Role-Play Simulation with AI in Education:
Historical Foundations, Current Trends, and New Frontiers

Abstract

This systematic literature review synthesized 67 scholarly publications to examine how the emergence of Artificial Intelligence (AI) transforms the pedagogical uses of role-play simulation (RPS) in education. We examined the trends of using RPS across major, education level, and learning outcomes before and after the emergence of AI. We explored the characteristics of using AI in RPS. The results indicated that AI-powered RPS included more disciplines and a broader range of educational levels, and led to the emergence of new learning outcome categories. Our findings revealed that the current RPS with AI relied heavily on closed-source, single-modal models. The paper discussed the limitations of existing research and proposed future directions for research on RPS in education.

Keywords: Role-Play Simulation, AI, GenAI, LLM, Education, Historical Development, Open/Closed Model, Single/Multi-modal

Introduction

Role-Play Simulations (RPS) is a pedagogical approach combining role-play techniques with simulation to foster deep, experiential learning (DeNeve & Heppner, 1997). It has been widely used in the education field for training teachers and motivating students, developing students' communication skills (Cowin et al., 2023; Park et al., 2024; Xu et al., 2025). However, research also found that traditional RPS had some limitations, including high cost, relying on human operators, and restrictions on physical locations. (Budin, 2024). The emergence of AI, particularly generative AI (GenAI), has further expanded the potential of RPS in education, highlighting its distinctive advantages in enhancing learning experiences and fostering instructional innovation. The integration of AI technologies has transformed the interaction model of RPS from predominantly human-to-human exchanges to human-AI interactions. The research has already developed AI-based applications for RPS such as TutorUp (Pan et al., 2025) to train teachers in a more engaging and risk-free context.

Despite the growing presence of using AI in RPS for educational purposes, existing research has primarily focused on evaluating its effectiveness within specific role-play scenarios, often through user-centered measures such as engagement, learning motivation, or learning outcomes (Talbot et al., 2012; Vieriu & Petrea, 2025). However, the literature lacked comprehensive reviews that systematically examined how the emergence of AI had transformed the design, implementation, and evolution of RPS in education. Moreover, previous review studies (Bhattacharjee & Ghosh, 2013; Burgues et al., 2024; Dai & Ke, 2022; Rosli & Zakaria, 2024) have primarily explored RPS from pedagogical perspectives, focusing on its learning outcomes and implementation strategies. However, none of these reviews have addressed the integration of GenAI within RPS in educational contexts, nor have they examined how the emergence of AI has driven historical changes in the design and use of RPS in education. Moreover, the emergence of GenAI in education has also raised concerns about data privacy and ethical issues, particularly in relation to the use of open- versus closed-source AI models as well as single- versus multi-modal systems. This literature review aims to bridge these gaps by systematically analyzing the historical evolution of RPS in education with the advent of AI, along with the characteristics of the underlying AI tools, including their source and model type. Below are our research questions:

1. What were the general trends in educational role-play simulation across academic majors, education levels, and learning outcomes before and after the emergence of AI?
2. What were the key characteristics of current research on the use of AI in educational role-play simulation?

Methodology

This study applied a systematic review methodology, following Wright et al. (2007). To ensure conceptual clarity, delineate the scope of the review, and support consistent analysis, this study defined key terms (see Table 1), which guided the selection of articles and framed the interpretation of findings.

Article Selection and Criteria

To analyze the results separately, the keywords searching for 'RPS with AI' and 'PRS without AI' in education were separated. For 'RPS without AI', we used the keywords with the boolean search string: ("role* play*" OR "simulat*" OR "role-play simulation" OR "simulation role play") AND ("education" OR "teach*" OR "learn*"). For studies on RPS with AI, the terms related to AI ("AI" OR "GPT" OR "LLM" OR "Chatbot") were included from this string.

Team members searched the keywords in six databases for with AI and without AI together: Web of Science, ScienceDirect, ProQuest Dissertation and Theses, ERIC, Education Source Ultimate, and Semantic Scholar, with additional resources on Google Scholar, more than 26,000 total articles were selected for the initial result (with AI and without AI). And based on the inclusion and exclusion criteria (see Table 2), after the second round screening of the title and related abstract, the final round retained 67 articles, which include 52 articles for SRP with AI, and 15 articles were similarly identified for analysis of RPS without AI.

Data Analysis and Procedure

The research team recorded the articles using two different tables, one for those without AI and another for those with AI, to clearly distinguish the RPS practices in traditional educational contexts from those influenced by AI technologies. We tracked the general trends in RPS with or without the use of AI in education by examining the major, education level, and learning outcomes for each paper to make a more accurate comparison.

An inductive approach (Braun & Clarke, 2006) was employed to identify emergent themes related to learning outcomes. Sub-categories were first identified by linking patterned meanings using constant comparative analysis (Glaser & Strauss, 1967), as illustrated in Table 2. Our analysis yielded 14 sub-categories in the non-AI context, with an additional 9 sub-categories emerging exclusively in the AI-supported context. By grouping the initial sub-categories for non-AI context, six initial categories were identified for articles without AI. As a result, three additional categories were identified in the 'With AI' dataset. Finally, all categories were organized by grouping similar learning outcomes into three-dimensional themes: 1) Behavior, 2) Cognition, and 3) Emotion (see Table 2). Additionally, the majors in higher education were merely coded since few articles reported in the K-12 period. To answer research question 2, building on the literature reviews, five categories: 1) which AI tools were used, 2) whether the tools were adopted for LLM, 3) what the baseline models were, 4) whether the models were closed or open source, and 5) what the application scenes are, were developed to systematically identify and compare the tools and model types that were used in RPS with AI.

