

A COMPARATIVE ANALYSIS OF AI MODELS FOR SCENARIO-BASED LEARNING

**Gabriel Borg, Keith Azzopardi, Matthew Montebello,
Vanessa Camilleri, & Luke Cardona**
Department of Artificial Intelligence, University of Malta (Malta)

Abstract

Scenario-Based Learning (SBL) is a pedagogical approach that situates learners within realistic narratives to promote skill transfer, decision-making, and contextual understanding (Clark, 2012). With the rise of generative artificial intelligence, Large Language Models (LLMs) have become central to enhancing SBL by enabling adaptive dialogue, dynamic narrative generation, and personalized feedback (Liu, Yang, & Xu, 2023; Zhang, Wang, & Li, 2024). This paper presents a comparative analysis of LLM applications in SBL across immersive (VR/digital) and non-immersive (text-based and 2D) environments, focusing on education and safety training domains. The study incorporates a review of SBL literature and evaluation frameworks, highlighting how learning outcomes, engagement, and realism are assessed in current research (Lee, Park, & Kim, 2024; Kim & Torres, 2023). We also examine design processes underpinning SBL systems, including human-centered design, iterative prototyping, and collaborative scenario development with educators (WHO, 2022; Ahmed, Brooks, & Liu, 2023; Hu & Chan, 2025). Findings from recent studies (Liu et al., 2023; Rajput, Seetharam, & Balasubramanian, 2019) indicate that LLM-driven systems support high levels of psychological fidelity and learner engagement, particularly in dialogue-rich contexts, though challenges such as prompt control and response reliability persist (Zhang et al., 2024). Finally, through this paper we aim to conclude that the most effective SBL implementations are built on hybrid AI architectures that combine the creativity and adaptability of LLMs with the precision and structure of rule-based or knowledge-driven components. Such hybrid designs enable educators to maintain control over content accuracy while benefiting from AI-driven flexibility. Additionally, robust evaluation frameworks that systematically assess learning outcomes, engagement, and realism are crucial to ensure both educational effectiveness and technical reliability. Together, these strategies provide a pathway toward developing realistic, adaptive, and safe learning environments powered by AI (Lee et al., 2024; Kim & Torres, 2023; Ahmed et al., 2023).

Keywords: *Scenario-Based Learning (SBL), Large Language Models (LLMs), virtual reality, educational technology, adaptive learning, human-centered design.*

1. Introduction

Scenario-Based Learning (SBL) has emerged as a powerful pedagogical approach that situates learners within authentic, context-rich narratives to promote skill transfer, decision-making, and deeper conceptual understanding (Clark, 2012). Traditionally, SBL environments relied on carefully designed branching scenarios or scripted interactions that provided realistic contexts for learners to practice complex tasks. With recent advances in generative artificial intelligence, particularly Large Language Models (LLMs), new opportunities have emerged to make SBL systems more adaptive, conversational, and personalized. Studies such as those by Liu, Yang, & Xu (2023) and Zhang, Wang, & Li (2024) demonstrate that LLMs can simulate dynamic dialogue, generate context-aware narratives, and respond intelligently to learner inputs in both virtual reality and text-based simulations, marking a notable shift from static scenario structures to interactive, learner-driven experiences.

Despite these advances, comparative evidence remains limited on how different LLMs influence key educational constructs such as psychological fidelity, learner engagement, and measurable performance outcomes across diverse SBL contexts. In parallel, prior research on SBL system design underscores the importance of incorporating human-centered and iterative development practices (Hu & Chan, 2025; Rajput, Seetharam, & Balasubramanian, 2019). However, there is currently a lack of comprehensive synthesis on how these design methodologies are adapted, or need to be adapted, when integrating

generative AI components into scenario workflows. Moreover, while existing evaluation models emphasize engagement, knowledge retention, and technical fidelity (Ahmed, Brooks, & Liu, 2023; World Health Organization, 2022), few frameworks capture the nuanced dimensions introduced by LLM-based adaptivity, such as conversational coherence, ethical interaction patterns, and learner agency.

