Smart scaffolds, smarter learners? Analysing the cognitive outcomes of AR and chatbot feedback

Jerry Chih-Yuan Sun, Even Yi-Wen Liu, Jocelyn Ting-Yu Shen
National Yang Ming Chiao Tung University, Taiwan

This study investigated the impact of chatbot and AR-based feedback scaffolding on university students' cognitive load during a scientific simulation inquiry. In a randomized experiment, 118 university students received feedback from one of four scaffolding conditions: traditional, chatbot, augmented reality (AR), or a combined chatbot-AR system. Cognitive load was assessed using a validated self-report scale, complemented with an open-ended questionnaire. Results indicated that chatbot scaffolding significantly increased intrinsic load, and the combined chatbot–AR scaffolding group showed significantly higher extraneous load than the AR-based group. These findings suggest that integrating multiple forms of scaffolding may elevate cognitive burden and should be implemented with caution to prevent excessive load on learners.

Keywords: chatbot, augmented reality, cognitive load, feedback scaffolding, STEM education

Introduction

With the rapid integration of emerging technologies into education, tools such as chatbots and augmented reality (AR) are increasingly employed to support inquiry-based learning. However, the simultaneous use of these technologies may introduce unintended cognitive demands. Thus, it is important to understand the impact of chatbot and AR-based feedback scaffolding in complex learning environments.

STEM education

Current STEM (Science, Technology, Engineering, and Mathematics) education emphasises engaging students in real-world inquiry to develop deep conceptual knowledge and reasoning skills. However, this is particularly challenging in disciplines like chemistry, where understanding requires connecting macroscopic observations to invisible particulate and symbolic levels. Scientific simulation software is a powerful tool to bridge this gap by visualizing complex and abstract phenomena. Nevertheless, its inherent interactive complexity can overwhelm learners, making it difficult to identify causal relationships while imposing an extraneous cognitive load that diminishes learning effectiveness (Chang & Yang, 2023; Pellas, 2023). To this end, this study implemented a scientific simulation inquiry teaching approach based on the STEM education conceptual model to tackle the identified challenges of using simulation software as mentioned above.

Chatbot feedback scaffolding and AR-based feedback scaffolding

Technology feedback scaffolding is pivotal in STEM education, providing the cognitive and motivational support for students to tackle complex tasks beyond their independent ability (Belland, 2017). Tools like chatbots can structure these tasks through targeted prompts to improve learning outcomes (Dever et al., 2024). However, while structured support can enhance inquiry skills, poorly designed guidance may increase extraneous cognitive load by introducing unnecessary interaction elements (Sweller, 2020). To mitigate the risk, the development of personalised scaffolding that responds to individual learner needs is essential. Chatbot feedback scaffolding offers adaptive support by delivering immediate, conversational, and tailored feedback, helping reduce cognitive overload and improve task performance (Brachten et al., 2020). Similarly, AR supports learning by overlaying virtual content onto real world environments, offering dynamic, visual feedback that fosters interactive engagement (Na & Sung, 2025). The use of pedagogical agents within AR environments further enables timely, contextual feedback aligned with student needs (Chen et al., 2025). While the integration of AR and chatbots holds promise, its combined impact in STEM inquiry learning remains underexplored. Therefore, this study designed and implemented feedback scaffolding that integrated chatbot and AR to support students within a scientific inquiry-based experimental learning context.