To ensure the credibility and validity of the analysis process, a triangulation strategy was used (Carter, 2014). Each article was coded in pairs by two members independently. Discrepancies were discussed and resolved through group consensus to ensure reliable findings.

Results

General Trends of RPS in Education

As shown in Figures 1(a) and (b), teacher education and language majors were two majors used in RPS. With AI, computer science (14.6%) emerged as a new major while the engineering major dropped from 25% to 9.8%. Moreover, the application of RPS research expanded to a more diverse range of majors, such as math and social work.

For the education level (see Figures 1(c) and (d)), the result showed that with the emergence of AI, the studies of RPS focused on a broader range of educational levels, from pre-k to vocational levels. The team also noticed an increase in the focus on undergraduate level, rising from 36.8% in studies without using AI to 64.8% in those using AI.

In terms of learning outcomes (see Figures 1(e) and (f)), the results indicated that behavior dimensions accounted for the largest proportion in both AI and non-AI RPS studies. We also found that the application in the emotional dimension increased with the integration of AI. However, with the emergence of AI, new learning outcome categories have also been identified, including using AI to manage tasks and enhance learning through in-time reflections (SFR, 9.4%, Baucon & de Carvalho, 2024; Hu et al., 2025), strengthening responsibility in one's learning processes (ESL, 2.4%, Malik, 2024; Ng et al., 2024) and receiving emotional assistance based on personalized needs (RSF, 11%, Son & Lee, 2024; Zhang et al., 2024). Moreover, we observed in Figures (g) and (h) that the proportion of studies focusing on ESSA remained the highest in both AI and non-AI contexts, while the emphasis on ILP decreased from 20.7% to 12.6%. Our analysis also identified nine new subcategories that emerged in AI, such as lowering stress in learning processes (PEE3, Lee et al., 2024a; Zhang et al., 2024) and supporting the management of emotions and specific needs during learning (RSF2, Pan et al., 2025; Xu et al., 2025). With the emergence of AI, although DU1 remained the dominant category, DU2 and DU3 also appeared, collectively representing the broader DU classification. In the ECA category, ECA1, ECA2, and ECA3 were present in studies without AI; however, this distribution shifted to ECA2, ECA3, and ECA4 in AI-supported contexts. Additionally, with the transition from non-AI to AI-integrated studies, the proportion of articles addressing all three themes in one paper increased substantially from 6.7% to 25%.

Key Characteristics of AI in Educational RPS

RPS has followed an interconnected but asymmetrical evolutionary trajectory, evolving from human-to-human interaction to AI-enhanced strategies, and most recently to human-GenAI interaction (see Figure 2). Our primary finding is that current educational RPS predominantly adopt GenAI with a significant reliance on closed-source, single-modal Large Language Models (LLMs), such as OpenAI's GPT family (Lee et al., 2024c; Park et al., 2024) and Google's Gemini (Strobel et al., 2024). Specifically, the public release of ChatGPT catalyzed a surge in educational applications (Lo, 2023). While a few studies explored the potential of open-source alternatives (Höhn et al., 2024; Lee et al., 2024b), the prevailing trend leans heavily towards relying on the advanced performance of closed models (Divanji et al., 2024; Gervás et al., 2023). However, this reliance introduces challenges regarding transparency and customization, while the focus on single-modality constrains the potential richness of the simulation by excluding visual or auditory interactive elements. To better align AI behavior with pedagogical goals and make models more adapted to specific educational contexts, some studies utilize customization techniques like sophisticated prompt engineering or fine-tuning (Son et al., 2024; Xu et al., 2025).

In addition, our study reveals that the role AI plays in RPS has undergone significant changes alongside the development of AI technologies. In contrast to the limited, predefined, or scripted roles before GenAI emerged, a wide variety of dynamic personas are now embodied, fostering more dynamic and interactive learning experiences. AI is currently integrated into educational RPS in several primary pedagogical roles: a simulated student (Pan et al., 2025; Mikeska & Howell, 2020), a teaching assistant or coach (Markel et al., 2023; Liu & M'Hiri, 2024), and a conversational partner (Yue et al., 2024; Du & Daniel, 2024).

Discussion and Future Directions

While teacher education and language studies have remained dominant, computer science has emerged as a new area of focus, likely driven by AI's alignment with computational thinking and technical skill development. In contrast, the decline in engineering-related RPS suggests a redistribution of simulation-based learning tools across fields. The expansion into disciplines such as mathematics and social work further indicates that AI has enhanced the adaptability and applicability of RPS across a broader range of educational contexts. Our review revealed that research on RPS education in AI predominantly targets the undergraduate level. We believe that integrating RPS and AI in education is not only feasible but also valuable at the K–12 level, and we call for more research focus in this area.

