An exploratory study of examining the effects of graduate-level learners' trust towards GAI-enhanced 3D virtual teachers on learning outcomes, behavioral transitions, and perceived cognitive load

Tinghui Wu [1], Manpreet Singh [2], Yanjie Song [2*]

[1] College of Education, Zhejiang University, Hangzhou, China

[2] Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR, China

[*] ysong@eduhk.hk

Abstract: *Recently, generative AI (GAI) has aroused great attentions in education. However, whether learners' trust concerning GAI would hinder self-regulated learning remain understudied. This study examines the effects of graduate-level learners' trust concerning GAI-enhanced 3D virtual teachers on their learning outcomes, behavioral transitions, and perceived cognitive load. Six graduate-level learners were recruited in a research-intensive university in Hong Kong, and they were grouped into high-level trust group (HT group), medium-level trust group (MT group), and low-level trust group (LT group) according to their beforehand self-reported data, respectively. Datasets were collected from final concept map products, process-oriented data, and survey data. We used a mixed approach to analyze these data. Results showed that the HT group exhibited the best learning outcomes, the most reasonable behavioral transitions, and appropriate perceived cognitive load in their learning. While learners in the LT group performed worse, the MT group's performance was average. This findings provide valuable insights for educators when integrated GAI in learning, guiding instructors and learners to use artificial intelligence to smooth their learning.*

Keywords: generative AI (GAI), trust, 3D virtual learning environments, exploratory study

1. Introduction

As virtual reality technology evolving, the three-dimensional VLE (3D-VLE) has been employed in facilitating learning performances due to the characteristics of embodiment, immerse, and society (Bogusevschi et al., 2020). In addition, generative artificial intelligence (GAI) has also gained significant attention and has shed lights on supporting self-regulated learning (Hsiao & Tang, 2024). Based on deep neural networks, GAI can be integrated into 3D-VLE to serve as a personalized assistant in self-regulated learning. Nonetheless, trust concerns regarding personal privacy and content accuracy may still hinder students' intention to use GAI-enhanced 3D-VLE (Gao et al., 2024). As previous studies stated, trust is a crucial factor influencing whether people interact with AI and their subsequent outcomes (Ueno et al., 2022). To the best of our knowledge, whether and to what extent learners' trust towards GAI affects their learning remains understudied. To fill this research gap, we conducted an exploratory study aiming to examine the effects of graduate-level learners' trust concerning GAI-enhanced 3D virtual teachers on learning outcomes, behavioral transitions, and perceived cognitive load within the context of GAI-enhanced 3D-VLE. Specifically, six graduate-level learners in a research-intensive university in Hong Kong were recruited to perform self-regulated learning based on a five-phase inquiry-based learning pedagogical design. They were grouped into high-level trust group (HT group), medium-level trust group (MT group), and low-level trust group (LT group) according to their beforehand self-reported survey data. Dataset was collected from learning-related outcomes, including their final concept maps, process-oriented data, and self-reported perceived cognitive load. To achieve the research objective, three research questions were proposed.

RQ1: Do the HT group, MT group, and LT group have different learning outcomes?

RQ2: What behavioral transitions do the HT group, MT group, and LT group demonstrate?

RQ3: What cognitive load levels do the HT group, MT group, and LT group perceive?

2. Methods

2.1. Research context, participants, and research procedure

This study referred to the IPCC Sixth Assessment Report (Kikstra et al., 2022) and chose the topic of "Atomic energy for replacing unsustainable sources" for learning. We adapted a well-crafted questionnaire survey from the study of Gulati et al. (2019) to measure their self-reported trust concerning GAI beforehand (Cronbach's α= 0.83), ultimately designating the top two highest-scoring participants as the high-level trust group (HT group), the two lowest-scoring participants as the low-level trust group (LT group), and the rest as the medium-level trust group (MT group). In this study, we employed the inquiry-based learning pedagogical design by referring to study of Song and Wen (2018), which included five phases in the GAI-enhanced 3D-VLE. It should be noted that each phase has respective GAI-enhanced 3D virtual teacher. Participants can interact with them at any time they wished. The research procedure includes three sections, i.e., the pre-survey section, the experimental section, and the post-test section.

2.2. Data collection and analysis

The study employed a mixed approach. In terms of the quantitative analysis, the concept map technique was introduced to help participants elaborate their learning understanding. This study used the quantitative content analysis (QCA) approach to assess participants' concept maps by referring to the coding scheme. Second, we used the Pupil eye-tracking device to record the process-oriented eye movement behaviors data. By referring to Zhou et al. (2024), this study focused on seven types of individual-interaction status derived from the eye movement data. Process mining (PM) was used to present learners' behavioral patterns over the inquiry-based learning. Lastly, after the experiment, all participants were required to finish a questionnaire survey concerning perceived cognitive load, adapting from the study of Andersen and Makransky (2021). In terms of the qualitative analysis, this study collected the participants' process-oriented and open-ended responses concerning their Know-Want-Learned text. From this perspective, we may determine whether participants have acquired something-related to the topic and to what extent.

