Ozzy and the Oddity Kingdom: A Generative AI Visual Novel Educational Mobile Game

[1]Nur Ameera Husna Binti Mahfizan, [2]Nurateela Atasya Binti Roziky Effendy, [3]Siti Nur Aisyah Binti Ruddy, [4]Sofea Shameera Binti Azrul Amri, *[5]Muhammad Haziq Bin Abd Rashid

[12345]Akademi Pengajian Bahasa, Universiti Teknologi MARA Shah Alam

***Corresponding Author**

DOI: https://dx.doi.org/10.47772/IJRISS.2025.925ILEIID000034

Received: 23 September 2025; Accepted: 30 September 2025; Published: 05 November 2025

ABSTRACT

Students majoring in English for Professional Communication, enrolled in the Introduction to Critical Thinking course, often struggle to grasp abstract concepts such as barriers to critical thinking, inductive and deductive reasoning, and the identification of logical fallacies. To address this gap, Ozzy and the Oddity Kingdom was developed as a Mobile-Assisted Language Learning (MALL) solution that transforms these challenging ideas into an engaging, interactive experience. Built in Unity using C#, the game adopts a visual novel format set in a rich fantasy world, enhanced by high-quality, AI-generated resources created through Midjourney. The design incorporates gamification elements such as branching storylines, distinctive characters, a map-based navigation interface, and a variety of formative and summative assessments including drag-and-drop, fill-in-the-blank, multiple-choice, and word scrabble challenges. A reward-and-punishment system and scalable difficulty levels encourage sustained engagement while reinforcing learning objectives. The choice of Unity ensures cross-platform deployment on both iOS and Android with minimal production costs, making the project both practical and sustainable. Beyond its immediate application in critical thinking education, the framework is adaptable to other subject areas, opening pathways for wider adoption. Its market potential lies in its ability to serve as a supplementary learning tool that blends academic rigour with narrative-driven play, offering educators and learners an accessible, portable, and customisable resource for deepening understanding in diverse educational contexts.

Keywords: mobile-assisted language learning, unity, generative ai, critical thinking, visual novel)

INTRODUCTION

Students in Introduction to Critical Thinking often struggle with abstract concepts such as barriers to sound judgment, the difference between inductive and deductive reasoning, and the classification of fallacies. These ideas require students to apply knowledge beyond textbook examples to real-life situations, which is challenging. A mobile-assisted approach offers a practical solution. Research shows that mobile-supported learning can lead to moderate to high improvements in engagement and achievement compared to traditional methods (Chen et al., 2020). However, reviews also warn against focusing only on technology without strong pedagogy (Burston, 2014). For critical thinking in EFL courses, studies recommend explicit teaching, guided questioning, and collaborative tasks, which work best in interactive and feedback-rich environments (Zhao et al., 2016). Based on this, Ozzy and the Oddity Kingdom was developed as a mobile game that applies these principles in an engaging format.

Our design uses a visual novel with branching storylines, supported by evidence that narrative-based environments can improve both engagement and learning (Rowe et al., 2010). Branching choices create decision-making opportunities that build perspective-taking and reasoning skills when supported by clear scaffolds (Fenici & Mosca, 2024). Visual novels also allow for proven teaching strategies such as choice-based dialogue, mini-games, and exploration, which we use to practice reasoning skills (Camingue et al., 2020). Gamification elements like progression and rewards are included but aligned with learning goals, as research shows that unstructured gamification has limited impact (Sailer & Homner, 2020; Huang et al., 2020).

Generative AI is used to create diverse content while maintaining teacher control. AI helps produce prompts,

examples, and assessment items quickly, which is valuable for mobile learning (Godwin-Jones, 2024). At the same time, ethical concerns such as privacy, bias, and overreliance require clear policies and human review (Williams, 2024). Ozzy and the Oddity Kingdom uses AI-generated art and text as drafts, refined by experts to ensure quality and alignment with learning goals. Combined with formative checks embedded in the story and mini-games, this approach makes abstract reasoning visible and improves it through timely feedback and repeated low-stakes practice (Sailer & Homner, 2020).

Problem Statement

Traditional lectures and worksheets focus on memorising rules rather than applying them, which limits transfer to authentic contexts. Research on mobile-assisted language learning shows moderate to large improvements compared to traditional methods, but reviews also note inconsistent study quality and an overuse of surface-level tasks (Chen et al., 2020; Mihaylova et al., 2022). In EFL courses, explicit modelling and guided questioning improve critical thinking, but these methods require continuous practice and feedback that static materials cannot provide (Zhao et al., 2016). There is currently no solution that embeds reasoning practice into meaningful interaction on a device students use daily (Burston, 2014).