The new categories and sub-categories in learning outcomes (e.g, ESL, RSF2) indicated that AI may foster more personalized learning in RPS education. Despite the continued dominance of behavior learning outcomes, the increased attention to emotion suggests a growing recognition of affective factors in learning. Future research should further examine how AI-driven tools can support these emerging outcomes, particularly in underexplored categories such as ESL1. Additionally, the substantial rise in studies addressing all three themes concurrently highlights the potential of AI to support a broader range of learning outcomes. Future research should explore how AI can facilitate multi-dimensional learning outcomes.

A key finding was the current significant gap that existed in the exploration of open-source models, like LLaMA (Touvron et al., 2023), Qwen (Bai et al., 2023), and the emergent reasoning model DeepSeek-R1 (Guo et al., 2025). Future work should conduct more studies on open-source models to evaluate their effectiveness, security, and accessibility in diverse educational domains. Such research is crucial for developing more equitable and adaptable AI-driven simulators.

Furthermore, the reviewed simulations relied almost exclusively on text-only interactions. This approach overlooked the pedagogical benefits of multimodal learning, where visual and auditory information can significantly enhance comprehension and engagement. Multimodal applications incorporating elements like images, animations, or synthesized speech could create more immersive and effective learning experiences that better cater to different learning styles and subject matter requirements.

References

Barry Issenberg, S., Mcgaghie, W. C., Petrusa, E. R., Lee Gordon, D., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical teacher*, *27*(1), 10-28.

Basri, W. S. (2024). Effectiveness of AI-powered Tutoring Systems in Enhancing Learning Outcomes. *Eurasian Journal of Educational Research (EJER)*, (110).

Baucon, A., & de Carvalho, C. N. (2024). Can AI get a degree in Geoscience? performance analysis of a gpt-based Artificial Intelligence System trained for earth science (geologyoracle). *Geoheritage*, *16*(4). https://doi.org/10.1007/s12371-024-01011-2

Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., Grundy, S., & Hassan, G. M. (2025). Implementing generative AI (genai) in Higher Education: A systematic review of case studies. *Computers and Education: Artificial Intelligence*, *8*, 100407.

Bhattacharjee, S., & Ghosh, S. (2013, April). Usefulness of Role-Playing Teaching in Construction Education: A systematic review. In *49th ASC annual international conference, San Luis Obispo, CA*.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2013). Successful qualitative research: A practical guide for beginners. (2014). *QMiP Bulletin*, *1*(18), 39–41. https://doi.org/10.53841/bpsqmip.2014.1.18.39

Budin, S. (2024). Three approaches to using mixed reality simulations for teacher preparation and recruitment of future teachers. *Education Sciences*, *14*(1), 75.

Burgues, M., Goujet, R., & Zaraik, J. (2024). Learning Soft Skills with an AI-Based Simulation Role-Play: A Literature Review. *EDULEARN24 Proceedings*, 6285-6293.

Burenkova, O. M., Arkhipova, I. V., Semenov, S. A., & Samarenkina, S. Z. (2015). Motivation within role-playing as a means to intensify college students' educational activity. *International Education Studies*, *8*(6). https://doi.org/10.5539/ies.v8n6p211

Canonigo, A. M. (2024). Levering AI to enhance students' conceptual understanding and confidence in mathematics. *Journal of Computer Assisted Learning*, *40*(6), 3215–3229.

Cao, Y., Li, S., Li, Y., Lu, Y., & Wei, X. (2023). A Survey on Generative AI. *arXiv preprint arXiv:2303.04226*.

Chang, M., Kuo, R., Hirose, M., Chen, G.-D., & Kinshuk. (2009). In *Learning by playing: Game-Based Education System Design and Development: 4th international conference on E-learning and games, edutainment 2009, August 9-11, 2009, Banff, Canada, proceedings*. Heidelberg; Springer.

Chien, C.C., Chan, H.Y., & Hou, H.T. (2024). Learning by playing with Generative AI: Design and evaluation of a role-playing educational game with Generative AI as scaffolding for Instant Feedback Interaction. *Journal of Research on Technology in Education*, *57*(4), 894–913. https://doi.org/10.1080/15391523.2024.2338085

Cowin, J., Oberer, B., Lipuma, J., Leon, C., & Erkollar, A. (2024). Accelerating higher education transformation: Simulation-based training and AI coaching for educators-in-training. *International Conference on Interactive Collaborative Learning*, 532-541. Cham: Springer Nature Switzerland.

Dai, C. P., & Ke, F. (2022). Educational applications of artificial intelligence in simulation-based learning: A systematic mapping review. *Computers and Education: Artificial Intelligence*, *3*, 100087.

Darmawansah, D., Hwang, G.-J., Lin, C.-J., & Febiyani, F. (2025). An artificial intelligence-supported GFCA learning model to enhance L2 students' role-play performance, English speaking and Interaction Mindset. *Educational Technology Research and Development*.

DeNeve, K. M., & Heppner, M. J. (1997). Role play simulations: The assessment of an active learning technique and comparisons with traditional lectures. *Innovative higher education*, *21*(3), 231-246. Stevens, R. (2015). Role-play and student engagement: reflections from the classroom. *Teaching in Higher Education*, *20*(5), 481-492.

Divanji, R., Dangol, A., Lombard, E. J., Chen, K., & Rubin, J. D. (2024). Togethertales RPG: Prosocial

Skill Development through digitally mediated collaborative role-playing. *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference*, 1012–1015.