Feedback scaffolding on cognitive load

Cognitive load refers to the demands placed on working memory during learning and includes three types: intrinsic load, stemming from the inherent complexity of the material; extraneous load, resulting from poorly designed instruction; and germane load, the productive effort dedicated to schema construction (Sweller, 2020; Sweller et al., 1998). Effective instructional design seeks to reduce extraneous load and promote germane load through structured support. However, balancing these demands remains challenging in AR learning environments, which may impose excessive load without personalised scaffolding. Segmenting tasks with feedback in inquiry-based learning, suggested by Moon and Brockway (2019), may ease such burden. While existing evidence suggests that chatbot scaffolding can enhance learner focus and reduce cognitive demands (Schmidhuber et al., 2021), some studies report increased cognitive load due to users' unfamiliarity with chatbot interfaces (Nguyen et al., 2022). These inconclusive findings underscore the need to further investigate the effectiveness of chatbot-based feedback scaffolding in complex digital learning environments. Distributed scaffolding offers multiple, complementary forms of support that can enhance learning outcomes (Puntambekar & Kolodner, 1998); however, its effectiveness hinges on contextual alignment, personalisation, and the rules governing adaptive support (Belland et al., 2022). Extending these insights, we assumed chatbots can reduce cognitive load in AR-based learning by incrementally presenting content and enabling interactive, personalised support. Accordingly, this study examined whether the use of chatbot and AR-based feedback scaffolding significantly influences university students' intrinsic, extraneous, and germane cognitive load during a scientific simulation inquiry. Figure 1 shows the study framework.

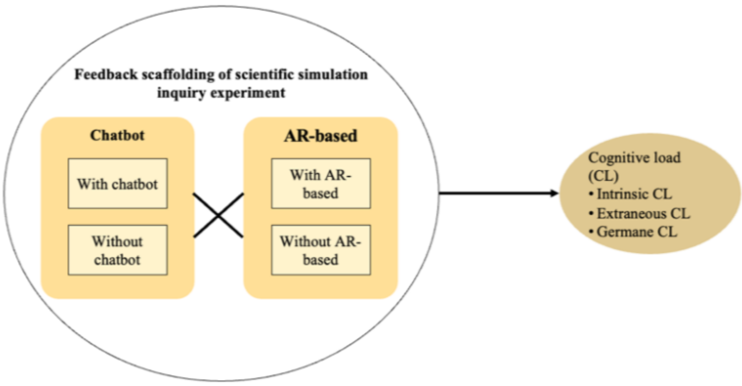

Figure 1. Research model

Method

In this study, 118 university students were recruited via social media by convenience sampling. After excluding invalid responses, 98 valid samples remained (48 males, 50 females; *M* age = 23.56, *SD* = 2.98). They were randomly assigned to the traditional feedback scaffolding group (T-FS group), the chatbot feedback scaffolding group (C-FS group), the AR-based feedback scaffolding group (AR-FS group), and the chatbot integrated with AR-based feedback scaffolding group (CAR-FS group). Cognitive load was measured using a 10-item, 6-point Likert scale adapted from Leppink et al. (2013), with three sub-dimensions: intrinsic, extraneous, and germane load. Cronbach's alpha values for each subscale were .88, .83, and .87, respectively, indicating good reliability. To complement the quantitative data, open-ended questions were included to elicit participants' reflections on cognitive experience and feedback presentation. Figure 2 illustrates the experimental procedure. All participants completed the same tablet-based scientific simulation tasks but received different feedback scaffolding. The T-FS group used a paper manual. The C-FS group accessed a chatbot via Line, an instant messaging application. Built on the Dialogflow platform, the chatbot provided conceptual and procedural knowledge through button-triggered replies tailored to students' learning needs. The AR-FS employed image-based AR developed via the UniteAR platform. Using the UniteAR app on their mobile devices, students scanned marked areas on tablets and textbooks to access real-time virtual texts and instructional videos.

Figures 3 and 4 present the user interfaces and operational schematics of the chatbot and AR-based scaffolding systems, respectively. The CAR-FS group combined both chatbot and AR scaffolding for integrated support across print and digital materials.