Another challenge is design coherence. Gamification is often used to keep students engaged, but research shows it only brings small to moderate learning gains unless game elements are closely tied to learning goals (Huang et al., 2020; Sailer & Homner, 2020). Narrative-based learning offers a better approach because well-designed stories can connect motivation with problem solving. Studies show that narrative engagement can improve learning outcomes (Rowe et al., 2010). Visual novels, in particular, allow for strategies such as choice-based dialogue, mini-games, and exploration, which can be structured to build reasoning skills step by step (Camingue et al., 2020). Current tools for this course do not use these strategies, leaving a gap between research and practice.

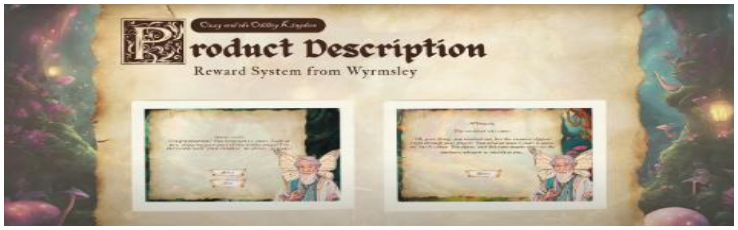
Assessment is another weakness. Current evaluations rely on multiple-choice questions that measure recognition rather than reasoning in action. Research on game-based assessment shows that embedded and unobtrusive measures can track skills during play and provide feedback that improves performance (Shute & Ventura, 2013; Shute et al., 2021). However, creating and maintaining such systems is expensive, so many mobile apps fail to deliver strong formative assessment. Generative AI can speed up content creation and item variation, but its use raises concerns about privacy, bias, and overreliance, which require human oversight and clear policies (Godwin-Jones, 2024; Williams, 2024). At present, there is no mobile solution that combines narrative-based design, structured gamification, and embedded assessment while meeting ethical standards for AI use.

Objectives

1. To design and develop a story-based mobile application that uses gamification and built-in formative assessment to help learners identify and apply critical thinking skills in real-life contexts.
2. To assess the effectiveness and commercial potential of a generative AI-supported content creation process for producing scalable and ethically sound educational resources within a mobile-assisted language learning framework.

PRODUCT DESCRIPTION & METHODOLOGY

Figure 1 shows One of The Available Features in The Application



Ozzy and the Oddity Kingdom is a cross-platform visual novel game built in Unity using C#. It places key critical thinking concepts within story-based choices that affect the game's outcome. Each chapter focuses on a specific skill, such as distinguishing inductive from deductive reasoning or identifying fallacies. Players move through branching storylines where their decisions influence the plot. Short mini-games are included, such as drag-and-drop argument mapping, fill-in-the-blank exercises, timed multiple-choice questions, and word scrabble for key terms. The game uses progressive difficulty, low-stakes retries, and rewards based on reasoning quality rather than speed. This design follows research showing that narrative-based environments can improve engagement and learning, and that visual novels support strategies like choice-based dialogue and mini-games (Camingue et al., 2020; Rowe et al., 2010). Gamification is used carefully to support learning goals, as studies show limited benefits when game elements are not aligned with instruction (Huang et al., 2020; Sailer & Homner, 2020). Branching choices are designed to develop perspective-taking and reasoning, not just to vary content (Fenici & Mosca, 2024).

Generative Content Workflow

Content creation uses a human-in-the-loop process. Generative AI produces draft artwork, dialogue, and question items, which are then reviewed for accuracy, cultural sensitivity, and fairness. This speeds up development while maintaining quality. Research shows that AI can support both independent and teacher-led learning by generating examples and feedback, but strict controls are needed to manage bias and privacy (Godwin-Jones, 2024). Ethical measures include local processing where possible, clear consent notices, and a review checklist for fairness and difficulty. This checklist ensures diverse representation, avoids stereotypes, and tracks all revisions. These steps follow current recommendations for responsible AI use in education, which stress transparency and human oversight (Williams, 2024).