Docter, M. W., de Vries, T. N. D., Nguyen, H. D., & van Keulen, H. (2024). A proof-of-concept of an integrated VR and AI application to develop classroom management competencies in teachers in training. *Education Sciences*, *14*(5), 540. https://doi.org/10.3390/educsci14050540

Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered Chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, *6*, 100230. https://doi.org/10.1016/j.caeai.2024.100230

Gemini Team, Google. (2023). Gemini: A Family of Highly Capable Multimodal Models. *arXiv preprint arXiv:2312.11805.*

Gervás, P., León, C., Kumar, M., Méndez, G., & Bautista, S. (2025). Prompting an LLM chatbot to role play conversational situations for language practice. *Proceedings of the 17th International Conference on Computer Supported Education*, 257–264.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Aldine.

Guo, D., Yang, D., Zhang, H., Song, J., … Wu, W., et al. (2025). DeepSeek‑R1: Incentivizing reasoning capability in LLMs via reinforcement learning. *arXiv preprint arXiv:2501.12948.*

Höhn, S., Nasir, J., Tozadore, D. C., Paikan, A., Ziafati, P., & André, E. (2024). Beyond pretend-reality dualism: Frame analysis of LLM-Powered Role Play with social agents. *Proceedings of the 12th International Conference on Human-Agent Interaction*, 393–395. https://doi.org/10.1145/3687272.3690894

Hu, B., Zhu, J., Pei, Y., & Gu, X. (2025). Exploring the potential of LLM to enhance teaching plans through teaching simulation. *Science of Learning*, *10*(1).

Khan, S., & Javed, M. (2025). AI-Powered Chatbots in Education: Evaluating Their Effectiveness in Student Engagement and Learning Outcomes. *ACADEMIA Tech Frontiers Journal*, *1*(1), 11-21.

Kincaid, J. P., & Westerlund, K. K. (2009). Simulation in education and training. *Proceedings of the 2009 Winter Simulation Conference (WSC)*, 273–280. https://doi.org/10.1109/wsc.2009.5429337

Lee, D., Son, T., & Yeo, S. (2024a). Impacts of interacting with an AI chatbot on Preservice Teachers' responsive teaching skills in math education. *Journal of Computer Assisted Learning*, *41*(1). https://doi.org/10.1111/jcal.13091

Lee, D., & Yeo, S. (2022). Developing an AI-based chatbot for practicing responsive teaching in Mathematics. *Computers & Education*, *191*, 104646.

Lee, J., Huang, J.-X., Cho, M., Roh, Y.-H., Kwon, O.-W., & Lee, Y. (2024b). Developing conversational intelligent tutoring for speaking skills in Second language learning. *Lecture Notes in Computer Science*, 131–148. https://doi.org/10.1007/978-3-031-63028-6_11

Lee, L.K., Chan, E. H., Tong, K. K.L., Wong, N. K.H., Wu, B. S.Y., Fung, Y.C., Fong, E. K., Leong Hou, U., & Wu, N.-I. (2024c). Utilizing virtual reality and Generative AI chatbot for Job Interview Simulations. *2024 International Symposium on Educational Technology (ISET)*, 209–212. https://doi.org/10.1109/iset61814.2024.00048

Lim, M. Y., Dias, J., Aylett, R., & Paiva, A. (2009). Intelligent NPCs for Educational Role Play Game. In *Agents for Games and Simulations*. essay, Springer Nature.

Liu, M., & M'Hiri, F. (2024). Beyond traditional teaching: Large language models as simulated teaching assistants in Computer Science. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, 743–749. https://doi.org/10.1145/3626252.3630789

Lo, C. K. (2023). What is the impact of chatgpt on education? A Rapid Review of the literature. *Education Sciences*, *13*(4), 410. https://doi.org/10.3390/educsci13040410

Lu, X., & Wang, X. (2024). Generative students: Using LLM-simulated student profiles to support question item evaluation. *Proceedings of the Eleventh ACM Conference on Learning @ Scale*, 16–27. https://doi.org/10.1145/3657604.3662031

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J., Ogata, H., ... & Tsai, C. C. (2020). Challenges and future directions of big data and artificial intelligence in education. *Frontiers in psychology*, *11*,