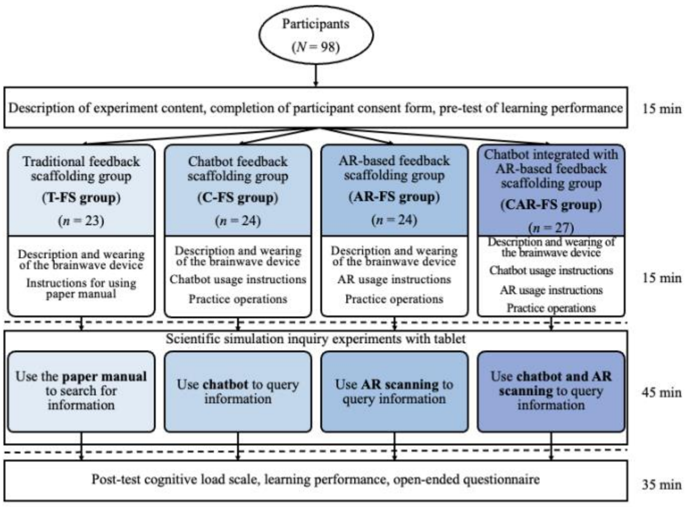

Figure 2. Experimental design

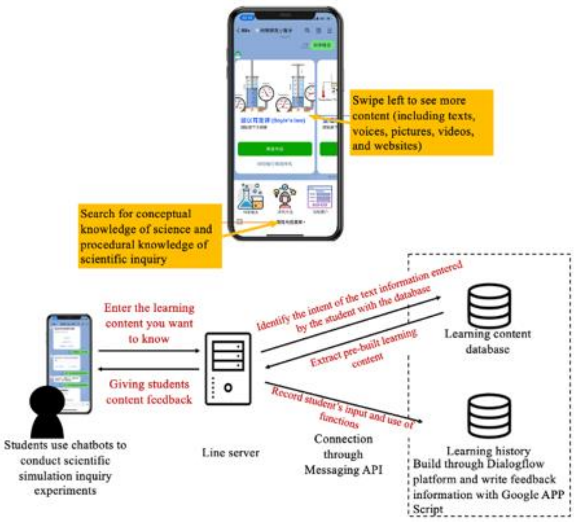

Figure 3. Interface and operation schematic of chatbot feedback scaffolding

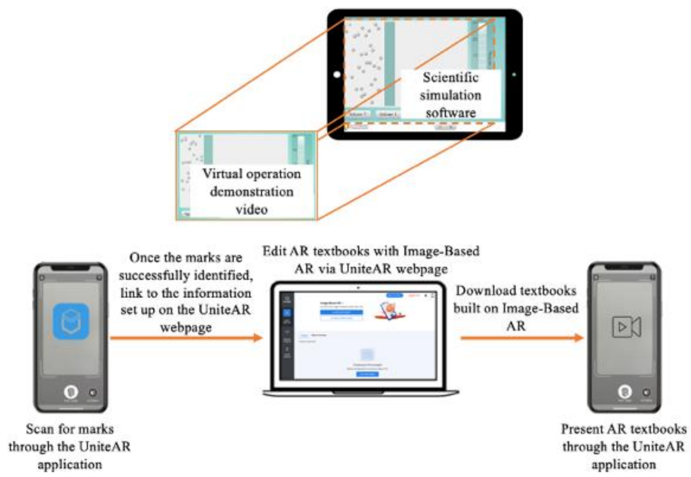

Figure 4. Interface and operation schematic of AR-based feedback scaffolding

Results and Discussion

As the three sub-dimensions of cognitive load represent distinct constructs, they were analysed separately rather than aggregated into a total score. A series of two-way ANOVAs were conducted for each dimension. Detailed analysis results are presented in Table 1. For intrinsic cognitive load, the main effect of chatbot feedback scaffolding was significant ($F = 4.10$, $p < .05$, $\eta^2 = .04$). Post-hoc comparisons indicated that students who received chatbot scaffolding reported higher intrinsic load ($M = 7.59$, $SD = 2.64$) than those without it ($M = 6.43$, $SD = 2.87$). Neither the main effect of AR-based feedback scaffolding nor the interaction effect reached statistical significance ($F = 0.44$, $p = .51$; $F = 3.16$, $p = .08$, respectively). For extraneous cognitive load, a significant interaction was observed between chatbot and AR-based scaffolding ($F = 4.06$, $p < .05$, $\eta^2 = .04$). Simple main effects analysis revealed that the chatbot scaffolding had a significant effect only when AR scaffolding was also present ($F = 7.88$, $p = .01$). Specifically, under the AR condition, students with chatbot scaffolding reported higher extraneous load ($M = 6.59$, $SD = 3.18$) than those without it ($M = 4.63$, $SD = 1.88$). No other simple main effects were statistically significant. For germane cognitive load, no significant main or interaction effects were found.