3. Results

3.1. RQ1: Do the HT group, MT group, and LT group have different learning outcomes?

Results demonstrated the detailed results of the categories of each group' concept maps. First, the results indicated that the HT group's concept maps had the highest number of nodes and links (node = 29, link = 27), followed by the MT group (node = 28, link = 26) and the LT group (node = 15, link = 14). Second, in terms of the "Subjects" category, the HT group had obvious multidisciplinary perspective in utilizing the atomic energy to address climate change crisis, especially indicating the value of Social science which other groups did not mention in this process. Third, for the "Promotions" approaches, both the HT group and the MT group referred to efforts from the international, national, social, and individual perspectives, while the LT group only demonstrated the social initiatives and individual efforts. Lastly, participants in the MT group and LT group exhibited a neutral stance towards the utilization of atomic energy as an approach to alleviate climate change crisis, while one of the participants in the HT group had a positive attitude.

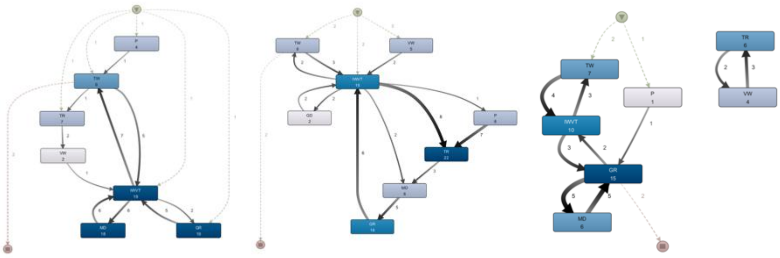
To have a better understanding of the learning outcomes among participants in different groups, we gained insights from the Know-Want-Learned text. To some extent, the qualitative text indicated that participants with a higher level of trust had a relatively positive attitude and tendency towards interacting with GAI when they developed an interest in the learning topic or encountered cognitive difficulties. Furthermore, their understanding of the learning topic was not limited to a single perspective. In contrast, participants with a lower level of trust may prefer to use traditional learning methods to solve cognitive deficiencies such as watching videos, with that they may find it challenging to summarize a comprehensive answer related to the inquiry or they did not know their answers were actually one-sided.

3.2. RQ2: What behavioral transitions do the HT group, MT group, and LT group demonstrate?

As shown in Figure 1, all the three groups exhibited significant behavioral transitions (frequency ≥ 5). The HT group exhibited such significant behavioral transitions as TW-->IWVT, IWVT-->TW, MD-->IWVT, IWVT-->MD, and GR-->IWVT (see Figure 1a), the MT group exhibited IWVT-->TR, MD-->GR, GR-->IWVT, and P-->TR (see Figure 1b), and the LT group exhibited GR-->MD and MD-->GR (see Figure 1c). In addition, the transition between VW and IWVT occurred in the HT and MT group, but did not occur in the LT group. The top-3 codes of the HT group were IWVT (19), GR (19) and MD (18), which in the MT group were TR (22), IWVT (16) and GR (16) and in the LT group were GR (15), IWVT (10) and TW (7). The code of the IWVT had the most bidirectional transitions with TW (i.e., TW-->IWVT, IWVT-->TW), GR (i.e., GR-->IWVT, IWVT-->GR), and MD (i.e., MD-->IWVT, IWVT-->MD) in the HT group. Compared to the other groups, the LT group had the simplest transitions loops (i.e., TW-->IWVT-->GR-->MD-->GR-->IWVT-->TW).

Figure 1

Process mining visualization of each group' behavioral transitions. From Left to Right: HT, MT, and LT group. Note. Interacting with GAI-enhanced virtual teacher (IWVT), text writing (TW), text reading (TR), video watching (VW), guidance sheet reading (GR), concept map drawing (MD), and atomic theories and models pairing (P).



(a) *(b)* *(c)*

3.3. RQ3: What cognitive load levels do the HT group, MT group, and LT group perceive?

The mean values of perceived cognitive load among participants in the HT group were 2.28 (SD = 1.15) and 1.78 (SD = 0.71), respectively; in contrast, the corresponding values in the MT group were 1.67 (SD = 0.88) and 3.11 (SD = 0.81), while in the LT group they were 1.22 (SD = 0.42) and 3.06 (SD = 0.78). The results indicated that participants in the HT group had a moderate perceived cognitive load during the self-regulated learning process in their inquiry-based learning in the GAI-enhanced 3D-VLE, while those in the MT group and LT group had perceptions of cognitive load that were either relatively high or low.