580820.
Lucci, S., Musa, S. M., & Kopec, D. (2022). *Artificial Intelligence in the 21st Century*. Mercury Learning and Information.
Luke, S. E., Ford, D., Vaughn, M., & Fulchini-Scruggs, A. (2021). Using mixed reality simulation and Roleplay to develop preservice teachers' metacognitive awareness. *Journal of Technology and Teacher Education*, *29*(3), 389–413. https://doi.org/10.70725/164279zjkiny
Malik, N. (2024). Interacting with chat GPT: Impact on students and teaching-learning. *Sigma Journal of Engineering and Natural Sciences – Sigma Mühendislik ve Fen Bilimleri Dergisi*, 1736–1748. https://doi.org/10.14744/sigma.2023.00147
Markel, J. M., Opferman, S. G., Landay, J. A., & Piech, C. (2023). GPTeach: Interactive TA training with GPT-based students. *Proceedings of the Tenth ACM Conference on Learning @ Scale*, 226–236. https://doi.org/10.1145/3573051.3593393
Matsuda, N., Yarzebinski, E., Keiser, V., Raizada, R., Stylianides, G. J., Cohen, W. W., & Koedinger, K. R. (2011). Learning by teaching simstudent – an initial classroom baseline study comparing with Cognitive Tutor. *Lecture Notes in Computer Science*, 213–221. https://doi.org/10.1007/978-3-642-21869-9_29
Mikeska, J. N., & Howell, H. (2020). Simulations as practice‐based spaces to support elementary teachers
in learning how to facilitate argumentation‐focused science discussions. *Journal of Research in Science Teaching*, *57*(9), 1356–1399. https://doi.org/10.1002/tea.21659
Mishu, A., Ahamed, M. M., Akan, M. F., Abdul-Rab, S. D., Chowdhury, G., Ahmad, J., & Sultana, I. (2025). How AI Is Ushering in a New Era in ELT: Teachers' Perspectives.*Theory & Practice in Language Studies (TPLS)*, *15*(2).
Ng, D. T., Xinyu, C., Leung, J. K., & Chu, S. K. (2024). Fostering students' ai literacy development through educational games: AI knowledge, affective and cognitive engagement. *Journal of Computer Assisted Learning*, *40*(5), 2049–2064. https://doi.org/10.1111/jcal.13009
Nygren, T., Samuelsson, M., Hansson, P.-O., Efimova, E., & Bachelder, S. (2025). Ai versus human feedback in mixed reality simulations: Comparing LLM and expert mentoring in Preservice teacher education on controversial issues. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-025-00484-8
OpenAI. (2023). *ChatGPT (GPT‑3.5/4) [Large language model]*. https://chat.openai.com/
Pan, S., Schmucker, R., Garcia Bulle Bueno, B., Llanes, S. A., Albo Alarcón, F., Zhu, H., Teo, A., & Xia, M. (2025). TutorUp: What if your students were simulated? training tutors to address engagement challenges in online learning. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–18. https://doi.org/10.1145/3706598.3713589
Park, J., Yang, S., Yoon, C., Si, J., Jung, Y., & Kim, S. (2024). AI English conversation coaching platform in the metaverse: Focused on Korean users. *2024 IEEE 18th International Conference on Application of Information and Communication Technologies (AICT)*, 1–6. https://doi.org/10.1109/aict61888.2024.10740456
Rao, D., & Stupans, I. (2012). Exploring the potential of role play in higher education: development of a typology and teacher guidelines. *Innovations in Education and Teaching International*, *49*(4), 427-436.
Rosli, R. R., & Zakaria, A. F. (2024). The impact of digital role-playing games in engineering education: A preliminary review. In *2024 9th International STEM Education Conference (iSTEM-Ed)* (pp. 1-6). IEEE.
Rumore, D., Schenk, T., & Susskind, L. (2016). Role-play simulations for Climate Change Adaptation Education and engagement. *Nature Climate Change*, *6*(8), 745–750. https://doi.org/10.1038/nclimate3084
Russell, S. J., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach* (3rd ed.). Pearson Education.
Shanahan, M., McDonell, K., & Reynolds, L. (2023). Role play with large language models. *Nature*,

623(7987),493-498.
Son, T., Yeo, S., & Lee, D. (2024). Exploring elementary preservice teachers' responsive teaching in mathematics through an artificial intelligence-based chatbot. *Teaching and Teacher Education*, *146*, 104640. https://doi.org/10.1016/j.tate.2024.104640
Strobel, J., Medina, M., Guzman, E. S., & Van den Bogaard, M. (2024). Exploring AI bots as simulators in Human Subject Research: A novel approach to ethical and efficient experimentation in Engineering Education Research. *2024 IEEE Frontiers in Education Conference (FIE)*, 1–9.
Talbot, T. B., Sagae, K., John, B., & Rizzo, A. A. (2012). Sorting out the virtual patient: how to exploit artificial intelligence, game technology and sound educational practices to create engaging role-playing simulations. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)*, *4*(3), 1-19.
Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., ... & Lample, G. (2023). LLaMA: Open and Efficient Foundation Language Models. *arXiv preprint arXiv:2302.13971*.
Vieriu, A. M., & Petrea, G. (2025). The impact of artificial intelligence (AI) on students' academic development. *Education Sciences*, *15*(3), 343.
Violakis, P. (2025). Leveraging large language models for enhanced simulation-based learning in police and law enforcement. *Policing: A Journal of Policy and Practice*, *19*.
Wright, R. W., Brand, R. A., Dunn, W., & Spindler, K. P. (2007). How to Write a Systematic Review. *Clinical Orthopaedics and Related Research, 455*, 23–29.
Wu, T., Lee, H., Chen, P., Lin, C., & Huang, Y. (2024). Integrating peer assessment cycle into ChatGPT for STEM education: A randomised controlled trial on knowledge, skills, and attitudes enhancement. *Journal of Computer Assisted Learning*, *41*(1). https://doi.org/10.1111/jcal.13085
Xu, S., Wen, H.-N., Pan, H., Dominguez, D., Hu, D., & Zhang, X. (2025). Classroom simulacra: Building contextual student generative agents in online education for learning behavioral simulation. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–26.
Yue, M., Jong, M. S. Y., & Dai, Y. (2022). Pedagogical design of K-12 artificial intelligence education: A systematic review. *Sustainability*, *14*(23), 15620.
Yue, M., Lyu, W., Mifdal, W., Suh, J., Zhang, Y., & Yao, Z. (2024). MathVC: An llm-simulated multi-character virtual classroom for mathematics education. *arXiv preprint arXiv:2404.06711.*
Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenič, A., Spector, M., Liu, J., Yuan, J., & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity*, *2021*, 18–18.
Zhang, D., Wu, J. G., & Fu, Z. (2024). From shy to fly: Facilitating EFL learners' willingness to communicate with an AI chatbot and an intelligent tutoring system. *System*, *127*, 103501.
Zhang, Z., Zhang-Li, D., Yu, J., Gong, L., Zhou, J., Hao, Z., Jiang, J., Cao, J., Liu, H., Liu, Z., Hou, L., & Li, J. (2025). Simulating classroom education with LLM-empowered agents. *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, 10364–10379. https://doi.org/10.18653/v1/2025.naacl-long.520
Zhao, J., Jingru, Z., & Lu, Y. (2025). Enhancing design historical education through AI virtual characters role-playing narratives in serious games. *International Journal of Gaming and Computer-Mediated Simulations*, *17*(1), 1–20.
Zhao, W. X., Zhou, A., Li, J., Ma, W., Zhang, B., Ding, M., ... & Wen, J. (2023). A Survey of Large Language Models. *arXiv preprint arXiv:2303.1822*