Assessment Design and Learning Supports

Figure 2 Shows the Types of Assessments in the Application

The game uses embedded assessments that track reasoning through player actions. Each skill is linked to observable behaviours, such as selecting counterexamples or repairing arguments. Data from choices and mini-games feed into a stealth assessment system that updates learner profiles in real time and provides targeted feedback, such as worked examples or guided retries (Shute & Ventura, 2013; Shute et al., 2021). Tutorials include modelling and guided questioning before independent practice, reflecting strategies proven effective in

EFL contexts (Zhao et al., 2016). The aim is to integrate assessment into gameplay, reducing test anxiety and supporting continuous improvement while maintaining reliability for research (Shute & Ventura, 2013).

Empirical Methodology

In the scheduled pilot study, the evaluation is planned to use a two-phase design-based research approach. Phase one involves classroom pilots to test usability, narrative flow, and assessment accuracy through think-aloud sessions and log analysis. Phase two is a quasi-experimental study comparing the game with standard teaching, using pre- and post-tests on argument analysis and transfer tasks. Engagement and user experience will be measured through surveys and interviews. This approach addresses common weaknesses in MALL research, such as short duration and small samples, while building on evidence of mobile learning benefits and the added value of narrative and gamification (Chen et al., 2020; Mihaylova et al., 2022; Rowe et al., 2010; Sailer & Homner, 2020). Instructor workload and implementation fidelity will also be documented to inform future scaling.

POTENTIAL FINDINGS AND COMMERCIALISATION

The game is expected to improve students' ability to analyse arguments and apply critical thinking skills compared to the usual teaching approach. Based on previous studies, mobile learning often shows better results than traditional methods, and narrative-based games can increase both engagement and learning (Chen et al., 2020; Rowe et al., 2010). The built-in assessment system should match well with external tests, showing that it measures reasoning accurately while giving feedback during play (Shute & Ventura, 2013; Shute et al., 2021). Students are also likely to make fewer mistakes with fallacies and better separate inductive and deductive reasoning because the game combines clear examples, guided practice, and independent tasks, which are effective in EFL settings (Zhao et al., 2016).

Game data should show that quick feedback after wrong choices helps students improve immediately and over time. This supports the idea that assessment should guide learning without interrupting the experience (Shute & Ventura, 2013). Engagement scores are expected to partly explain learning gains, which agrees with research showing that strong stories and interactive choices keep students motivated and focused on learning (Rowe et al., 2010; Camingue et al., 2020). Fairness checks should confirm that the game works for students with different language levels, which is important for first-year courses with mixed abilities (Chen et al., 2020).

The content creation process using generative AI should speed up development without lowering quality when human review is included. More scenarios can be produced in less time, while checks ensure accuracy and cultural sensitivity. This matches findings that AI can help create language learning materials if teachers control quality and protect privacy (Godwin-Jones, 2024). Interviews may show concerns about ethics and overuse of AI, so clear policies and consent steps will be important (Williams, 2024). If successful, this process will make it easier to update and localise content for different contexts.

Figure 2 shows The List of Commercialisation Potentials of the Application

For commercialisation, the plan is to offer the app to universities first, with features like progress dashboards for instructors and easy cross-platform access. Later, a freemium model can reach a wider audience, with basic chapters free and advanced content as paid add-ons. The AI-supported pipeline will allow fast updates and

seasonal content while keeping costs low and quality high (Godwin-Jones, 2024; Williams, 2024). Evidence from trials, including cost per learner and teacher workload, will support adoption by institutions looking for affordable, research-based tools (Mihaylova et al., 2022; Sailer & Homner, 2020)

NOVELTY AND RECOMMENDATIONS

This project is unique because it combines a story-based visual novel with built-in assessment for critical thinking. The branching choices are designed to make students practice reasoning skills like spotting fallacies and distinguishing between inductive and deductive arguments. Feedback is given during the story so learning happens without breaking the flow. Research shows that strong narratives can improve both engagement and learning, and visual novels work well when they include choices and mini-games (Rowe et al., 2010; Camingue et al., 2020). Gamification is used carefully to support learning goals, which is important because studies show that random game elements have little effect on learning (Huang et al., 2020; Sailer & Homner, 2020). The game also uses stealth assessment to track skills and give timely feedback, following best practices in game-based learning (Shute & Ventura, 2013).

The project uses a human-in-the-loop process with generative AI to create dialogue, scenarios, and assessment items faster while keeping quality and cultural accuracy. This approach combines the speed of AI with teacher review to avoid bias and errors. It also addresses common problems in mobile learning projects, such as short lifespan and poor scalability, by making it easier to update and localize content (Godwin-Jones, 2024; Williams, 2024). The design meets both main goals: improving critical thinking through interactive learning and testing an ethical, scalable way to produce educational content.