Addressing these gaps is essential for understanding how LLM-enhanced SBL systems perform, how they are designed, and how their impacts can be rigorously evaluated. To guide this investigation, our study is structured around the following research questions:

1.1. Research questions

1. How do current Large Language Models (LLMs) contribute to enhancing realism, adaptability, and learner engagement within Scenario-Based Learning (SBL) environments?
2. What design processes and pedagogical strategies have researchers employed to integrate LLMs effectively into SBL systems across educational and safety training domains?
3. Which frameworks and evaluation criteria are most effective for assessing the performance and learning impact of LLM-driven SBL systems?

The structure of our research is organized to explore how AI Models (LLMs) support Scenario-Based Learning (SBL). In Section 2, we review relevant literature on SBL pedagogy, early AI integrations, recent LLM applications, design approaches, and evaluation frameworks, identifying current gaps in comparative evidence. Section 3 outlines the methodology used for the comparative analysis, including study selection criteria, analytic dimensions, and synthesis procedures. In Section 4 we present our results and summarise how different AI models perform across immersive and non-immersive environments, assessing their effects on engagement, realism, adaptability, and learning outcomes. We then discuss these findings relative to the research questions in Section 5, highlighting design practices, evaluation trends, and implications for educational and safety training domains. Finally, we conclude in Section 6 by summarizing key insights and emphasizing directions for future work on hybrid architectures and comprehensive evaluation frameworks.

2. Background

2.1. Scenario-Based Learning (SBL) theoretical foundations

Scenario-Based Learning (SBL) is grounded in constructivist and experiential learning theories that emphasize authentic task engagement, situated decision-making, and reflective practice. Classic conceptualizations of SBL argue that learning is optimized when learners interact with realistic scenarios that require active problem solving and contextual application of skills, rather than passive reception of information (Clark, 2012). Contemporary educational technology research further situates SBL within broader theoretical frameworks such as adaptive learning and learner autonomy, demonstrating how narrative complexity, feedback loops, and cognitive engagement support higher-order learning outcomes (Liu et al., 2025). Together, these theoretical positions provide the basis for understanding how SBL environments can mediate cognitive, affective, and behavioral dimensions of learning, and why extensions, such as the integration of generative AI, must preserve alignment with pedagogical principles to maintain instructional integrity.

2.2. AI in SBL: From early tutoring to immersive LLM-Driven environments

AI in education has evolved from early intelligent tutoring systems (ITS) with rule-based adaptive feedback to modern LLM-enhanced environments. Traditional ITS frameworks (Tan, L. et al., 2025) supported learner engagement and procedural performance by diagnosing learner states and providing tailored hints, but interactivity and narrative flexibility were limited by static scripts.

Generative AI and LLMs now enable naturalistic dialogue, context-aware narratives, and real-time personalized feedback, extending beyond fixed responses (El Hajji, M. et al., 2025). In immersive VR, LLM-driven non-player characters scaffold exploration and enhance presence through adaptive dialogue (Gianni, A. et al., 2025). Design approaches emphasize human-centered, iterative development, integrating educator input, prompt engineering, and continuous feedback to align AI behaviors with pedagogical goals (Harder, N. et al., 2025).

Immersive and non-immersive SBL environments offer distinct advantages. VR supports embodied interaction, situational realism, and improved retention in professional and safety training (Hemminki-Reijonen, U. et al., 2025). Non-immersive 2D and text-based systems prioritize accessibility and scalability, leveraging LLMs to generate branching narratives and adaptive prompts (Stenseth, H. et al., 2025). Hybrid designs that combine narrative-driven LLM layers with optional immersive overlays can balance fidelity, engagement, and broad accessibility.