Table 1
Definitions and Terminologies

Terminologies	Definitions
Artificial Intelligence (AI)	The study of agents that receive percepts from the environment and perform actions, with the goal of creating systems that act rationally and adaptively in complex environments (Russell & Norvig, 2010).
Generative AI (GenAI)	A branch of AI that learns patterns from large datasets to autonomously create novel content—text, images, audio, video, or code—that closely resembles human-produced artifacts (Cao et al., 2023).
Large Language Models (LLMs)	A class of powerful deep learning models trained on massive amounts of text data, typically employing Transformer-based neural network architectures with billions to trillions of parameters, trained on vast text corpora to generate and comprehend human language with high fluency across diverse NLP tasks (Zhao et al., 2023).
Role-Play (RP)	A teaching strategy where students assume specific roles and interact within simulated scenarios to achieve learning objectives, practice skills, and explore different perspectives in a dynamic and interactive manner (Stevens, R., 2015).
Simulation	In an educational context, this technique replicates real-world situations or events, allowing learners to engage in authentic tasks, make decisions, and experience consequences in a controlled and safe environment to acquire knowledge, skills, and attitudes (Issenberg et al., 2005).
Role-Play Simulation (RPS)	In education, it typically refers to the combined application of both role-play and simulation techniques. It provides learners with an immersive, experiential learning environment where they assume specific roles, interact and practice within a complex scenario that mimics a real-world context, thereby fostering deep learning and skill development (DeNeve & Heppner, 1997).

Table 2
Inclusion and Exclusion Criteria

Criterion	Inclusion	Exclusion
Settings	Focus on role-play simulation before and after using AI separately in some specific educational fields, especially in STEM education	Role-play simulation in health, medical, business, and counselling-related fields in education
Language	English	Languages other than English
Publication Type	Peer-reviewed journal articles, Conference proceedings, Book chapters	Posters, Preprints, news articles, stories
Type of Study	Empirical Study	Other studies besides the empirical study

Table 3
Learning Outcomes Theme

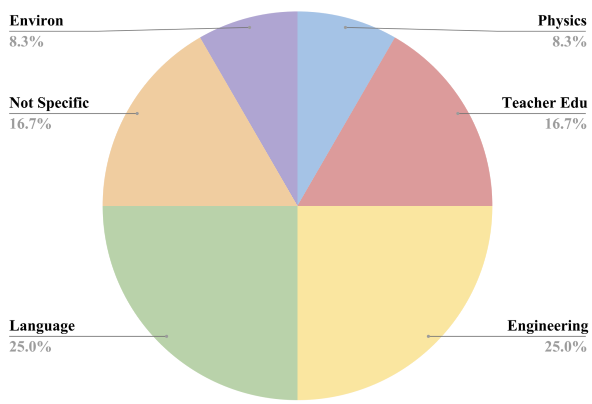
Theme	Category	Sub-Category
	Improve Learning Performance (ILP): Enhancing academic tasks	● ILP1: The quality of completing academic tasks, such as '*academic performance*', '*content*', '*context*'; ● ILP2: The level of competence, especially in language learning, such as '*proficiency*'; ● ILP3: The ability to apply to real-world scenarios, such as '*practical*'.
Behavior: Focuses on actions and improvement in task performance and interpersonal abilities	**Enhance Soft Skills and Ability (ESSA)**: Developing interactive and self-management competencies	● ESSA1: The ability to express information and interact with other people, such as '*communication*,' '*engagement*,' '*teamwork*,' '*interaction*', '*leadership*'* (only appear in AI); ● ESSA2*: The ability to analyze and evaluate ideas, such as '*critical thinking*'; ● ESSA3: The ability to plan and use proper methods to achieve goals, such as '*strategy*' and '*decision-making*'.
	Support Facilitation & Reflection (SFR)*: Using AI to manage tasks and enhance learning through in-time reflections	● SFR1*: Being aware of issues, such as '*noticing*'; ● SFR2*: The ability to organize and construct the assessments, such as '*task management*' and '*test design*'.