4. Discussion, conclusion, limitations, and future works

This exploratory study aims to examine the effects of graduate-level learners' trust on learning outcomes, behavioral transitions, and perceived cognitive load in a GAI-enhanced 3D-VLE. This study contributes to the existing literature from the perspectives of theory and pedagogy. Theoretically, the impact of learners' trust on learning outcomes resonates with the theory of self-efficacy, which suggests that an individual's self-efficacy can be enhanced through external sources. In other words, when individuals perceive a certain pedagogical support as trustworthy or professional, the messages conveyed by this support are likely to benefit their self-efficacy improvement. In this study, if learners trust GAI, they are likely to accept its positive feedback, thereby enhancing their own self-efficacy and subsequent learning-related outcomes. Emphasizing their initiative and intent for self-regulated development in inquiry-based learning, the study found that learners' trust levels successfully predicted their learning outcomes and also influenced their behavioral transitions and perceived cognitive load, further confirming the importance of trust in human-AI interaction. Pedagogically, this study highlights the role of graduate-level learners' trust concerning GAI in improving the quality of inquiry-based learning in 3D virtual learning environments. To achieve better pedagogical effects, educators should make efforts to foster learners' trust concerning GAI by designing reliable and consistent technology-assisted learning environments.

The results need to be accepted under certain limitations. First, this study only involved six graduate-level learners, which may lead to a threat to generalizing the findings. Future studies should extend the research sample. The second limitation would be the data collections method. In this study, there is a lack of other multimodal data. Future studies should highlight the importance of multimodal learning analytics in examining learners' learning process. Finally, the learning environment in this study may have certain technological shortcomings, which could also affect the learners' final outcomes and perceptions. Future research should be conducted under mature technological conditions.

Acknowledgements

We thank all the people who participated in this study.

References

Andersen, M. S., & Makransky, G. (2021). The validation and further development of a multidimensional cognitive load scale for virtual environments. Journal of Computer Assisted Learning, 37(1), 183-196. https://doi.org/10.1111/jcal.12478

Bogusevschi, D., Muntean, C., & Muntean, G. M. (2020). Teaching and learning physics using 3D virtual learning environment: A case study of combined virtual reality and virtual laboratory in secondary school. Journal of Computers in Mathematics and Science Teaching, 39(1), 5-18.

Gao, Y., Wang, Q., & Wang, X. (2024). Exploring EFL university teachers' beliefs in integrating ChatGPT and other large language models in language education: a study in China. Asia Pacific Journal of Education, 44(1), 29-44. https://doi.org/10.1080/02188791.2024.2305173

Gulati, S., Sousa, S., & Lamas, D. (2019). Design, development and evaluation of a human-computer trust scale. Behaviour & Information Technology, 38(10), 1004-1015. https://doi.org/10.1080/0144929x.2019.1656779

Hsiao, C. H., & Tang, K. Y. (2024). Beyond acceptance: an empirical investigation of technological, ethical, social, and individual determinants of GAI-enhanced learning in higher education. Education and Information Technologies, 1-26. https://doi.org/10.1007/s10639-024-13263-0

Kikstra, J. S., Nicholls, Z. R., Smith, C. J., Lewis, J., Lamboll, R. D., Byers, E., ... & Riahi, K. (2022). The IPCC Sixth Assessment Report WGIII climate assessment of mitigation pathways: from emissions to global temperatures. Geoscientific Model Development, 15(24), 9075-9109. https://doi.org/10.5194/gmd-15-9075-2022

Song, Y., & Wen, Y. (2018). Integrating various apps on BYOD (Bring Your Own Device) into seamless inquiry-based learning to enhance primary students' science learning. Journal of Science Education and Technology, 27, 165-176. https://doi.org/10.1007/s10956-017-9715-z

Ueno, T., Sawa, Y., Kim, Y., Urakami, J., Oura, H., & Seaborn, K. (2022, April). Trust in human-AI interaction: Sco** out models, measures, and methods. In CHI Conference on Human Factors in Computing Systems Extended Abstracts (pp. 1-7). https://doi.org/10.1145/3491101.3519772

Zhou, Q., Suraworachet, W., & Cukurova, M. (2024). Detecting non-verbal speech and gaze behaviours with multimodal data and computer vision to interpret effective collaborative learning interactions. Education and Information Technologies, 29(1), 1071-1098. https://doi.org/10.1007/s10639-023-12315-1