AI-generated gamified quizzes in an online special education module: Evaluating question quality, student engagement, and its potential to identify at-risk students. *Education and Information Technologies*, 30(17), 25335–25357. <https://doi.org/10.1007/s10639-025-13765-5>
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave Macmillan.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow, and immersion in game-based learning. *Computers in Human Behavior*, 54, 170–179. <https://doi.org/10.1016/j.chb.2015.07.045>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, 50(4), 258–283. <https://doi.org/10.1080/00461520.2015.1122533>
- Prensky, M. (2001). *Digital game-based learning*. McGraw-Hill.
- Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review*, 32(1), 77–112. <https://doi.org/10.1007/s10648-019-09498-w>
- Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13–35. <https://doi.org/10.1016/j.compedu.2018.09.009>
- Shelly, G. B., & Vermaat, M. E. (2012). *Discovering computers: Fundamentals*. Cengage Learning.
- Soriano-Sánchez, J. G., Quijano López, R., & Airado Rodríguez, D. (2026). The impact of game-based learning on motivation, self-efficacy, and academic achievement in the natural sciences: A meta-analysis. *Education Sciences*, 16(1), Article 122. <https://doi.org/10.3390/educsci16010122>

- Vaughan, T. (2011). *Multimedia: Making it work* (8th ed.). McGraw-Hill.
- Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), 186–204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249–265. <https://doi.org/10.1037/a0031311>