2.3. Evaluation frameworks in SBL

Evaluating SBL effectiveness necessitates frameworks that capture multiple learning dimensions including knowledge acquisition, engagement, fidelity, and transfer. Historically, evaluation in SBL research (Lee, A. et al., 2023) has incorporated performance tasks, retention tests, and subjective measures of presence and satisfaction, particularly in immersive environments where psychological fidelity correlates with motivation and engagement. Safety training research (Hemminki-Reijonen, U. et al., 2025), for example, has applied mixed-methods evaluation designs that combine task performance (e.g., correct procedural steps) with subjective self-efficacy and system acceptance measures to ascertain both skill outcomes and learner experience in complex, high-risk scenarios.

However, the integration of generative LLMs introduces new evaluation needs that extend beyond traditional metrics. This is because of the adaptive, dialogic nature of LLM output that requires assessment of conversational coherence, alignment with learning objectives, ethical compliance, and interaction quality, which are not fully addressed by conventional frameworks (Voultsiou, E. et al., 2025). Emerging meta frameworks (Mustafa, M. et al., 2024), such as multidimensional pyramids incorporating safety, trustworthiness, and operational effectiveness, offer promising starting points for capturing AI-specific constructs like reliability and human-AI interaction quality alongside classic pedagogical outcomes. These frameworks advocate a layered approach where core instructional safety and ethical compliance anchor higher-order assessments such as usability, personalization quality, and learning impact. For comprehensive evaluation in LLM-augmented SBL, melding these AI-centric constructs with established educational outcomes enables a more nuanced interpretation of both what learners accomplish and how they engage with generative adaptive systems.

3. Methodology

The core of this research involves a systematic review of literature on AI applications in Scenario-Based Learning (SBL), conducted following PRISMA 2020 guidelines (Haddaway et al., 2022). Database searches were performed in IEEE Xplore, SpringerLink, and ScienceDirect using Boolean queries targeting SBL and LLM applications published between 2022 and 2025 (see Appendix 1). Records were deduplicated, screened by title and abstract, and assessed for full-text eligibility based on predefined inclusion criteria. The PRISMA flow diagram (Appendix 2) documents the screening and selection process to ensure transparency and reproducibility. Searches yielded 97 open-access review articles from SpringerLink, 1 from IEEE Xplore, and 4 from ScienceDirect, totalling 102 articles. Following title and abstract screening, 50 articles were retained, and full-text assessment resulted in 10 studies selected for detailed synthesis. These studies were tabulated in Figure 1 and are described in Section 4.

4. Results

The ten literature articles were categorised by the title of the review, author, their immersiveness, which directly maps to immersive or non-immersive scenarios, and their use case, which explains the field that the literature was directly impacting.

Figure 1. Tabulated results of the literature overview used for synthesis.

Title	Authors	Immersive / Non-Immersive	Use Case
AI-Supported VR Ethical Dilemmas	Liu et al., 2025	Immersive	Education - Ethics
VR Simulation in Teacher Training (FCT)	Hu & Chan, 2025	Immersive	Education
AR Sustainability Scenarios	Ahmed et al., 2023	Immersive	Education
“Bill”: AI Clinical Scenario Coach	Rajputh, Seetharam, & Balasubramanian, 2019	Immersive	Medicine
AI-GBL Leadership Simulations	Lee et al., 2024	Immersive	Education
GenAI Patient Simulator versus VR	Lee, Park, & Kim, 2024	Immersive	Medicine
ChatGPT for Pre-Service Teacher Scenarios	Liu, Yang, & Xu, 2023	Non-Immersive	Education
AI-Generated Scenario Lessons	Zhang, Wang, & Li, 2024	Non-Immersive	Education
Online Nursing Scenario Modules	Kim & Torres, 2023	Non-Immersive	Medicine
Generative AI Literacy in SBL	Rajputh et al., 2025	Non-Immersive	Education

5. Discussion

5.1. Realism, adaptability, and learner engagement

Large Language Models enhance Scenario-Based Learning by increasing cognitive realism, adaptability, and engagement across immersive and non-immersive environments. In VR and AR scenarios, LLM-driven agents dynamically respond to learner actions, supporting realistic decision-making and professional skill rehearsal (Lee et al., 2024; Ahmed et al., 2023). In non-immersive text-based or 2D environments, LLMs provide personalized feedback and coherent narrative branching, improving perceived authenticity and motivation despite limited sensory immersion (Liu, Yang, & Xu, 2023). Across modalities, LLMs strengthen social and cognitive fidelity, enabling meaningful learning transfer.