Teachers should use the game as part of a critical thinking unit, starting with short lessons and then guided play. Instructors need training on how to read the dashboard, adjust difficulty, and run quick discussions after gameplay. Researchers should measure both learning outcomes and engagement, and check if feedback inside the game improves results. Larger studies across different campuses are recommended to confirm effectiveness and fairness (Mihaylova et al., 2022; Zhao et al., 2016). Assessment accuracy should also be tested against external tasks to ensure it measures real reasoning skills (Shute & Ventura, 2013).

Start with university licenses that include teacher dashboards and clear privacy policies. Later, use a freemium model with free basic chapters and paid advanced content. The AI content pipeline should include templates, bias checks, and localization guides to keep updates fast and ethical (Godwin-Jones, 2024; Williams, 2024). Partnerships with language centers can help adapt content for EFL learners. Finally, share data on cost, teacher workload, and learning gains to support adoption. Evidence from mobile learning and gamification research can strengthen the case for investment (Chen et al., 2020; Sailer & Homner, 2020).

ACKNOWLEDGEMENTS

The authors would like to thank Universiti Teknologi MARA for providing academic guidance and resources that supported the development of this project. Appreciation is also extended to the course instructors and peers from the Introduction to Critical Thinking class for their constructive feedback during the design and testing phases.

REFERENCES

1. Burston, J. (2014). MALL: The pedagogical challenges. Computer Assisted Language Learning, 27(4), 344–357. https://doi.org/10.1080/09588221.2014.914539
2. Camingue, J., Melcer, E. F., & Carstensdottir, E. (2020, September). A visual novel route to learning: A taxonomy of teaching strategies in visual novels. In Proceedings of the International Conference on the Foundations of Digital Games (FDG '20). https://doi.org/10.1145/3402942.3403004
3. Chen, Z., Chen, W., Jia, J., & An, H. (2020). The effects of using mobile devices on language learning: A meta-analysis. Educational Technology Research and Development, 68(4), 1769–1789. https://doi.org/10.1007/s11423-020-09801-5

4. Fenici, M., & Mosca, I. (2024). Gamebooks and branching narratives in education: Fostering sustainability competences to promote positive social change. Frontiers in Education, 8 (2023), Article 1335605. https://doi.org/10.3389/feduc.2023.1335605
5. Godwin-Jones, R. (2024). Distributed agency in second language learning and teaching through generative AI. Language Learning & Technology, 28(2), 5–30. https://doi.org/10.64152/10125/73570
6. Huang, R., Ritzhaupt, A. D., Sommer, M., Zhu, J., Stephen, A., Valle, N., Hampton, J., & Li, J. (2020). The impact of gamification in educational settings on student learning outcomes: A meta-analysis. Educational Technology Research and Development, 68(4), 1875–1901. https://doi.org/10.1007/s11423-020-09807-z
7. Mihaylova, M., Gorin, S., Reber, T. P., & Rothen, N. (2022). A meta-analysis on mobile-assisted language learning applications: Benefits and risks. Psychologica Belgica, 62(1), 252–271. https://doi.org/10.5334/pb.1146
8. Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C. (2010). Integrating learning and engagement in narrative-centered learning environments. In Proceedings of the 10th International Conference on Intelligent Tutoring Systems (pp. 166–177). https://intellimedia.ncsu.edu/wp-content/uploads/sites/42/rowe-its-2010.pdf
9. Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. Educational Psychology Review, 32, 77–112. https://doi.org/10.1007/s10648-019-09498-w
10. Shute, V. J., Lu, X., & Rahimi, S. (2021). Stealth assessment. In J. M. Spector (Ed.), The Routledge Encyclopedia of Education (pp. 1–9). Routledge. https://files.eric.ed.gov/fulltext/ED612156.pdf
11. Williams, R. T. (2024). The ethical implications of using generative chatbots in higher education. Frontiers in Education, 8 (2023), Article 1331607. https://doi.org/10.3389/feduc.2023.1331607
12. Zhao, C., Pandian, A., & Singh, M. K. M. (2016). Instructional strategies for developing critical thinking in EFL classrooms. English Language Teaching, 9(10), 14–21. https://doi.org/10.5539/elt.v9n10p14