Table 1
Summary of the two-way analysis of variance results for cognitive load

Cognitive Load	**Source of Variation**	***df***	***F***	***p***	**Partial η²**	**Post-hoc**
Intrinsic	Inter-group					
	Chatbot	1	4.10	< .05	.04	With chatbot > Without chatbot
	AR	1	0.44	.51	.01	
	Chatbot*AR	1	3.16	.08	.03	
	Intra-group (error)	94				
Extrinsic	Inter-group					
	Chatbot	1	3.52	.06	.04	
	AR	1	0.01	.92	< .001	
	Chatbot*AR	1	4.06	< .05	.04	
	Intra-group (error)	94				
	Simple main effect					
	Chatbot feedback scaffolding					
	With AR	1	7.88	.01		With chatbot > Without chatbot
	Without AR	1	.01	.92		
	AR-based feedback scaffolding					
	With chatbot	1	1.67			
	Without chatbot	1	2.56			
Germane	Inter-group					
	Chatbot	1	0.04	.84	< .001	
	AR	1	1.29	.26	.01	
	Chatbot*AR	1	1.04	.31	.001	
	Intra-group (error)	94				

This study examined the impact of chatbot and AR-based feedback scaffolding on students' cognitive load during a scientific simulation inquiry. For intrinsic cognitive load, students who received chatbot scaffolding reported significantly higher levels than those who did not, regardless of AR integration. Although the chatbot was designed to break down complex tasks into manageable components (Moon & Brockway, 2019), open-ended responses suggested that the overall volume and complexity of content remained overwhelming, resulting in cognitive overload despite structured guidance. In terms of extraneous cognitive load, a significant interaction effect showed that increased load emerged only in the condition where both chatbot and AR scaffolding were implemented. While multiple scaffolding is generally intended to enhance learning opportunities (Puntambekar & Kolodner, 1998), and chatbots are often shown to reduce cognitive burden (Schmidhuber et al., 2021), participants' unfamiliarity with these technologies may have increased cognitive demands (Nguyen et al., 2022). Qualitative responses further confirmed that students' limited prior experience with the technologies introduced operational difficulties, increasing their perceived cognitive load. Regarding germane cognitive load, no significant main or interaction effects were observed across the groups. Although well-designed instructional scaffolding could enhance germane load by fostering deeper cognitive engagement (Paas et al., 2003), open-ended responses revealed that students across all conditions were able to focus on the essential scientific concepts and tasks, likely due to the smooth guidance offered by feedback scaffolding. This may explain the lack of significant differences in germane cognitive load among groups.

Conclusion

This study investigated the effects of chatbot and AR-based feedback scaffolding on university students' cognitive load during a scientific simulation inquiry. Results showed that chatbot scaffolding significantly increased intrinsic load, while the combined chatbot–AR condition significantly elevated extraneous load. These outcomes suggest that integrating multiple scaffolds may overwhelm learners, whereas using a single scaffold—preferably AR-based—may better support cognitive processing in inquiry-based learning. However, several research limitations should be noted. First, individual learner characteristics (e.g., learning styles) may have influenced outcomes. Second, unfamiliarity with chatbot or AR tools could have affected cognitive load. Third, the study relied solely on self-reported data and examined only two scaffolding types, limiting generalisability. Future research should explore varied scaffolding designs and account for learner differences to strengthen applicability of the results.

Reference

Belland, B. R. (2017). *Instructional scaffolding in STEM education: Strategies and efficacy evidence*. Springer Nature.