5.2. Design processes and pedagogical strategies for AI and LLM integration

Effective LLM integration relies on human-centered, iterative design. Structured scenario authoring, prompt scaffolding, and educator oversight ensure pedagogical alignment and reliability (Rajput et al., 2019; Hu & Chan, 2025). Pedagogically, LLMs are most effective when embedded in guided experiential learning: as intelligent agents in immersive environments or adaptive tutors in non-immersive contexts (Ahmed et al., 2023; Kim & Torres, 2023). Hybrid AI architectures combining generative flexibility with rule-based controls are consistently preferred.

5.3. Evaluation frameworks for LLM-Driven SBL systems

Evaluation frameworks in recent SBL research commonly assess learning outcomes, engagement, and performance, with immersive studies additionally emphasizing presence and psychological fidelity (Lee et al., 2024; Kim & Torres, 2023). However, the introduction of LLM-based adaptivity necessitates expanded criteria, including conversational coherence, feedback quality, learner agency, and ethical AI behavior (Ahmed et al., 2023; Hu & Chan, 2025). The literature supports the use of mixed-methods evaluation approaches and highlights the need to extend existing frameworks to systematically assess both educational impact and AI reliability in LLM-driven SBL systems.

6. Conclusion

In our study, we sought out to review how AI models and Large Language Models enhance Scenario-Based Learning by supporting adaptive interaction, cognitive realism, and learner engagement across both immersive and non-immersive environments. However, their educational effectiveness depends on deliberate integration rather than unrestricted generation. The reviewed literature indicates that hybrid AI architectures, combining the generative adaptability of LLMs with rule-based or knowledge-driven controls, provide the most effective balance between flexibility, instructional accuracy, and safety. These findings directly inform the evaluation criteria adopted in our work, which extend beyond traditional learning outcomes to include learner engagement, realism, conversational coherence, adaptive feedback quality, and learner agency. Future research should empirically validate these criteria through controlled comparative studies of different AI architectures, including large general-purpose and domain-specific models, while also examining long-term learning transfer and ethical considerations in safety-critical training contexts to ensure that AI-enhanced SBL systems remain both effective and trustworthy.

References

- Ahmed, S., Brooks, R., & Liu, F. (2023). Frameworks for evaluating scenario-based learning in VR: A mixed-methods approach. *Journal of Computer Assisted Learning*, 39(4), 987–1002. <https://doi.org/10.1111/jcal.12753>
- Clark, R. C. (2012). *Scenario-based learning: Evidence-based guidelines for online workforce learning*. Pfeiffer.
- El Hajji, M., Ait Baha, T., Berka, A., Ait Nacer, H., El Aouifi, H., & Es-Saady, Y. (2025). An architecture for intelligent tutoring in virtual reality: Integrating LLMs and multimodal interaction for immersive learning. *Information*, 16(7), 556. <https://doi.org/10.3390/info16070556>
- Gianni, A. M., Nikolakis, N., & Antoniadis, N. (2025). An LLM-based learning framework for adaptive feedback mechanisms in gamified XR. *Computers & Education: X Reality*, 7, 100116. <https://doi.org/10.1016/j.cae.xr.2025.100116>

- Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Systematic Reviews*, 18, e1230. <https://doi.org/10.1002/cl2.1230>
- Harder, N., Ali, F., Turner, S., Workum, K., & Gillman, L. (2025). Comparing artificial intelligence-enhanced virtual reality and simulated patient simulations in undergraduate nursing education. *Clinical Simulation in Nursing*, 105, 101780. <https://doi.org/10.1016/j.ecns.2025.101780>
- Hu, W., & Chan, T. (2025). Human-centered design of AI-enhanced scenario-based learning systems. *Interactive Learning Environments*, Advance online publication, 1–15. <https://doi.org/10.1080/10494820.2025.0000000>
- Hemminki-Reijonen, U., Hassan, N. M. A. M., Huottilainen, M., Koivisto, J.-M., & Cowley, B. U. (2025). Design of generative AI-powered pedagogy for virtual reality environments in higher education. *NPJ Science of Learning*, 10, Article 31. <https://doi.org/10.1038/s41539-025-00326-1>
- Kim, J., & Torres, A. (2023). Evaluating presence and engagement in VR-based classroom simulations. *Interactive Learning Environments*, 33(5), 741–758. <https://doi.org/10.1080/10494820.2022.2141234>
- Lee, A. V. Y., Tan, S. C., & Teo, C. L. (2023). Designs and practices using generative AI for sustainable student discourse and knowledge creation. *Smart Learning Environments*, 10, Article 59. <https://doi.org/10.1186/s40561-023-00279-1>
- Lee, S., Park, H., & Kim, J. (2024). Immersion and presence in VR fire safety training: Impacts on evacuation skills. *Safety Science*, 168, 106924. <https://doi.org/10.1016/j.ssci.2023.106924>
- Liu, Y., Yang, H., & Xu, J. (2023). Large language models in teacher training simulations: Enhancing classroom management skills. *Computers & Education*, 195, 104673. <https://doi.org/10.1016/j.compedu.2022.104673>
- Mustafa, M. Y., Tlili, A., Lampropoulos, G., Huang, R., Jandrić, P., Zhao, J., Salha, S., Xu, L., Panda, S., & López-Pernas, S. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): A roadmap to a future research agenda. *Smart Learning Environments*, 11, Article 59. <https://doi.org/10.1186/s40561-024-00350-5>
- Rajput, R., Seetharam, R., & Balasubramanian, K. M. (2019). A structured framework for designing scenario-based e-learning systems. *International Journal of Emerging Technologies in Learning*, 14(21), 229–237. <https://doi.org/10.3991/ijet.v14i21.11795>
- Stenseth, H. V., Steindal, S. A., Solberg, M. T., Ølnes, M. A., Sørensen, A. L., Strandell-Laine, C., Olaussen, C., Farsjø Aure, C., Pedersen, I., Zlamal, J., Gue Martini, J., Bresolin, P., Wang Linnerud, S. C., & Nes, A. A. G. (2025). Simulation-based learning supported by technology to enhance critical thinking in nursing students: Scoping review. *Journal of Medical Internet Research*, 27, e58744. <https://doi.org/10.2196/58744>
- Tan, L. Y., Hu, S., Yeo, D. J., & Cheong, K. H. (2025). Artificial intelligence-enabled adaptive learning platforms: A review. *Computers & Education: Artificial Intelligence*, 9, 100429. <https://doi.org/10.1016/j.caeai.2025.100429>
- Voultsiou, E., & Moussiades, L. (2025). A systematic review of AI, VR, and LLM applications in special education: Opportunities, challenges, and future directions. *Education and Information Technologies*, 30, 19141–19181. <https://doi.org/10.1007/s10639-025-13550-4>
- World Health Organization. (2022). *WHO guideline: Recommendations on digital interventions for health system strengthening*. WHO. <https://www.who.int/publications/i/item/9789240052192>
- Zhang, X., Wang, S., & Li, P. (2024). Generative VR for education: Using LLMs to power adaptive NPCs in scenario-based learning. *British Journal of Educational Technology*, 55(2), 423–438. <https://doi.org/10.1111/bjet.13356>