Cognition: Focuses on self-learning, deeper understanding, and awareness of thinking	**Deepen Understanding (DU)**: Enhancing depth in comprehension via the learning process	● DU1: Grasping the meanings and representing the information to aid comprehension, such as '*understanding*' and '*visualization*'; ● DU2*: Processes in reasoning in the mind, such as '*thinking*'; ● DU3*: Evaluating and reviewing one's behaviors, thoughts, and assessments, such as '*self-reflection*' and '*reflection*'.
	Enhance Cognitive Awareness (ECA): Developing awareness of one's thinking processes	● ECA1: The awareness of controlling and handling one's own learning processes, such as '*metacognition*'; ● ECA2: Evaluating one's own performance, such as '*self-assessment*'; ● ECA3: Conscious recognition and tracking of ideas or actions, such as '*awareness*' and '*monitoring*'*; ● ECA4: Belief in the ability to succeed, such as '*self-efficacy*'*.
	Enhance Self-Learning (ESL)*: Strengthening responsibility in one's learning processes	● ESL1*: Managing behaviors, thoughts, and direction to achieve the goals, such as '*self-regulation*', '*strategy use*', '*learner control*', and '*goal setting*'.
Emotion: Focuses on expressing positive experiences, empathy, and emotional support in learning processes	**Positive Emotion Experience (PEE)**: Supporting emotional status to enhance learning processes	● PEE1: Internal drive and belief in learning engagement and capability, such as '*confidence*' and '*motivation*'; ● PEE2: Feelings of enjoyment, satisfaction, or achievement, such as '*positive experience*'; ● PEE3*: Lowering stress in learning processes, such as '*reduce anxiety*'.
	Empathy and Perspective (EP): The ability to understand and connect with other people's perspectives	● EP1: Valuing and seeing other people's views and backgrounds, such as '*perspective*,' '*diversity*', and '*enrich dialogues*'*; ● EP2: Understanding and sharing other people's feelings, such as '*empathy*'.
	Receive Support and Feedback (RSF)*: Receiving emotional assistance based on personalized needs	● RSF1*: Emotional responses to evaluate the performance, such as '*feedback*'; ● RSF2*: Supporting the management of emotions and specific needs during learning, such as '*emotional scaffolding*' and '*personalized support*'.

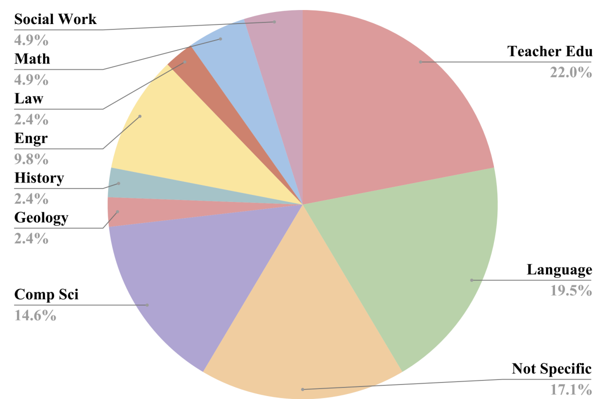
Note. The symbol * means the categories and sub-categories are only shown with AI in RPS.

Figure 1

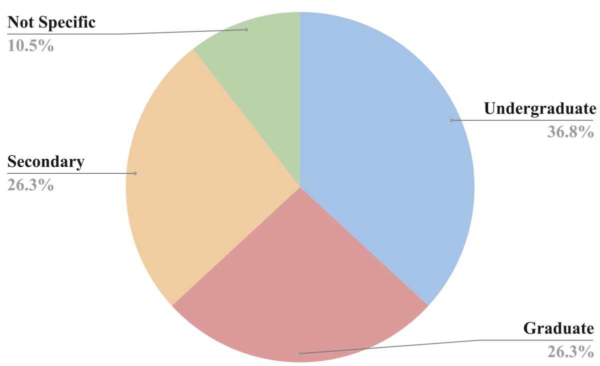
The Comparison of Distributions of Aspects in Education with/without AI



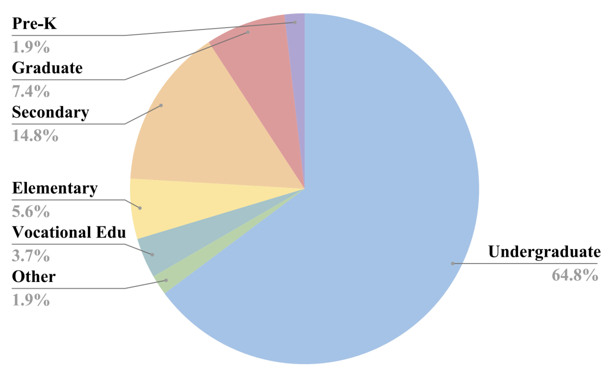
(a) Majors in Higher Education without AI