Belland, B. R., Lee, E., Zhang, A. Y., & Kim, C. (2022). Characterizing the most effective scaffolding approaches in engineering and technology education: A clustering approach. *Computer Applications in Engineering Education*, *30*(6), 1795-1812. https://doi.org/10.1002/cae.22556

Brachten, F., Brünker, F., Frick, N. R., Ross, B., & Stieglitz, S. (2020). On the ability of virtual agents to decrease cognitive load: An experimental study. *Information Systems and e-Business Management*, *18*(2), 187–207. https://doi.org/10.1007/s10257-020-00471-7

Chang, C.-C., & Yang, S.-T. (2023). Learners' positive and negative emotion, various cognitive processing, and cognitive effectiveness and efficiency in situated task-centered digital game-based learning with different scaffolds. *Interactive Learning Environments*, *32*(9), 5058-5077. https://doi.org/10.1080/10494820.2023.2209600

Chen, C., Jamiat, N., Rabu, S. N. A., & Du, S. (2025). Effects of Augmented Reality with Pedagogical Agent on EFL Primary Students' Vocabulary Acquisition, Motivation, and Technology Perceptions. *International Journal of Human–Computer Interaction*, 1-19. https://doi.org/10.1080/10447318.2024.2443534

Dever, D. A., Wiedbusch, M. D., Romero, S. M., & Azevedo, R. (2024). Investigating pedagogical agents' scaffolding of self-regulated learning in relation to learners' subgoals. *British Journal of Educational Technology*, *55*(4), 1290-1308. https://doi.org/10.1111/bjet.13432

Leppink, J., Paas, F., van der Vleuten, C. P., van Gog, T., & van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1

Moon, J. A., & Brockway, D. (2019). Facilitating learning in an interactive science simulation: The effects of task segmentation guidance on adults' inquiry-based learning and cognitive load. *Journal of Research on Technology in Education*, *51*(1), 77–100. https://doi.org/10.1080/15391523.2019.1566038

Na, H., & Sung, H. (2025). Learn math through motion: a technology-enhanced embodied approach with augmented reality for geometry learning in K-12 classrooms. *Interactive Learning Environments*, 1-14. https://doi.org/10.1080/10494820.2025.2450663

Nguyen, Q. N., Sidorova, A., & Torres, R. (2022). User interactions with chatbot interfaces vs. Menu-based interfaces: An empirical study. *Computers in Human Behavior*, *128*(2022), 107093. https://doi.org/10.1016/j.chb.2021.107093

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Pellas, N. (2023). Exploring relationships among students' computational thinking skills, emotions, and cognitive load using simulation games in primary education. *Journal of Computer Assisted Learning*, *39*(5), 1576-1590. https://doi.org/10.1111/jcal.12819

Puntambekar, S., & Kolodner, J. L. (1998, December 16–19). *Distributed scaffolding: Helping students learn in a "learning by design" environment* [Paper presentation]. The third International Conference of the Learning Sciences (ICLS'98), Atlanta, GA. https://www.cc.gatech.edu/projects/lbd/pdfs/distscaffold.pdf

Schmidhuber, J., Schlögl, S., & Ploder, C. (2021, September 8–10). *Cognitive load and productivity implications in human-chatbot interaction* [Paper presentation]. The 2021 IEEE 2nd International Conference on Human-Machine Systems (ICHMS), Magdeburg, Germany.

Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development*, *68*(1), 1–16. https://doi.org/10.1007/s11423-019-09701-3

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296. https://www.jstor.org/stable/23359412

ASCILITE 2025

Future-Focused:

Educating in an Era of Continuous Change

Sun, J. C-Y., Liu, E. Y-W. & Shen, J. T-Y. (2025). Smart scaffolds, smarter learners? Analysing the cognitive outcomes of AR and chatbot feedback. In S. Barker, S. Kelly, R. McInnes & S. Dinmore (Eds.), *Future-focused: Educating in an era of continuous change*. Proceedings ASCILITE 2025. Adelaide (pp. 361-366)
https://doi.org/10.65106/apubs.2025.2676