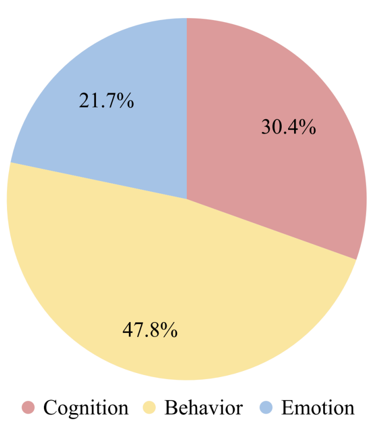
(b) Majors in Higher Education with AI



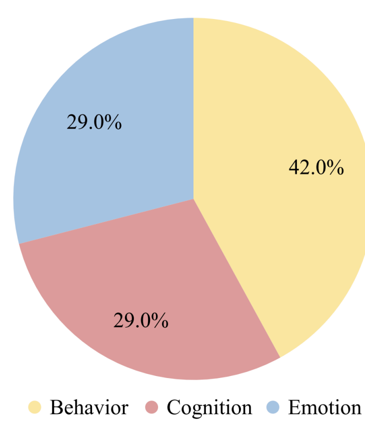
(c) Educational Levels without AI



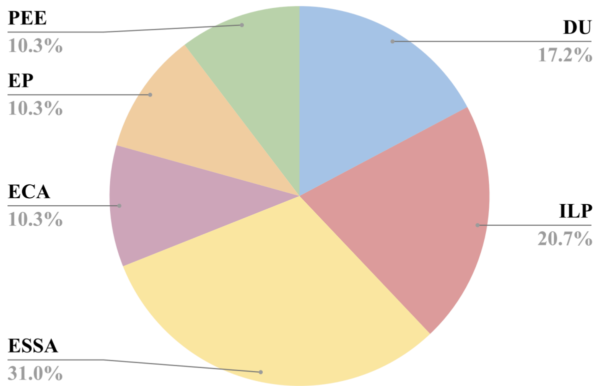
(d) Educational Levels with AI



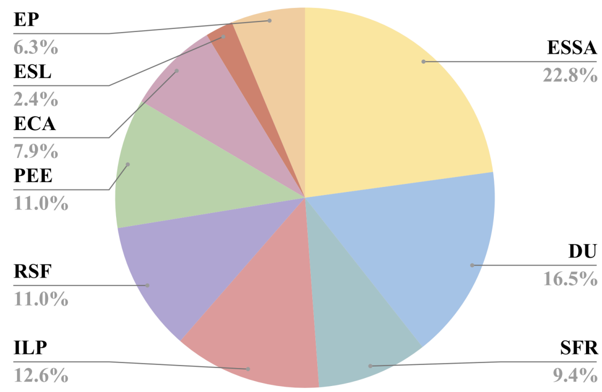
(e) Learning Outcome Themes without AI



(f) Learning Outcome Themes with AI



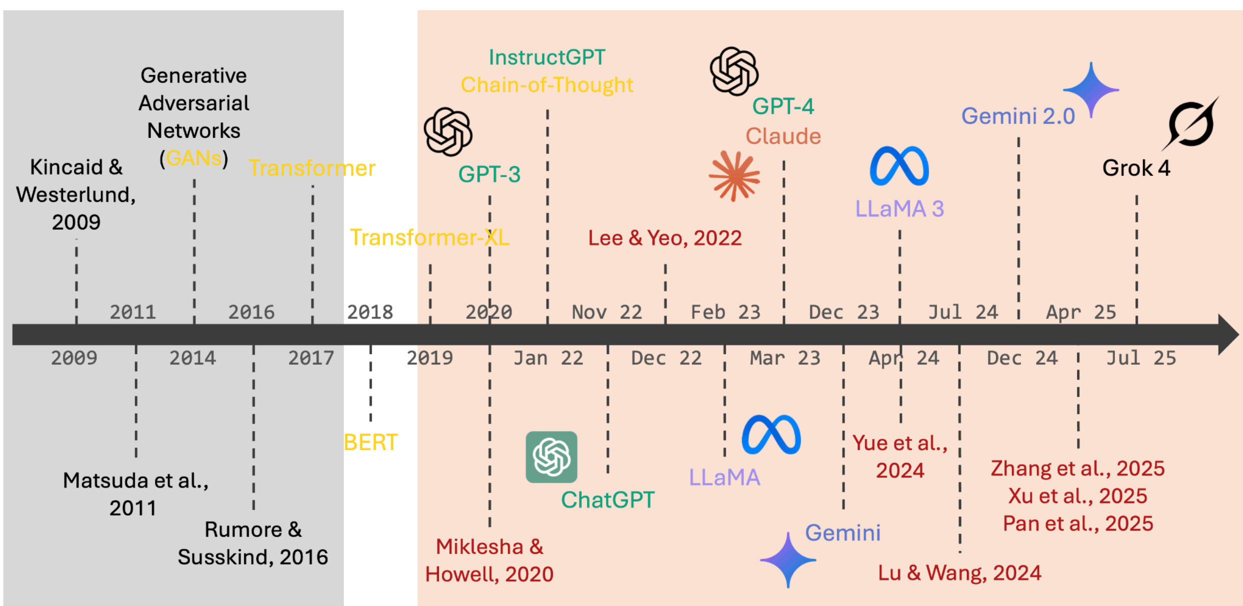
(g) Learning Outcome Categories without AI



(h) Learning Outcome Categories with AI

Figure 2

Timeline of AI and Edu Simulators' Development



Note. The black colored texts with gray background and red colored texts with pink background respectively represent education papers or applications without and with AI.